ACTIVITY AND SIGN

Grounding Mathematics Education

Edited by
Michael H.G. Hoffmann
Johannes Lenhard
Falk Seeger

Springer

Activity and Sign

Michael H.G. Hoffmann
Johannes Lenhard
Falk Seeger
(Editors)

Activity and Sign

Grounding Mathematics Education

Springer

Michael H.G. Hoffmann, Georgia Institute of Technology, U.S.A.
Johannes Lenhard, University of Bielefeld, Germany
Falk Seeger, University of Bielefeld, Germany

Library of Congress Cataloging-in-Publication Data

Activity and sign: grounding mathematics education / Michael Hoffmann, Johannes Lenhard, Falk Seeger (editors).
p. cm.
Includes bibliographical references and index.
ISBN 10: 0-387-24269-4 ISBN 13: 9870387242699 (acid-free paper)—
ISBN 10: 0-387-24270-8 ISBN 13: 9870387242705 (E-book)
1. Mathematics—Study and teaching. 2. Mathematics—Philosophy. I. Hoffmann, Michael (Michael H.G.) II. Lenhard, Johannes. II Seeger, Falk.

QA11.2.A28 2005
510'.71—dc22

2004066226

Printed in the United States of America.

9 8 7 6 5 4 3 2 1 SPIN 11375494

springeronline.com

TABLE OF CONTENTS

ACKNOWLEDGEMENT VII
MICHAEL HOFFMANN, JOHANNES LENHARD and FALK SEEGER / Grounding Mathematics Education – Michael Otte's contribution 1
MICHAEL OTTE / Mathematics, Sign and Activity 9

A. SIGN PROCESSES

PAUL ERNEST / Agency and Creativity in the Semiotics of Learning Mathematics 23
SUSANNA MARIETTI / The Semiotic Approach to Mathematical Evidence and Generalization 35
MICHAEL HOFFMANN / Signs as Means for Discoveries. Peirce and His Concepts of "Diagrammatic Reasoning," "Theorematic Deduction," "Hypostatic Abstraction", and "Theoric Transformation" 45
WILLIBALD DÖRFLER / Diagrammatic Thinking. Affordances and Constraints 57
FALK SEEGER / Notes on a Semiotically Inspired Theory of Teaching and Learning 67

B. SIGN PROCESSES IN THE MATHEMATICS CLASSROOM

MARIA G. BARTOLINI BUSSI, MARIA ALESSANDRA MARIOTTI AND FRANCA FERRI / Semiotic Mediation in the Primary School: Dürer's Glass 77
HEINZ STEINBRING / Do Mathematical Symbols Serve to Describe or Construct "Reality"? – Epistemological Problems in Teaching Mathematics in the Field of Elementary Algebra 91
NORMA PRESMEG / Metaphor and Metonymy in Processes of Semiosis in Mathematics Education 105
ANNA SIERPINSKA / On Practical and Theoretical Thinking and Other False Dichotomies in Mathematics Education 117
LUIS RADFORD / The Semiotics of the Schema. Kant, Piaget, and the Calculator 137

C. MATHEMATICS EDUCATION AS A SCIENCE

KENNETH RUTHVEN / Towards a Normal Science of Mathematics Education? 153
GUY BROUSSEAU / The Study of the Didactical Conditions of School Learning in Mathematics 159
ROLAND FISCHER / The Formal, the Social and the Subjective: Variations on a Theme of Michael Otte 169

D. CROSSING BOUNDARIES

BERND FICHTNER / Reflective learning – Problems and Questions Concerning a Current Contextualization of the Vygotskian Approach 179
RAINER BROMME / Thinking and Knowing about knowledge: A Plea for and Critical Remarks on Psychological Research Programs on Epistemological Beliefs 191
THOMAS MIES / The Cognitive Unconscious. Recalling the History of the Concept and the Problem 203

E. HISTORY OF MATHEMATICS AND MATHEMATICS EDUCATION

HANS NIELS JAHNKE / Hilbert, Weyl, and the Philosophy of Mathematics 215
THOMAS MORMANN / Mathematical Metaphors in Natorp's Neo-Kantian Epistemology and Philosophy of Science 229
KARL-NORBERT IHMIG / Newton's Program of Mathematizing Nature 241
MIRCEA RADU / Did Hermann and Robert Graßmann Contribute to the Emergence of a Formal Axiomatics? 263
GERT SCHUBRING / A Case Study in Generalization: The Notion of Multiplication 275
CIRCE MARY SILVA DA SILVA DYNNIKOV / Some German Contributions to Mathematics Research in Brazil 287

F. MAKING PHILOSOPHY OF MATHEMATICS RELEVANT

ANDREAS DRESS / Data Structures and Virtual Worlds: On the Inventiveness of Mathematics 299
HERMANN DINGES / Variables, in Particular Random Variables 305
JOHANNES LENHARD / Deduction, Perception and Modeling: The Two Peirces on the Essence of Mathematics. 313
C. ULISES MOULINES / Models of Data, Theoretical Models and Ontology: A Structuralist Perspective 325
MARCO PANZA / Some Sober Conceptions of Mathematical Truth 335
JEAN PAUL VAN BENDEGEM / Can There Be an Alternative Mathematics, Really? 349

G. CODA

ROLAND FISCHER / An Interview with Michael Otte 361

INDEX 379

ACKNOWLEDGEMENT

This book would not have been successfully completed without the work of many hands and heads and hearts. The editors would like to thank especially Christel Biere and Herta Ritsche for their work on various versions of the chapters, Jonathan Harrow and Günther Seib for their efforts to make translations from non-English speaking authors into sounding like they were from native speakers, and Roland Fischer for his brilliant idea to do an interview with Michael Otte.

MICHAEL H. G. HOFFMANN
JOHANNES LENHARD
FALK SEEGER

GROUNDING MATHEMATICS EDUCATION

Michael Otte's contribution

Mathematics education has a long past, but only a relatively short history as an institutional effort. Even though already Plato's "geometry" or medieval arithmetic books have been exceptionally "didactical" in their approach to present themselves to the reader, it was only in the 1960s that the institutionalization of mathematics education as a scientific discipline started on a larger scale. Following major changes of the role and place of science in society, the universities began at that time to change their organizational structures or added new, often interdisciplinary, organizational units focusing on applied or basic problems of research.

In a surprisingly consonant manner, this development in mathematics education was from the beginning an international one, crossing the then still existing boundaries between East and West. Without disregarding previous work of didacticians of mathematics in the 19th and early 20th century and their influence, this can certainly be viewed to a larger degree as the result of the work of the International Commissions on Mathematical Instruction (ICMI) inspired and chaired by Felix Klein working from 1908 to the twenties on a survey comparing mathematics education in many countries of the world. The International Commissions prepared the ground for a worldwide attempt to reform and revolutionize mathematics education. The reform movement of the early seventies in a way took up the international spirit of the International Commissions – particularly in the form of the ICME-conferences taking place every four years with the 1972 Exeter conference being a signal for a new beginning. Most of the contributors to this volume have played a role in this historical period – some more central, some more peripheral.

Michael Otte's scientific life and career is situated in this field of historical forces and developments. It was a founding period for systematic research and a new disciplinary self-image of mathematics education. Also in Germany, the institutionalization of basic research in mathematics education within university departments and faculties began. At that time the "new math" movement which also had been a truly international affaire just had run aground, and the Volkswagen foundation had called for a proposal to establish a central research institute in Germany to reach a deeper scientific understanding of the disastrous failure of the new math approach in German primary schools. The underlying idea was as deceptively simple as attractive: finding out the reasons and reconstruct them theoretically would also provide a

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 1 – 7.

platform or a ground to anchor a new idea of how mathematics teaching and learning should look like in the future. The original design for the Institute included five chairs furnished with opulent financial means, in terms of manpower as well as other facilities. Michael Otte was to become appointed as one of the three chairs directing the *Institut für Didaktik der Mathematik* of Bielefeld University in the next 25 or so years.[1]

Very soon after the Institute began its work, two of his important books appeared. The first one was *Mathematiker über die Mathematik* [Mathematicians on Mathematics] in 1974, and *Mathematik, die uns angeht* [Mathematics Concerning Us] in 1977. Two ensuing volumes were on mathematics and learning from text (see Keitel, Otte and Seeger 1980), and on epistemology, history and science (see Jahnke and Otte 1981). These four volumes like landmarks claim the big territory Michael Otte is covering in his scientific efforts, while the programmatic volume on *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik* [The formal, the social and the subjective: An introduction into the philosophy and pedagogy of mathematics, see Otte 1994], in a sense, is trying to sum up what had been important ideas in the work of the preceding years.

Broad coverage is, however, not the main point or intention of Michael Otte's work. We find his approach unique in his focus on grounding mathematics education as a discipline, including an emphasis on understanding the nature of disciplinarity and interdisciplinarity.

At the beginning of the 70s, it was suggestive to begin the discourse on how to ground mathematics education as a scientific discipline with the question of what exactly was the scientific nature of that enterprise. The previous roads to understanding the idea and mission of mathematics education were blocked, partly as a result of the failure of the recent reform in the primary classroom, partly as a result of a growing demand to make teacher training more up-to-date, more "scientific." The reform movement of the "new math" found its final justification in the structural identity or parallelism of the structure of cognition and the structure of mathematics and with this orientation, surprisingly enough, supported the old idea that the logic of teaching had to follow the logic of mathematical structure. It was a surprise because Piaget who was the author of the assumption of a structural similarity between the INRC-group and the "mother"-structures found in algebra, order, and topology by the Bourbaki, was so utterly successful in creating a developmental approach and a pedagogical vision on the idea of the *operative* nature of learning – an idea which in principle moved away from the emphasis on the structure of the content and underlined the importance of learning activity instead. It was too ironic that the new math reform ultimately confirmed the old idea that the structure of mathematics delivered a blueprint for the teaching of mathematics. In addition, the Bourbaki and their aim to find a common grounding for mathematics in a sense reinforced again the idea to find a common, universal grounding for all matters mathematical to be taught in schools.

Geometry seemed to be the price which had to be payed in order to reach the summit of modern mathematics in the primary and secondary classroom. "Euclid must go!" voiced by Dieudonné (1961) was the battle cry of an approach which should among other things underestimate the role of the teacher in any kind of educational reform.

Only ten years later, on the 1972 International Congress for Mathematics Education in Exeter, René Thom was actually deriding the emphasis on structure in the teaching of mathematics in a plenary speech entitled "Modern mathematics: does it exist?" He claimed that rigour was not particularly important in mathematics, and that he would prefer "meaning" rather than "rigour." Finally, he claimed that geometry, not algebra, is the natural and perhaps irreplaceable stuff mediating between language and mathematics. The juxtaposition of his ideas could not differ more from the structure-oriented "new math." With Thom's challenge the necessity became obvious to search for a conceptual and disciplinary grounding in mathematics education.

Michael Otte and his group of co-workers have, in a sense, taken up the challenge of grounding a new discipline of mathematics education very much in the sense of Thom's *dictum* that "teaching problems have to be solved fundamentally." Equally important for Michael Otte's work was Thom's famous quote: "In fact, whether one wishes it or not, all mathematical pedagogy, even if scarcely coherent, rests on a philosophy of mathematics." (Thom 1973, p. 204)

But from the beginning this form of grounding a discipline had a perspective extending beyond the pressing tasks at hand. Gaining a long-term perspective on this matter was made particularly difficult through the enormous complexity of the subject matter of mathematics education. In the centre of this development of mathematics education into a scientific discipline, two huge problems could be found – and still continue to exist. One problem has to do with finding an understanding of what "interdisciplinarity" could mean for mathematics education, this is the problem of grounding a *disciplinary identity*. The other problem has to do with the fact that the scientific study of mathematics teaching and learning has to understand and learn from that same form of *praxis* it is about to study. The grounding efforts thus have always to be twofold. An important part of the work has to be done to "justify" the autonomy and the self-image of mathematics education as a discipline. Here, the risk is enormous to just claim to *possess* autonomy in the study of that specific scientific subject, mathematics teaching and learning, which provides and grounds the autonomy. From the beginning of the work of the IDM, Michael Otte has pointed out that interdisciplinary work in the case of mathematics education can only be grounded in a deeper understanding of its subject matter. From a more comprehensive, and thus deeper, understanding of teaching and learning mathematics it becomes obvious and reasonable to study this subject from a multitude of perspectives and from the angle of diverse disciplines. The related disciplines, however, cannot simply study phenomena of mathematics education through the application of their methods without coming to understand what they are doing. Proving a deeper understanding is the task of mathematics education proper. It may even be so that after having studied things of mathematics teaching and learning in an interdisciplinary effort, the disciplines may go back home and start asking new questions about their own subject matter. The autonomy of mathematics education, thus, is not grounded in fencing its territory in and controlling access. It is paradoxically grounded in being dependent on other scientific disciplines.

The other major task in grounding mathematics education is to find a "secure base" in its attachment to the *praxis* of teaching and learning mathematics. The subject matter of mathematics is intimately bound to the multiple ways of practical experience to teach and learn mathematics. These forms of practical experience can not be understood as being "practical" in the sense of "not conscious" and "not theoretical." In the contrary, these forms of experience and knowledge are in themselves largely organized as theories. For mathematics education it seems utterly important to find access to these practical forms of theoretical knowledge, to find ways to communicate, to exchange and to open up a discourse. Again, we find a form of paradoxical relation here between theory and practice.

In the attempt to identify core ideas of mathematics education, guiding principles and essential problems, Michael Otte came up with a completely new and unusual picture of the territory of science and of the relations the disciplines have to each other and to that newly established discipline mathematics education. He insisted that mathematics education cannot survive without lively relations to the *Bezugsdisziplinen*, like sociology, education, history, psychology and so on. One of the most attractive ideas for making these things work is the idea of complementarity as can be seen in Renuka Vithal's recent book (Vithal 2003) taking up Otte's idea and elaborating further implications.

Complementarity in Michael Otte's work seems like a methodological heuristic. It is equally well known from dialectics as the coexistence and co-occurrence of contradiction, as from the Copenhagen interpretation of quantum mechanics by Niels Bohr. There we can find descriptions of the basic process either as waves or as particles, for example. Both descriptions are necessary; however, they exclude each other simultaneously. Otte has used both forms of complementarity in a lot of variations, e.g., in his 1994 book where he describes mathematics as a paradigm defining and using the complementarity of the formal/algorithmic and the historical/cultural. It is important to see that the complementarity heuristic is not satisfied with eliminating contradictions. For Michael Otte, in the contrary, the location of complementarity often seemed to indicate that a sufficient degree of analytical precision has been accomplished. Following from this, an understanding of grounding has been reached which is not guided by the idea of eliminating contradictions but finding a synthesis embracing truly contradictory forces, entities or concepts.

It is, of course, interesting to ask what will be the future of this research perspective and what will be possible applications in the pursuit of research questions in mathematics education. While we have tried in the preceding pages to sketch how Michael Otte's contribution to the advancement of mathematics education is historically situated, we would like to give in a second section an overview on a possible research perspective which could emerge as a consequence from the bundle of contributions made to this volume.

Only in recent years, Michael Otte and his group have started to work on a research perspective attempting to understand better the role of signs and representations in relation to mathematical activity and communication about mathematics. These efforts have to be seen as related to the internationally just emerging and recent approach to rephrase problems of mathematics education in terms of "semiotics", the "theory of signs." A semiotic grounding of mathematics education seems to be a promising approach.

Signs and Representations have an essential role in mathematics. It could even be said that the essence of mathematics consists in working with representations: *Mathematization* means to represent problems or facts by mathematical representational means, *calculation* is transforming such representations according to the rules of a certain system of representation, *proof* is representing a theorem as implied by other theorems within a consistent system of representation, and *generalization* is restructuring such systems of representation to include new, symbolically designated ideal objects (not implying any ontological commitments).

As it is impossible to directly grasp and experience the ideal objects and the objectivity of mathematics, we need signs and representations. Mathematical cognition is mediated by representations. The latter are on the one hand the "objects" proper of mathematical activity, and they are means to develop mathematical knowledge further on the other. This "complementarity" of means and object which Michael Otte (1994, 275 ff.) speaks of in this connection, appears to be essential for the possibility of mathematical generalization: Introducing ideal objects by means of "hypostatic abstraction" – as Peirce calls it – not only creates ever new mathematical "objects," but at the same time new "means" for the next stage of generalization. The crucial element here is the recursive character of thought which is expressed in the fact that a thought or an action is "hypostasized" to become the object of another "thought."

The semiotic dimension is in particular essential for *learning* and *teaching* mathematics. In mathematics instruction, children learn for the first time to operate exclusively with signs, they learn that the world of concrete objects and activities can be represented and understood *mathematically*, they are confronted with the problem that there are often quite different possibilities of representing the same situation, and they can see that a change of such representations often makes possible new insights. The creativity proper of mathematics results from this very fact.

Each representation is characterized by a richness of interpretation possibilities which is in principle infinite. While learning mathematics, on the one hand, involves taking over the conventional meanings of mathematical signs, it depends also on switching between different possibilities of interpretation – on seeing an "*A* as a *B*," as Michael Otte (2003, 233 ff.) says. Changing the point of view is an essential prerequisite both for learning processes and for the dynamics of theories in the sciences.

Beyond that, *communication* and *interaction* in mathematics instruction will always necessarily be mediated by signs. One learns *by* signs and *with* signs, signs are in focus of social interaction in the classroom. With regard to that, the crucial problems of learning mathematics can be formulated in semiotic terms: How shall the ideal mathematical objects and situations be *understood* if they cannot be grasped without signs on the one hand, but cannot be simply identified with certain representations either? How shall we deal with the problem, central for understanding signs, that a sign's meaning – as the only epistemologically convincing semiotics shows, the Peircean semiotics –, is always constituted by interpreting this sign, and not before?

The semiotic approach is leading to a wealth of new questions, partly reformulating old problems. The semiotic approach offers an intriguing new way to deal with complementarity, pushing forward the binary contradictory relations into the multitude of sign-object-interpretant triads. There is also the hope that the semiotic perspective could offer a new promising way to an interdisciplinary grounding of mathematics education. Here, one can think of the disciplines as interpretants of the relation between activity and sign, opening up a way to see and understand the unity in the diversity, the constancy in the variation of the various disciplines.

Even though the present volume was meant to be a *Festschrift*, a volume dedicated to an academic scholar by his colleagues, companions, friends and students on the occasion of his retirement, it turned out to be impossible to only look back at past achievements of Michael Otte.

Actually, after having delved into the writings of Charles Sanders Peirce, Michael Otte began a new chapter in his scholarly career dealing with semiotics, representation and sign processes. Accordingly, after presenting Otte's approach in a pivotal paper on "Mathematics, Sign and Activity", the first two chapters of this volume focus primarily on "Sign Processes" and "Sign Processes in Mathematics Education."

The third chapter, "Mathematics Education as a Science," takes up Otte's quest of grounding mathematics education by asking from different vantage points what it might mean to speak of mathematics education as a scientific discipline. The heading of the fourth chapter, "Crossing Boundaries," refers mainly to contributions which are located in the context of learning theory and psychology, giving an example of how distant the echo of Otte's approach is carrying. The fifth and the sixth chapter give evidence that grounding mathematics education rests on a deepened understanding of mathematics as the core reference. It seems self-evident that this understanding includes the history of mathematics and of mathematics teaching, because the view is basically an evolutionary one. While the fifth chapter links the "History of Mathematics and Mathematics Education", the sixth chapter, finally, collects different attempts of "Making Philosophy of Mathematics Relevant."

It becomes apparent already from this short overview that this volume celebrating Michael Otte as one of the founders of mathematics education as a scientific discipline is not bidding farewell but attempting to define a new beginning. The perspectives elaborated here are for the greatest part *motivated* originally by the impressing variety of Otte's thoughts and the many inputs he gave as a friend and as an always stimulating and challenging dialogue partner. The aim of the perspectives presented here is not to look back, but to find out where the research agenda might lead us in the future.

Institut für Didaktik der Mathematik, Universität Bielefeld

NOTES

[1] Heinrich Bauersfeld and Hans-Georg Steiner were the two other long-time directors at the Institut für Didaktik der Mathematik

REFERENCES

Dieudonné, J. (1961). New thinking in school mathematics. In Organisation for European Economic Co-Operation (OEEC) (Ed.), *New Thinking in School Mathematics*, 31-45. Paris: OEEC.

Otte, M. (Ed.).(1974). *Mathematiker über die Mathematik* [Mathematicians on mathematics]. Berlin: Springer-Verlag.

Otte, M., Steinbring, H., & Stowasser, R. (1977). *Mathematik die uns angeht* [Mathematik concerning us]. Gütersloh: Bertelsmann.

Otte, M. (1994). *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik* [The formal, the social and the subjective: An introduction into the philosophy and pedagogy of mathematics]. Frankfurt am Main: Suhrkamp.

Otte, M. (2003). Mathematik, Zeichen und Tätigkeit [Mathematics, sign and activity]. In M. H. G. Hoffmann (Ed.), *Mathematik verstehen – Semiotische Perspektiven* [Understanding mathematics - Semiotic perspectives], 206-241, Hildesheim: Franzbecker.

Jahnke, H. N., & Otte, M. (Eds.).(1981). *Epistemological and Social Problems of the Sciences in the Early Nineteenth Century*. Dordrecht: D. Reidel.

Keitel, C., Otte, M., & Seeger, F. (1980). *Text, Wissen, Tätigkeit. Das Schulbuch im Mathematikunterricht* [Text, knowledge, activity: The textbook in mathematics teaching]. Königstein: Scriptor.

Thom, R. (1973). Modern mathematics: Does it exist? In G. Howson (Ed.), *Developments in Mathematical Education*, 195-209. Cambridge: Cambridge University Press.

Vithal, R. (2003). *In Search of a Pedagogy of Conflict and Dialogue for Mathematics Education.* Dordrecht: Kluwer.

MICHAEL OTTE

MATHEMATICS, SIGN AND ACTIVITY

1. FUNDAMENTAL PROBLEMS

The following proposes the thesis that certain fundamental problems of mathematics that commonly appear difficult to understand can be represented more or less clearly, and hence understood, from a semiotic perspective. Thinking is not just mental, but is realized through semiotic activity. The sign process is, however, not just a continuous flow of meaning, but is interrupted and broken up by catastrophes (Thom). For us, fundamental problems are primarily the following: first, the problem of the mathematical objects; second, the paradox of proof.

Concerning the first problem, mathematical objects are not objective in the sense in which we habitually speak about existing concrete objects belonging to our empirical environment and to our everyday experience. They are not given to us in an immediate way. They are always objects of mathematical activity, and beyond that, even cultural artifacts. On the other hand, mathematics, inasmuch it is understood as an activity, does indeed possess objects of its own, and is no linguistic science based on the continuously oscillating meaning of its concepts. In his emphatic manner, Cassirer has expressed this by saying that mathematical cognition "sets in precisely at that point where the idea breaks through the cloak of language – but not in order to be from now on virtually naked, without any symbolic cover, but rather to transcend into a principally different symbol form" (Cassirer 1977, 396).

But while Cassirer understood the passage of cognition by language to be liberation from the boundaries "of intuitive representation and representability as such" (ibid., 398), Kant postulated a special form of intuition to characterize the mode of being of mathematical objects. Both authors note humanity's basic ability to distinguish between symbols and things. However, both the process of associating meanings and the Kantian construction of quantity (or of function) remain properly speaking, within the presemiotic area, as long as they are not approached from mathematical activity itself as a system to be determined. This question needs to be considered from the perspective of the two problems named above, or, in other terms, from the perspective of both the genesis and the foundation of mathematical knowledge. Everything we construct conceptually is distinct and preordained for distinction; everything we perceive is vague or continuous, and hence something general. "The man who mistook his wife for a hat" (Oliver Sacks 1970) was unable to perceive anything, requiring specific individual characteristics or *tokens* even to identify persons known to him. In some way, this man is similar to the pure mathematician who works on the basis of definitions, rather than concepts or ideas. Conceptual judgments are the cornerstones of knowledge. To know means to judge, and this, in

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 9 – 22.

turn, means to relate a particular experience to a concept (a predicate) or to a rule (a law), as there is no reasoning from particulars to particulars. Thus, to know implies, in any case, to relate a particular to a general; it means to generalize. (cf. Otte 1994, 75).

However, regardless of whether we focus on the genesis of new knowledge or on questions of foundation and proof, activity will always move between the singular and the general, between what exists and is explicitly defined on the one side, and what is vague or metaphorical on the other. In mathematics instruction, there is often the belief that the important thing is a precise language and conceptuality narrowed down in its meaning. What is so rigid and fixed, however, becomes a "private language" (Wittgenstein), and completely loses its communicative function. In short, from a semiotic perspective, the relation between indexical and iconic signs or representations becomes a crucial question of mathematical philosophy. In fact, every sign has some iconic and some indexical aspects. Take, for instance, the sentence, "It rains." Here, Peirce writes, "the icon is the mental composite photograph of all the rainy days the thinker has experienced. The index, is all whereby he distinguishes that day, as it is placed in his experience. The symbol is the mental act whereby [he] stamps that day as rainy" (CP 2.438).

Linked to the discussion of this question, as a rule, is a dispute about whether application and problem-solving on the one hand, or proving and theoretical coherence on the other, are to provide the essential orientations for mathematics. There exist, in fact, two different "cultures" in mathematics (Gowers 2000, Otte 2003)

Whereas Kant's ideas had been entirely repressed until the 1990s following the arithmetization program driven by the pure mathematics of the 19^{th} century and the concurrent "crisis of intuition," and they had fallen into oblivion until recently experiencing a certain renaissance, there remains the question what shall form the ultimate foundation of cognition: either the act of will and the concrete sign it sets, or the continuum, respectively, space, "as the primitive form of all material existence" (Cassirer 1977, 402). In semiotic terms, the conflict is between either constructing representations or recursively interlinking operative and receptive aspects of cognitive activity. This field of debate has recently seen, in mathematics education as well, an upswing of those positions emphasizing the significance of visual metaphors. Under the influence of the cognitive sciences and of the new means of cognition (computers), firstly, the belief that theory and science are also independent of our intuitions increases in importance, making one inclined to agree with the chemist H. Primas' remark that "a good theory is consistent, confirmed, and intuitable" (Primas 1981, 19). Secondly, however, the new intuition is recursively interlinked with the operative and symbolic elements of cognition, insofar as it is not directed toward a statically given world, or is the latter's reflex, but relates to the media of sign and representation themselves. Signs always have a general meaning as well, that is, they form a unity from the concrete thing and the general idea or perspective. The sign seems to represent a "contradiction" in itself, being on the one side an object – a sign after all needs to be presented as a token, that is, as a particular object or event – and, on the other side, having no existence, having only a meaning. Meanings are not things, but universals. A universal, however, has to function as a universal to be so considered. Thus, a sign is a sign only if it functions as such. A sign is not a thing, as said above, but it is not a function or bundle of various functions either. This is a crucial point not recognized by the currents of analytical philosophy and of idealist epistemology that prevail today. As a rule, these claim that reasoning is a-modal.

Now it can be said that generalization is the essential feature of the mathematical, and also that, in this, the signs, in the twofold sense already mentioned, are the object of activity. Whereas mathematical generalization consists ultimately in introducing ideal objects, the process also depends essentially on the concrete symbolic innovations, because ideas are indeed not given in themselves. It may even be said that the fundamental fact that *no* unmediated relation to reality is possible leads to the situation that theories and their languages, in the dynamics of scientific discoveries, appear in a close and indissoluble relation to one another. This is also most clearly explicated in a text of the eminent physicist and Nobel laureate Richard Feynman. Feynman compares three different forms of presenting classical mechanics, noting that they are of exactly equal value:

> Mathematically each of the three different formulations, Newton's law, the local field method and the minimum principle, gives exactly the same consequences. ... But psychologically they are very different ... because they are completely inequivalent when you are trying to guess new laws. As long as physics is incomplete, and we are trying to understand the other laws, then the different possible formulations may give clues about what might happen in other circumstances. (Feynman 1965, 53)

In this case, for instance, only Hamilton's formulation of classical dynamics permits the transition to wave theory, and this is a generalization that later became decisive in quantum theory (cf. Bohm 1977, 383f). In verification, be by logic, proof, or empirics, the double nature of the sign is often forgotten; only concrete verification and indexical signs being considered meaningful. We shall come back to this in presenting the paradox of proof. Mathematics operates with special signs, and an object is what is being designated and presented. The question what this is will then be answered in the framework of the respective mathematical activity.

Mathematical objects are at first nothing but objects of activity (e. g., problems) represented by indexical signs whose meaning unfolds in the elaboration of the structural and lawful determinations to which they are subject. Insofar, whereas mathematical objects are given to activity, they are "given as tasks" to understanding. This position is pre-established in modern axiomatics in Hilbert's sense. In this context, the question what a number is is answered simply by pointing to the arithmetic axioms: Number is everything that is embodied in a sign and that becomes an object of arithmetic activity; this activity appearing to be regulated by the axioms. The properties of the numbers (as objects) manifest themselves in the logical inferences from the axioms.

In Hilbert's axiomatics, however, all justification of the axioms is at first absent. They are little more than mere indices of mathematical objects. This view seems to suggest that the intended applications, when we interpret axiomatic structures in models, contribute something essential to the objects' contents (making it possible, for instance, to prove their consistency). It does not make sense, however, to "illustrate" the abstract! In the present case, that of numbers by concrete examples like pie charts, as in the empirical didactics of old, the important thing is rather to construct artificial imaginative worlds, by means of various types of play, for instance, wherein numbers occur as really existent, and, in this way, to apply and embody abstract structures. This means that the meaning of the concepts involved is to a considerable part fixed in the axioms.

How, however, shall application justify the applied, that is, the structures? To Kant, this seemed impossible, and for this purpose he made space and time, as forms of pure intuition, into subjective determinations. On the basis of similar ideas, modern axiomatics and logical theory of proof emerged as completely independent of any semantic reference and any ontological commitments, until Gödel's incompleteness theorems taught us to correct things, leading to re-instating the rights of the intended applications, or model theory. Even if the intensions of the mathematical concepts are in their essence established in the axioms, it does not follow that the same are the object, or describe it completely. Mathematical axioms do not present particular objects, but rather classes or types of these. An axiomatized theory, therefore, is an intensional theory; and the theory and its language becomes indistinguishable. It was, of course, impossible to return to a fundamentalism of classical character, neither to a constructivist one in Kant's sense, nor to a Platonist one in Bolzano's sense, but what resulted here was a so-to-say paradoxical linkage between condition and conditioned that can be rationally understood only from an evolutionary perspective. Concepts are both the condition and the goal of mathematical activity.

In founding the concept of number, for instance, there was an intense dispute at the end of the 19th century, respectively at the turn to the 20th, between those who held the view that the concern of arithmetics was to unfold the *contents* or intension of the concept of number – according to which the concept of number was to be erected exclusively on the notion of ordinal number – and others who held the view intending to obtain the number concept abstractly via the cardinality of sets, thus having the concept of number positioned at the beginning of all treatment of arithmetic (cf. Cassirer 1969, 67 ff). To settle this dispute reported by Cassirer, it could be said that the mathematical concept is always used attributively and referentially at the same time. Because in formal axiomatized theories, the concept's content presents itself precisely as the theory developed from the axioms and models, one would enter into conflict with Gödel's incompleteness theorem if one intended to advocate a purely intensional view of theory, and hence a purely attributive use of concepts. In case of attributive use of concepts, these appear mainly in their function in arguing and proving, whereas the focus of the referential use of concepts is mainly on the question of truth.

It has often been held in this context that Gödel's theorem shows that we are no machines. Machines, it is said, can only compute, whereas human thought is about substantial truths (cf.. Otte 1994, 221 ff., in part. 227). Webb has argued against this as follows:

> The incompleteness theorem shows that as soon as we have finished any specification of a formalism for arithmetic we can, by reflecting on that formalism (Hilbert's *"Wechselspiel"*), discover a new truth of arithmetic which not only could not have been discovered working in that formalism, but – and this is the point that is usually overlooked – which presumably *could not have been discovered independently of working with that*

> *formalism*. The very meaning of the incompleteness of a formalism is that it can be effectively used to discover new truths inaccessible to its proof-mechanism, but these new truths were presumably undiscoverable by any other method. How else would one discover the "truth" of a Gödel sentence other than by using a formalism metamathematically? We have here not only the discovery of a new way of using a formalism, but a proof of the eternal indispensability of the formalism for the discovery of new mathematical truths. (Webb 1980, 127)

Webb is certainly right here. This is already shown by elementary examples like the fact that the assumption of the real solvability of the equation $x^2 + 1 = 0$ leads to a contradiction that can be overcome by an extension of the concept of number. Webb's view, however, by no means implies that we human beings did not think intuitively, or that we did not need intuition; it is only directed against that classical reductionist concept of intuition. Our intuitions, just like the reasoning of computers, are only means of the activity of constructing representations.

In his famous Paris lecture on future problems of mathematics, Hilbert emphasized that arithmetic consists of nothing but the explication of mathematical intuitions. If we relate intuition, in contrast to the classical concept of application, to the signs and diagrams on which mathematicians base their activity, and which continuously accompany this activity, this process of explication will never come to its end. It can never be closed, because new intuitions are given, with new representations and diagrams. As Peirce says "a great distinguishing property of the icon is that by the direct observation of it other truths concerning its object can be discovered than those which suffice to determine its construction" (CP 1.179, see, also, Peirce NEM III, 749). Here again, the distinctive character of the icon is indicated, namely, that it is the only sign by which we can enlarge our knowledge. Under all circumstances "each Icon partakes of some more or less overt character of its Object" (CP 4.531). This partaking can be of a complex sort, and need not be completely determinable. This is nothing but a pointer to that which is implicitly and intuitively given and observable in the icon and lends itself to formulation after the fact in logical relationships and axioms.

Castonguay, who, like us, is in favor of a dualist theory of meaning according to which meaning is "an inseparable tissue of convention and fact," speaks of a "heuristic component" of mathematical meaning that represents a source of inspiration

> for the positing of relations between variously (and possibly referentially perceived mathematical concepts or entities, relations which may eventually crystallize, through more exact formulation and deductive corroboration, into objective relations of entailment between linguistically expressed concepts (Castonguay 1972, 3).

In our way of speaking, a mathematical theory's heuristic component would probably be the totality of intended applications or possible models, inasmuch it fills, in the absence "of an authentic referential pole for meaning in mathematics," the role in the concept of meaning that is complementary to intension.

Because we must conceive of intuition semiotically as of a means of formal inference, seems appropriate to treat this once more in more detail; and this is why we shall specifically treat mathematical deduction, as in the last section.

2. FOUNDING AND PROVING

This brings us to the second question in this essay, to the paradox of proof. It can be formulated as follows: On the one hand, a proof can prove something only if the knowledge concerned possesses a firm tautological structure, and if proving ultimately consists in sequencing immediate identities or equalities. In doing so, proof, on the other hand, reduces the knowledge to be conveyed to the knowledge already present in the recipient, and it is not seen how new knowledge can be created in the learner (cf. Otte/Bromme 1978, 20f.). If proof then is meant to produce knowledge – and mathematical knowledge cannot be obtained in any other way – it cannot be a tautological process that exerts a material or causal coercion, but must be a semiotic process instead. Proof does not characterize an interaction between reactive systems, but rather one between cognitive systems. Proof and cognition, then, require not only that general rules and procedures of proof, or logical arguments, be stated, but also that a certain perspective or idea be appropriated as one's own. Finally, as a third element, proof requires not only that a sign in the twofold sense be developed but also that it be applied to a situation of which the proof's recipient is perfectly aware. Proof thus always implies generalization, and a verification or application. The problem has been presented by Lewis Carroll in a most beautiful text: What the tortoise said to Achilles (reprinted in Hofstadter 1985, 47ff).

> Achilles and the tortoise talk about Euclid's elements and about the proofs encountered there. One of the examples they consider is the following:
>
> A) If two things are equal to a third, they are equal to one another.
> B) The two sides of this triangle are equal to another.
> Z) The two sides of this triangle are equal to another.
>
> Every reader of Euclid will probably admit that Z follows logically from A and B, so that everybody who accepts A and B must accept Z as true, Achilles claims. But in order to compel the tortoise accept this mode of inference, and in particular to accept Z, if it accepts A and B, he has no other option than to write down precisely this claim as a new rule.
>
> C) If A and B are true, Z must be true. And, further:
> D) If A and B and C are true, Z must be true, etc.

This is where it becomes clear that the infinite regression can only be overcome if the rule (or the idea, respectively the concept, as a scheme of action) were to be identical with its own application. This is how intuitive reasoning is traditionally characterized. In the famous *heureka!* or aha moment of intuitive insight, the fact presents itself in immediate identity with the establishment of its truth.

Hence, it is seen now that verification is threatened by the same regression as is generalization. The sentences "P" and "P is true" refer to the same judgment. They are different sentences, however; and Bolzano used this fact to construct an infinite totality of sentences. At this point, the impossibility of seeing the truth from the sentence itself, or of establishing a criterion of truth linguistically, leads to the tendency of repeating the predicate "... is true" ever more emphatically. We cannot define truth in a way that would permit us to decide about the truth or falseness of a sentence immediately upon its presentation. The sentence is a sign, too. This, however, is obviously a further sketchy expression of Gödel's incompleteness theorem. While truth is due to sentences, it cannot be established linguistically, but rather belongs to relations between language and the world that cannot be characterized by speech. Truth is unprovable and cannot be defined. This had already been stressed by Kant (Kant, *Critique of Pure Reason*, B 83). The same is true for the concept of existence (a finding also claimed by Kant: B 626).

It might be concluded from what has been said on intuition, a conclusion often drawn by referring to Gödel's incompleteness theorem, that intuition, in Kant's or Descartes' sense, should ultimately be the last instance of decision. Intuition and emotionality are doubtlessly of decisive import for the activity of cognition, which would not proceed at all without them. Our intuitions, however, are very misleading, and one may even claim that, without experience, they would err in the majority of cases. As we see from Carroll's parable, factual information alone, on the other hand, is not sufficient to correct this. One may indeed understand the above comment on Caroll's parable as a hint that a position purely aligned to intuitive truth and a deductive view, obligated merely to formal consistency, are identical. This once again concerns a complementarity of the attributive and referential use of mathematical concepts.

Mathematical cognitions, too, even if ultimately constituted by formal proofs, are dependent on experience, and mathematics thus must offer an opportunity for experience. Experience is obtained by the natural scientist, just as in everyday life, from experimental practice. In mathematics, there are no experiments, but mental experiments. Mental experiments, again, are signs, and not just internal imaginations. They are bound to certain concrete representations or models and thus permit certain experiences to be had when dealing with these.

Mental experiments have again and again played a decisive role in the development of physics or of chemistry. But, as Thomas Kuhn says, it is

> by no means clear how they could ever have significant effects. Often ... they have to do with relationships which have not been examined in the laboratory. Sometimes, ... they assume situations which cannot be completely studied at all and which need not even occur in nature. ... The main problems in connection with mental experiments can be formulated as a number of questions. Firstly: The situation imagined in a mental experiment must obviously not be completely arbitrary. (Kuhn 1977, 327)

Secondly, one must ask oneself how new cognitions of nature can emerge from the mental experiment if it does not produce any new information at all, as a real experiment does.

> Lastly, the third and shortest question: What new cognitions can be obtained in this way? (Thomas Kuhn, ibid.)

We are unable to present Kuhn's very differentiated and manifold answer to these questions in detail here. One thing, however, should be clear: Mental experiments are situations in which general rules and cognitions must be applied to particular constellations, and this is precisely where experience is obtained. Experience always means to experience a reality's resistance, and the latter can already come about by it not being clear which of two possible contradictory rules should be applied here. Experience indeed has to do with the interchange of general representations and individual perceptions, as well as with their objectiveness. In semiotic terms, indices are signs compelling us to make certain determinations, whereas icons relate to our ideas. The result thus is that mathematics cannot only operate with concepts, but must also use iconic and indexical signs.

The foundation of mathematical knowledge (i. e., proof) finds itself ultimately confronted with the same problematic as the genesis of mathematical cognition. This problematic consists in mathematical cognition being an activity, and that the point thus is not just to have an idea or to know a rule, but always to apply ideas, concepts, rules, and guesses to specific situations. This sameness of problematic is also illustrated by Plato's paradox of learning. Plato had formulated his argument in *Menon* from the perspective of the not-yet-knowing: how can one seek something when one does not even know what it is (Menon 80 d ff). The paradox is that if one knows what one is searching for, one no longer needs to search for it, and that if one does not know, the search becomes impossible.

This presentation, however, is incomplete inasmuch as the point is to imagine what one seeks. But, while such an imagination is a necessary, it is by no means a sufficient condition for what is sought. It may well be that one has the right idea for conducting a proof, but does not know exactly how one is to apply it. It is known, for instance, that the theorem about Euler's line in the triangle, because it contains only projective determinations in its claim, must lend itself to be proved from the axioms and theorems of projective geometry. Possibly, one even knows that this is a special case of Desargues' theorem (respectively its inversion). One does not know, however, how one is to apply this knowledge in the present case and to the given constellation. A theorem's premises are both an indexical hint, and a presentation of the intended situation, albeit a very incomplete and one-sided one. Conversely: once one has an idea of proof at one's disposal, one might ask what can conducting the proof then still add to this. *Nothing*, is the answer, if it only repeats the idea without sophisticating and specifying it. The first idea can never be completely right; otherwise the problem would be solved. Ideas are "pure," and hence one-sided and not adapted to reality at all.

Accordingly, one may also advocate the thesis that only a very limited role is due to intuition, or to the intuitive idea, or to the *heureka*. Intuition is as deceptive as it is important. It is always directed toward observing a representation, and it discovers something in it. What this is will only be shown in a new, transformed representation, and can only be shown effectively in this way. The idea, we shall claim, always is the idea or basis of a representation. Ideas are things possible; they have a meaning, but no factual existence. What is possible cannot be identified by the totality of its representations, because this totality actually does not exist a priori. But the possible cannot be separated from the totality of representations and understood as mere intuition or Platonic idea either.

From what has been said, we now obtain: Reality that is to be understood must be represented. A things' idea or essence thus is itself the essence of a representation of the thing. A representation's essence is again a transformed representation, for to interpret is to represent. Whereas the essence is something relative, something mediated, it is also objective. This objectivity shows in the continuum of all representations, for objectivity can be ascribed in an intelligible and fruitful sense only to a general object. The singular will, at best, operate as a constraint.

What has been formulated here as the *central thesis of a semiotic epistemology of mathematics* – that is to say, that signs are meaningful because they represent processes of interaction between the general and the particular – has appeared in manifold forms in the history of philosophy. The underlying problem shows clearly how post-Kantian idealism (i. e., Fichte, Schelling, Schleiermacher, and Hegel, too) treated the Kantian concept of intuition and the duality of concept and intuition that had been fundamental for Kant's epistemology. One spoke in this connection of an intellectual intuition, of a unity of construction and constructed, of an intuition of intuition, or also of the hermeneutic circle of interpretation that is cognitively based on the fact that the fundamental concepts and basic ideas are both the foundation and the result of interpretation or cognition.

It is correct that every new information, every new knowledge, must be related, just like every new idea, to the system of the cognitions and information already in existence, or – in psychological terms – must be integrable into the developed cognitive structure. To have experiences, to exploit information, to head toward goals, or to confront problems requires a frame, a perspective, and idea under which all this can be executed. If really new knowledge is to be acquired, however, this perspective, or this idea, must, on the other hand, be furnished at least partly by the new content itself. If something new is to be introduced into thought, this new thing must to a certain degree itself provide the perspective and the foundation of its development in reasoning. The theoretical concept must, so to say, deliver the basis of its own explanation. If this were not possible, it would be difficult to understand how something new can be learned, because the sole remaining measure would be to see whether the new ideas and the new concepts are similar to the old or not. This is nothing but a variation of the two paradoxes we have formulated, namely, the paradox of proof on the one side, and Plato's paradox of learning on the other.

3. ANALOGY, CONTINUITY, AND GENERALIZATION

The reports of scientists on their own work again and again stress the role of perceiving analogies and structural similarities as a means to obtain new things. The role of the concrete representations, respectively, the fact that one and the same idea must also be explicated and represented in a form as varied as possible, is seen more rarely (compare, in contrast, our above quote of Feynmann). In this respect, the dispute between the more traditional psychology of association from Hume across Helmholtz and Poincaré up to Ziehen (1914 <1902>) and Ebbinghaus (1908) on the one hand, and the Gestalt psychology of the Würzburg School (Bühler or Wertheimer) on the other is very informative and revealing. Whereas the psychology of association stresses the importance of the continuity principle on the basis of Hume's distinction between *associations by similarity* versus *associations by contiguity*, Gestalt psychology points out the determining character of the problem situation.

Ziehen describes the principle of continuity, which he calls the "neighbourhood principle," as follows:

> Every representation calls forth as its successor either a representation that is similar to it with regard to contents, or one with which it itself, or with whose basic sensation its own basic sensation has often appeared simultaneously. The association of the first order is called an internal association, that of the second order also an external one. (Ziehen 1914, 309)

As an essential element going beyond that, Gestalt psychology has considerably added the hint at the determining tendencies that emanate from a task to be solved, showing that a task presented or a problem considerably accelerates the course of all processes of reasoning. Here again, we encounter a variation of Plato's paradox. It is well known that it is much easier to prove a theorem that one knows to be true than one that is entirely unclear. It is also simpler to solve a task that possesses a solution. Moreover, N. Ach has pointed that the determining tendencies that emanate from the task are sometimes more important for the course taken by the representation than the external stimuli and the associative connections. The determining tendencies emanating from the task create new associations between the representations (cf. Ach 1905). This, however, is obviously mediated by the respective representation of the problem or the problem situation. It may thus probably be said that the decisive aspect in the transition from the psychology of association to Gestalt theory was to extend the understanding of the principle of continuity effected by liberating the latter from interpretations that confined it to mere representations or perceptions.

In empirical contexts, we observe certain regularities, like distributions of values measured, and we seek the principle that generates them. This is not attainable in a purely inductive way, but also requires, alongside the data, certain ideas on the form of the laws sought. Eventually, what was first assumed only hypothetically must be verified. In mathematical contexts, particularly in arithmetic and algebra, we are familiar with the transformations or constructive mechanisms, and we look for patterns or regularities in what is produced. Circle, ellipse, and parabola, for instance, are all second-order algebraic curves, and, as such, species of a genus, and this subsequently provides the basis for using the principle of continuity as a device of proof just as Leibniz used it, or later, to a larger degree, Poncelet. Here, the principle of continuity is at the service of a relational reasoning, of an analytic ideal of cognition according to which the objects are not seen in their individuality and in their distinction, but rather in their similarity and their connection. Now it can be said that the empirical sciences aim at regularities or laws as well, and it may even be claimed that they obtain the same actively inasmuch as they conduct experiments, leading Peirce, in an early manuscript of 1878, to designate the foundation of synthetic conclusions as follows: "Experiences whose conditions are the same will have the same general characters" (Peirce CP 2.692). Leibniz intended nearly the same thing when he wrote in 1687:

> If, in the series of the given quantities, two cases approximate one another continuously so that one transcends into the other, necessarily the same must occur in the series of derived or dependent quantities. (Leibniz HS I, 62)

Now I do not simply insert the generating principle into an experimental context, for otherwise, experimental research in its entirety would make no sense at all, as it would amount to a mere self-affirmation, but I observe a lawful regularity whose causes I seek. In the context of arithmetic, number theory, or algebra, in contrast, I have a formula and seek to describe what is generated by it.

In any case, we should therefore understand the representation of the respective task as a sign, and apply the principle of continuity, or the principle of neighborhood, to the relations between signs, rather than to mere intentions. This is what we had already explicated in our proposal to understand intuition as a means of semiotic activity, that is, to interpret the idea as a basis or essence of a sign or representation. The process of solving a problem thus consists in a gradual correction of ideas or generalization involving ever new concrete representations. It is seen here again that intuition is expressed in applying a general argumentation to a particular constellation, that is, in constructing a representation, and that this representation changes the intuition. This is why Peirce calls perceptual judgments an extreme case of abductive reasoning:

> The abductive suggestion comes to us like a flash. It is an act of *insight*, although of extremely fallible insight. It is true that the different elements of the hypothesis were in our minds before; but it is the idea of putting together what we had never before dreamed of putting together which flashes the new suggestion before our contemplation. On its side, the perceptive judgment is the result of a process, ... If we were to subject this subconscious process to logical analysis, we should find that it terminated in what that analysis would represent as an abductive inference, resting on the result of a similar process which a similar logical analysis would represent to be terminated by a similar abductive inference, and so on ad infinitum. This analysis would be precisely analogous to that which the sophism of Achilles and the Tortoise applies to the chase of the Tortoise by Achilles, and it would fail to represent the real process for the same reason. Namely, just as Achilles does not have to make the series of distinct endeavors which he is represented as making, so this process of forming the perceptual judgment, because it is sub-conscious and so not amenable to logical criticism, does not have to make separate acts of inference, but performs its act in one continuous process. (CP 5.181)

What we propose here is to, nonetheless, decompose this process, to interrupt the continuum by intermediate stages, to relativize the flash character and the immediateness of insight in order to make something teachable and learnable, in other words, communicable, that otherwise would seem to evade every communication. Intuition is not eliminated by this, but it is deprived of its quasi paradoxical character, inasmuch as many things are seen or perceived more easily than others. A bold hypothesis or a conclusion drawn from afar is decomposed into stages just as described already by Aristotle in his *Analytica posteriora* II 23 as the compression of the mean (cf. for this Detel 1993, 302 ff.). To prove the sum of angles' theorem in the triangle, for instance, I draw a straight line parallel to the base through the triangle's top, and then conclude that the theorem results from this by saying if such a parallel straight line is given, then ... etc., etc. This drawn diagram is precisely such a middle element suggested by the principle of continuity.

We have seen that all the problems named – object problem, problem of proof, problem of learning and of generalization – show the same basic structure, and that we encounter this problem structure already in the instant in which we try to explain the connection between perception and cognition, or the emergence of the perceptive judgment as a process in which action (representation) and reception (intuition) are becoming recursively intertwined. Peirce himself observed in one of his later manuscripts on the essence of pragmatism:

> I do not think it is possible fully to comprehend the problem of the merits of pragmatism without recognizing these three truths:
>
> 1. that there are no conceptions which are not given to us in perceptual judgments, so that we may say that all our ideas are perceptual ideas. This sounds like sensationalism but in order to maintain this position it is necessary to recognize,
> 2. that perceptual judgments contain elements of generality; so that Thirdness is directly perceived; and finally I think it of great importance to recognize
> 3. that the Abductive faculty, whereby we divine the secrets of nature is, as we may say, a shading off, a gradation of that which in its highest perfection we call perception. (Peirce MS 316)

The essential thing in this transition from perception to abduction seems to be the generalization of the singular and factual to the general connection as it is represented in an analogy or in a metaphor. In the case of a metaphor, the basis of the relation of similarity must be found first of all. What is more important here is the sameness of genus or of family, which implies a transition from the purely empirical to the theoretical. To begin with a simple example: parabola and catenary are empirically so similar that Galileo still took them to be the same; the difference being elaborated only by Huyghens. On the other hand, circle, ellipse, and parabola are of the same genus of family, but empirically quite dissimilar. In geometry, they are nonetheless considered to belong together.

Conceiving of the principle of continuity in its relation to a sameness of genus or family type marks an essential element of the scientific revolution of the 17th to the 19th century. This is when reasoning and intuition began to pass from the things to the laws determining them. The laws themselves are deemed to be anchored in logic, and in God's mind. Leibniz, albeit always aligned to the problem of individuation in his quest for cognition, considered pure mathematics to be an analytic science concerned with the general concepts of genus and with the laws that are to be valid in all possible worlds. These laws or relational structures, however, also determine the reasoning in analogies or metaphors, that is, they provide the aspect under which reasoning approximates the existing world. And this is the very purpose for which the principle of continuity has been conceived. In order to use the metaphorical as part of a mathematical-natural science methodology, one must draw on the principle of continuity in the sense of that Aristotelian quest for middle elements or intermediate steps. Peirce describes the method extensively:

> When a naturalist wishes to study a species, he collects a considerable number of specimens more or less similar. In contemplating them, he observes certain ones which are more or less alike in some particular respect. They all have, for instance, a certain S-shaped marking. He observes that they are not precisely alike, in this respect; the S has not precisely the same shape, but the differences are such as to lead him to believe that forms could be found intermediate between any two of those he possesses. He, now, finds other forms apparently quite dissimilar – say a marking in the form of a C – and the question is, whether he can find intermediate ones which will connect these latter with the others. This he often succeeds in doing in cases where it would at first be thought impossible; whereas, he sometimes finds those which differ, at first glance, much less, to be separated in Nature by the non-occurrence of intermediaries. In this way, he builds up from the study of Nature a new general conception of the character in

> question. He obtains, for example, an idea of a leaf which includes every part of the flower, and an idea of a vertebra which includes the skull. I surely need not say much to show what a logical engine is here. It is the essence of the method of the naturalist. How he applies it first to one character, and then to another, and finally obtains a notion of a species of animals, the differences between whose members, however great, are confined within limits, is a matter which does not here concern us. The whole method of classification must be considered later; but, at present, I only desire to point out that it is by taking advantage of the idea of continuity, or the passage from one form to another by insensible degrees, that the naturalist builds his conceptions. Now, the naturalists are the great builders of conceptions; there is no other branch of science where so much of this work is done as in theirs; and we must, in great measure, take them for our teachers in this important part of logic. And it will be found everywhere that the idea of continuity is a powerful aid to the formation of true and fruitful conceptions. By means of it, the greatest differences are broken down and resolved into differences of degree, and the incessant application of it is of the greatest value in broadening our conceptions. (CP 2.646)

Mathematical activity, and this is the thesis proposed here, is about transferring these methods to the world of the representations of mathematical facts.

Institut für Didaktik der Mathematik, Universität Bielefeld

NOTE

Citations from German editions were translated by Günter Seib.

REFERENCES

Ach, N. (1905). *Über die Willenstätigkeit und das Denken.* Göttingen.

Bohm, D. (1977). *The Structure of Scientific Theories.* Urbana: Univ. of Illinois Press

Cassirer, Ernst (1969). *The problem of knowledge: philosophy, science, and history since Hegel* / by Ernst Cassirer. Transl. by William H. Woglom and Charles W. Hendel. New Haven, London: Yale Univ. Press

Cassirer, E. (1977). *Philosophie der symbolischen Formen,* 3. Teil, 7. ed. Darmstadt: Wissenschaftliche Buchgesellschaft.

Castonguay, C. (1972). *Meaning and Existence in Mathematics.* Wien: Springer.

Detel, W. (1993). Einleitung. Aristoteles. Analytica Posteriora. *Werke in deutscher Übersetzung* vol. 3, part 2. Berlin: Akademie-Verl. 1, 103-438.

Ebbinghaus, H. (1908). Abriss der Psychologie. Leipzig: Veit.

Schlagwörter (RSWK). *Psychologie / Einführung.*

Feynman, R. P. (1965). *The character of physical law.* Cambridge, MA: MIT Press.

Gowers, W. T. (2000). The Two Cultures of Mathematics. In V. Arnold et al. (Eds.), *Mathematics: Frontiers and Perspectives.* AMS.

Hofstadter, D. (1985). *Gödel Escher Bach.* Stuttgart: Klett-Cotta.

Kant, Immanuel: *Critique of pure reason.* Cambridge Univ. Pr.

Kuhn, T. (1977). *The essential tension. Selected Studies in Scientific Tradition and Change.* Chicago: University of Chicago Press.

Leibniz, G. W. (HS) (1996). *Hauptschriften zur Grundlegung der Philosophie,* ed. E. Cassirer, 2 vol. Hamburg: Meiner.

Otte, M. (1994). *Das Formale, das Soziale und das Subjektive.* Frankfurt am Main: Suhrkamp.

Otte, M., & Bromme, R. (1978). Der Begriff und die Probleme seiner Anwendung. In Bloch u. a. (Eds.) *Grundlagenkonzepte der Wissenschaftskritik als unterichtsstrukturierende Momente.* IPN-Arbeitsberichte vol. 29. Kiel.

Otte, M. (2003). Construction of Existence. In *Anais da SBHM.* Rio Claro-SP.

Peirce (CP) (1958). *Collected Papers of Charles Sanders Peirce* (Vol. I-VI ed. Ch. Hartshorne & P. Weiss, 1931 – 1935, Vol. VII-VIII ed. A. W. Burks, cited acc. To volume and paragraph). Cambridge, MA: Harvard UP.

Peirce (MS) (1967). *The Charles S. Peirce Papers, Manuscript Collection in the Houghton Library. Harvard University* (counting acc. to R. S. Robin, Annotated Catalogue of the Papers of Charles S. Peirce. Worcester, Mass.: The University of Massachusetts Press).

Peirce (NEM) (1976). *The New Elements of Mathematics by Charles S. Peirce,* Vol. I-IV (ed. C. Eisele). The Hague-Paris/Atlantic Highlands, NJ: Mouton/Humanities Press.

Sacks, O. (1970). *The Man Who Mistook His Wife for a Hat.* New York: Harper Perennial.

Webb, J. C. (1980). *Mechanism, Mentalism, and Metamathematics: An Essay on Finitism.* Dordrecht: Reidel

Ziehen, T. (1914). *Leitfaden der physiologischen Psychologie* [1902]. Jena.

PAUL ERNEST

AGENCY AND CREATIVITY IN THE SEMIOTICS OF LEARNING MATHEMATICS

Abstract. Semiotics provides a way of conceptualising the teaching and learning of mathematics driven by a primary focus on signs and sign use. It considers patterns of sign use and production, and the contexts and social rules underlying sign use. It attends to agency in the learner's personal appropriation of signs and the meaning structures embodying the relationships between signs. Learner agency is manifested in communicative activity involving sign 'reception' (listening, reading) and sign production (speaking, writing, sketching). It is most marked in individual creativity in sign use, which is manifested at all levels in schooling and in the activities of the working mathematician.

Key words: Semiotics, semiotic systems, mathematical activity, agency, creativity, teaching and learning of mathematics, sign transformations, appropriation, conventionalization .

INTRODUCTION

A semiotic perspective of mathematical activity provides a way of conceptualising the teaching and learning of mathematics driven by a primary focus on signs and sign use. In providing this perspective it offers an alternative to any psychological perspective that focuses exclusively on mental structures and functions. It also rejects any straightforwardly assessment or performance focussed perspective concerned only with student behaviours. Instead it offers a novel synthesis that encompasses but also transcends these two types of perspective, driven by a primary focus on signs and sign use in mathematics. Beyond the traditional psychological focus on mental structures and functions it considers the personal appropriation of signs and the underlying meaning structures embodying relationships between signs. Beyond behavioural performance it is concerned with patterns of sign use and production, including individual creativity in sign use, and the underlying social rules and contexts of sign use. Thus a semiotic approach draws together the individual and social dimensions of mathematical activity which are understood as mutually dependent and constitutive aspects of the teaching and learning of mathematics.

The primary focus in a semiotic perspective is on communicative activity in mathematics utilising signs. This involves both sign 'reception' and comprehension via listening and reading, and sign production via speaking and writing or sketching. While these are conceptually distinct, in actualisation these two activities overlap and are mutually shaping in conversations (semiotic exchanges between persons within a social context). Sign production or utterance is primarily an agentic and often a creative act. For the speaker has to choose and construct texts to utter on the

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 23 – 34.

basis of their appropriated and learned repertoire of signs. In so doing, speakers are taking risks in exposing themselves to external correction and evaluation against the rules of appropriate utterances. 'Text' denotes more than a piece of writing here. As is widespread in semiotics, it is a compound sign made up of constituent signs, and can be uttered or offered in a conversation in many ways. It may be spoken, written, drawn, represented electronically and may include gestures, letters, mathematical symbols, diagrams, tables, etc., or some combination.

Texts, signs and their use need to be understood as part of more complex systems. First of all, sign use is always socially located and is a part of social and historical practice. In Wittgensteinian (1953) terms sign use comprises 'language games' embedded in social 'forms of life' (Ernest 1998). Second, signs are never used individually. Signs are always manifested as part of semiotic systems, with reference implicitly or explicitly, to other signs. The term semiotic system is used here to comprise the following three components:

1. A set of signs, the tokens of which might possibly be uttered, spoken, written, drawn, or encoded electronically.
2. A set of relationships between these signs based on an underlying meaning structure (or structures) embodying these relationships,
3. A set of rules of sign production, for producing or uttering both atomic (single) and molecular (compound) signs. (These rules are in most cases implicit, acquired by 'case law'.)

The social and historical embedding of semiotic systems concerns both their structural dimension (Saussure's *langue*) and in their functional role (Saussure's *parole*). These dimensions, while theoretical separable, are woven together in historico-social practice. The evolution of semiotic systems can be examined historically in terms both of these dimensions. Such developmental processes result in knowledge systems, such as school mathematics, that provides the underlying structure to the planned learning environments for students. However, just as semiotic systems change and develop over history, so too the semiotic systems mastered by learners develop and change over the course of their learning careers, becoming more elaborated and providing the basis for more complex and abstract systems. Mastering these enlarging semiotic knowledge systems constitutes learning. This feature is also a basis for a some learning difficulties, for the semiotic systems mastered by learners are never static. As they near mastery of a particular system, the teacher extends the system with new signs, relationships, rules or applications. For example, for a young child mastering elementary calculation $3 - 4$ is impossible. But later $3 - 4 = -1$. Similarly 3 divided by 4 (3/4) is at first impossible. Later it is not only possible but ¾ names the answer. These, together with more complex changes in the rules that occur (are imposed) as semiotic systems are extended, and the problems they cause, have been named epistemological obstacles (Bachelard 1951, Sierpinska 1987). Thus a structural view of semiotic systems can provide only a freeze-frame picture of a growing and life-like entity. Indeed in practice it is difficult to clearly distinguish and demarcate the range of semiotic systems encountered in school mathematics because of their growth and their mutually constitutive inter-relationships.

Successful mathematical activity in school requires at least partial mastery of some of the semiotic systems involved in schooling at the appropriate level. A number of different but interrelated and overlapping semiotic systems are important in learning, and mastery of the following systems usually constitute significant stages in learning school mathematics the way it is currently organised: Numbers and counting; Numerical computation, Fractions (rational numbers) and their Operations, Elementary linear algebra (solving equations), Analysis (calculus) and Abstract (axiomatic) group theory. Clearly, the semiotic systems chosen from university mathematics are more arbitrary than those chosen from the earlier years of schooling, in the sense that some university students could study mathematics but not analysis or abstract groups. The topic areas could be identified differently, but nevertheless they constitute a central part of taught mathematics straddling the years from kindergarten to university study. Naturally, there are further overlapping semiotic systems in school mathematics learned in parallel with these (e. g., geometry and probability) and even from this perspective the mathematics curriculum could be 'cut up' into different semiotic systems.

Semiotic systems are incorporated in all human communicative activities, and are inextricably woven into the fabric of all social activities and institutions. So the question can be posed: what is unique about their nature and deployment within school mathematics? A number of mathematically specific systems (topic areas) with dedicated sign systems, meanings and rules of use are mentioned above. But more than this, I want to suggest that there is an underlying characteristic shared by most if not all semiotic systems in school mathematics and more widely, by mathematics itself. In brief, my claim is that these systems are fundamentally sequential and procedural. In a sense this is an empty or superficial description, because at the heart of mathematics are its meanings, its purpose as a device for meaning-making, and this is driven by its social and human aims and context. But to treat these further issues in addition to its means of signification requires an in-depth discussion of historical and philosophical issues that are not only too complex and elaborate for the space here, but which are also clouded by centuries of metaphysical and ideological preconceptions about mathematics. However, the view of mathematical signs as sequential and procedural in nature of helps explain a well-known pathological outcome of education in which learners only appropriate surface characteristics without managing to transform them into part of a larger system of personal meanings.

My claim is that texts in the semiotic systems of mathematics are representative of sequences of actions (physical or textual), and the signs stand for steps (the individual results of procedures), actions on these steps (the procedures themselves), sequences of steps linked by procedures, and collections of these entities. My description includes the so-called entities involved themselves, whereas in mathematics and school mathematics we have almost nothing but the signs that stand for these steps, procedures and collections. Physical actions (such as enumerating a sequence of ordinals in counting a collection of tangible objects: 1, 2, 3, 4, 5, ...) which have an extended temporal existence become rapidly replaced by spatially extended sequences of signs, which themselves can become embodied into truncated 'super'-signs (cardinal numbers in the example). Such a replacement of process signs by product signs in mathematics (the reification of constructions) is discussed in the philosophy of mathematics (Machover 1983, Davis 1974, Ernest 1998), and in mathematics education (Dubinsky 1988, Ernest 1991, Sfard 1993). In linguistics, there is a well known parallel in the process of nominalisation, in which verbs designating actions and activities are transmuted into nouns, which representing the names of entities (Chomsky 1965). What is unique in mathematics is the great height to which these towers of abstraction rise, with each level reifying actions on lower level entities and processes into new entities. My claim is that all that there is (above the very basic ground floor of physical actions) is signs or names, and actions upon them.

The claim that mathematics is fundamentally procedural is lent some support by the philosophical position of intuitionism, which regards the objects and sentences of mathematics as representing constructions (Troelstra and van Dalen 1988, Heyting 1956). Although the intuitionist philosophy only has a minority of mathematicians as adherents, one of its achievements has been to translate a very significant part of mathematics including the content of all elementary (i. e., school) mathematics and much of advanced mathematics into transparently constructive (i. e. procedural) form Bishop (1967). Of course the Intuitionists do not accept that virtually all is signs or actions on them, for they posit some transcendent subjective (but universal) domain of meanings. Supporting, and in large part inspiring my account, Rotman's (1993) semiotic theory of mathematics also interprets mathematical inscriptions as recipes, instructions, or claims about the outcomes of procedures, without the need to posit entities beyond our social and cultural constructions.

What I am claiming (fully aware of the ontological implications) is that the so-called objects of mathematics are themselves the products of sequential actions and procedures. However, the tendentious nature of this statement is neutralised by the adoption of a semiotic perspective, rather than a philosophical one (for the moment). For my universe of discourse here is populated primarily by signs (and the persons who use them) rather than any abstract objects of mathematics.

A valuable feature of semiotics is that it is neutral towards representationalism. No assumption need be made that a sign must mirror the world or some mathematical reality. Semiotics regards signs, symbols, texts and all of language as constitutively public. However, meanings and imagery can be and are appropriated, elaborated and created by individuals and groups as they adopt, develop and invent sign-uses in the contexts of teaching, learning, doing and reflecting on mathematics, and all of the other important activities of life. Thus semiotics rejects the simple subjective/objective dichotomy that consigns mathematical knowledge to 'in here' or 'up there.' It provides a liberating perspective from which to study mathematics and education. It opens a new avenue of access to the concepts that have been developed for mathematics education in the social sciences and the other sciences, including psychology, but it also allows access to the intellectual resources and methods of the arts and humanities.

AGENCY AND CREATIVITY

Learners are human beings with all the complexity and moral aspects this involves. Human beings are constitutively social beings and this entails a widespread range of capacities concerning interpretation and sense-making in social or interpersonal situations. Focussing on the classroom, learners understand in their own ways the roles and asymmetric power relations of the teacher-student relationship, the aims and purposes of school mathematical activity and tasks (both espoused and enacted, where these differ[1]), and many other relevant aspects of micro-social context. Into the shifting and multifaceted context of the classroom learners brings their own historically formed subjectivity, sense of self, and capacities for meaning making. Acknowledging this formative background, the features I wish to focus on here are the central ones of agency and creativity.

Agency is the central capacity all human beings have for initiating (and continuing) activities, including the possibility of inaction. In focussing on learner agency I am not assuming that students or persons in general are rational beings making rational choices. All sorts of psychological factors can drive choices and behaviour, but this is irrelevant to the present discussion. In learning mathematics, the activities involved are primarily communicative involving mathematical sign systems, notably sign 'reception' (listening, reading) and sign production (speaking, writing, sketching). Creativity in such activities or conversations may be conceptualised as the ultimate expression of agency. In a minimal sense, almost any semiotic sign production can be classified as creative, because it involves first making a selection from the semiotic repertoire available, which includes signs and modes of expression, and then putting together and making a new public utterance. In practice the selection, combination and utterance of signs may very well be woven inseparably into a single action. By definition, any sign utterance is new because of its unique temporal and contextual location in conversation. However such usage trivialises the term creativity through making it universally applicable. By analogy with problem solving (a significant analogy, especially in the domain of mathematics) routine utterances can be distinguished from non-routine utterances. In the latter, semiotic elements (including the context) are combined in a novel and non-routinised way in the utterance. It is cases like this that are better characterised as creative.

Manifestations of agency in sign system usage are understood here, based here on a Wittgensteinian (1953) perspective, participation in *language games* embedded within social *forms of life*. Thus communicative activity involving mathematical sign systems is always encompassed within the social. Furthermore, the component activities of sign reception and production involved in language games are woven together within the larger epistemological unit of *conversation* (Ernest 1991, 1994, 1998). The way in which these two activities are mutually shaping in is shown in the model (Figure 1) of sign appropriation (reception) and sign use (production).

Figure 1 is based on Harré's (1983) model of 'Vygotskian space', previously applied to mathematics in Ernest (1998).[2] Evidently it embodies the well known dictum of Vygotsky 1978, 128)

> Every function in the child's cultural development appears twice, on two levels. First, on the social and later on the psychological level; first between people as an interpsychological category, and then inside the child as an intrapsychological category.

In the figure these two levels are represented, at least in part, first by the top left corner, for the socio-cultural is both public and collective, and secondly, by the bottom right corner, for the (intra)psychological is both individual and private. The other two corners are crossing points on the boundary between the two levels, and these are the locations where learner semiotic agency is acted out.

SOCIAL LOCATION

MANIFESTATION	Individual		Collective
Collective	Learner's public utilisation of sign to express personal meaning (Public & Individual)	*Conventionalization* →	Social (teacher & others) negotiated and conventionalised (via critical acceptance) sign use (Public & Collective)
	Publication ↑		↓ *Appropriation*
Private	Learner's development of personal meanings for sign and its use (Private & Individual)	← *Transformation*	Learner's own unreflective response to and imitative use of new sign utterance (Private & Collective)

Figure 1. Model of Sign Appropriation and Use

Following the processes in the model, signs and sign systems become adopted by the individual learner first in the process of *appropriation*. This leads to the learner's own unreflective response to and imitative use of a single sign, be it atomic or compound, or of a set of sign utterances. The learner has thus appropriated a collective sign into something for herself that is private. This is also the route by means of which learners appropriate the rules of sign-use, mostly through observing their exemplification in practise. Agency is manifested in several ways at this stage, including attending to the public sign utterance, becoming aware, to a greater or lesser extent, of the immediate context and associations of the sign use, and using the sign in an imitative way. The privately initiated uses of the sign, albeit possibly in response to another's request or command, are a public manifestation of learner agency. In such use the whole cycle is brought into play in miniature, because the sign as utilised in a personal performance is manifested publicly, and would normally be subject to social acceptance or correction (conventionalization). Such use corresponds in great part to Skemp (1976) and Mellin-Olsen's (1981) notion of *instrumentalism*, because of the simple imitative performativity involved. I avoid the term 'instrumental understanding' here, because of the commonly associated ideological assumption that locates knowledge and understanding 'inside' the private minds of individuals rather than as primarily manifested in public performances (which can also be rehearsed in private thought). Through the conventionalization of performance (applied to sign utterances) at this stage the learner also can become aware of restraints and restrictions applying to sign use, that is some of the rules of sign production that constitute part of the overall sign system.

When the next stage is achieved for a particular sign, which may follow a whole sequence of related appropriations, performances and conventionalization s in the mini-cycle described above, the learner will usually develop personal meanings for the sign and its use. This transforms it into something that is individual as well as private, because of the personal meanings associated with the sign. This will typically include a whole nexus of associations including a sense of where and how the sign is to be used acceptably. Such associations are primarily tacit, manifested in usage, but can include rationalisations and explanations about the limits, nature and purposes of sign usage. These may be appropriated from teacher and peer explanations prior to transformation into the meaning nexus, very likely tested and corrected by further mini-cycles involving publication and conventionalization . The successful appropriation and transformation of a sign, with its nexus of associated meanings and meta-discourse, finds a parallel Skemp's (1976) notion of 'relational understanding' in mathematics. This involves not only being able to use the sign correctly, that is, mostly corresponding to conventionally accepted usage within the micro-community of the classroom under the authority of the teacher, but also being able to offer a rationale or explanation for the usage. It may be inappropriate to describe the transformational process in which a meaning nexus is elaborated privately by the individual as manifestation of agency, as many of the processes are unconscious and involuntary. However the attention, persistence, and repeated performances in both sign utterances and explanatory meta-discourse evidently are manifestations of agency.

The third phase illustrated in Figure 1 is that of *publication*. In this process the individual learner engages in a conversational act in publicly performing or making a sign utterance. Mathematically this could vary from a quick, spontaneous verbal, gestural or written response to a question or other stimulus, through to constructing an extended text elaborated and revised over a period of time, prior to offering it to others. A group of learners can elaborate such a text co-operatively, but this process will have subsumed many sub-cycles in which individuals have communicated (offered signs) to others in the group in an extended conversation giving rise to a jointly elaborated, negotiated and agreed text.

It is in the publication stage of the overall cycle that agency is manifested most evidently and clearly. For the individual must initiate and produce a public sign utterance. At the simplest level this is an act of participation or even will, mediated through semiotic and social capabilities. More complex sign productions and utterances involve an elaborate series of meaning-attentive and meaning-driven voluntary actions. Agency is involved in interpreting the context and in choosing the mode, type and particular sign response and in making it. However, many psychological and social factors can inhibit, distort or enhance this performance, including such things as the learners self-confidence, perception of the surrounding others, classroom climate and so on.

Finally, the overall cycle is completed through the process of *conventionalization*. In this phase learner sign productions having been fed into the social milieu (the classroom conversation) are subjected to attention, critique, negotiation, reformulation and acceptance, or sometimes rejection, by the teacher and others. The outcome is an agreed or imposed conventionalization which is both public and collective. Because of the power and authority asymmetry in the classroom (and indeed in virtually all interpersonal contexts, but especially in socially sanctioned learner-teacher relationships) teacher approval will normally be the final arbiter of acceptance, rather than majority or learner agreement. Typically the conventionalized sign that is accepted will need to satisfy the following criteria.

1. **Relevance**. The sign or text is perceived to be a relevant response or putative solution (or possibly an intermediary stage to one) to a recognized (i. e., sanctioned) starting sign which has the role of a task, question or exercise. This might be teacher imposed or otherwise shared and authorized.
2. **Justification**. The mode of and steps in the derivation of the sign from the authorised 'starting point' will normally be exhibited as a semiotic transformation of signs, that is employing accepted or acceptable rules or means of sign transformations within the semiotic system, or justified meta-linguistically.[3]
3. **Form**. Both the signs and their transformations (where offered) will normally exhibit teacher-acceptable form, thus conforming to the rhetoric of the semiotic system involved as realized and defined in that classroom. This system could be that of spoken verbal comments, drawn and labeled diagrams, numerical calculations, algebraic derivations, or some combination of these or other sign types.[4]

These criteria primarily apply at the object language level, that is they directly concern mathematical tasks or contents. However they can also be applied meta-linguistically as comments on rather than as additions to object language level utterances in the classroom conversation.

If the public sign utterance deviates in relevance, justification or form a central aspect of the conventionalization stage will be the criticism, rejection or correction of the sign for its lack of acceptability in these dimensions. Such a process may involve 'degoaling', i. e., switching to a new goal, target or task (Hughes 1986) which could be intended to serve as an intermediate step towards the original goal, or which might be a shift in the discourse to a new subject matter. Conversation, even in its formal and controlled manifestation as it occurs in the mathematics classroom can be fluid and shifting in its actualisation, just as it can be rigid and one-sided. It can be 'live' in which near spontaneous verbal responses as well as other modes of response are sought and encouraged by the teacher and expressed by learners, or it can be highly formalised and regulated with the teacher directing attention to written tasks and requiring (and allowing) only formal written responses to them at determined moments.

The process of conventionalization is the stage in the cycle that is most public. For it often acts on a sign uttered or presented by the learner and involves the critical acceptance, correction or rejection of the sign. This is where the teacher's agency is at work, directed at the capabilities involved in skilled sign production. Indeed, the teacher may initiate the semiotic cycle at this point by introducing her own sign or text. (Mostly this will refer to previously introduced signs and conversations, but it may have a variety of functions beyond task setting, including explanation or scene setting to aid learners in the creation of meaning.) Skemp (1979) has described a central aspect of the teachers' aim as being the development of *logical understand-*

ing in the learner, to cap the instrumental (performance orientated) and relational (meaning elaboration and justification production) capabilities. The learner manifests logical understanding in this sense through being able to utilise and produce signs using the correct mode of expression and 'grammatical form', thus demonstrating a growing mastery of relevant aspects of the rhetoric of school mathematics. Through participation in and experience of conventionalization the learner first appropriates and then transforms into a personal aspect of her individual agency the capability of a critical and corrective perspective on signs. This involves not only the ability to produce signs in accordance with the (growing) set of rules of sign production manifested in the classroom, but also the capability to critically review and correct signs to conform to these rules. Ultimately the successful learner develops and adopts the aspect of agency corresponding to the role of the critic; the ability to make judgements concerning the correctness of sign utterances (with respect to relevance, justification or form) as is appropriate to the context. This involves the appropriation of a social role, a mode of 'voice', first experienced in the actions of others in conversation.

Traditionally in linguistic research two modes of sign usage are distinguished: listening/reading and speaking/writing. From the perspective of mathematical learner agency we might also distinguish two levels of functioning: lower level (responsive) and higher level (autonomous). Lower level functioning involves responding to signs or texts 'literally'. In listening/reading in school mathematics this means taking the signs as simply presenting routine tasks or instructions, or less commonly, as informational. In speaking/writing this usually involves simply offering an utterance in a response to some semiotic stimulus (spoken or written) delimited by the perceived constraints of the social context of utterance. In mathematics typically this involves simply performing a routine task. This usually necessitates applying one or more semiotic transformations to a sign, resulting in a sequence of signs (e. g., counting vocally or subvocally, performing column addition, solving a linear equation) resulting in a terminal sign, the 'answer'. Underpinning this is the ability to make sense of mathematical signs and texts, to interpret them as tasks and to apprehend their object, purpose and goals, within a variety of contexts, most notably, in the school context. Where these abilities are lacking or not fully developed it is the role of conversations directed by the teacher or more capable others, following the model in Figure 1, to further develop them.

Higher level or autonomous functioning means responding to signs in a more reflective way. In listening/reading this means spending time and making more effort to explore and create meanings for signs and also engaging in self-monitoring and self-reflection in the process. As the term reflection suggests, this involves elements of inward or self-directed dialogue. The metaphor of examining one's image in a mirror suggests stepping outside oneself and viewing oneself from the perspective of another, adopting an outsider's viewpoint. In dialogue, a person can adopt two opposite roles. First there is the role of proponent (or friendly listener) presenting (or following) sympathetically a text, a line of uttered or privately rehearsed argument or thought experiment, for exploratory or understanding purposes (Peirce 1931 – 58, Rotman 1993). By 'sympathetic' I mean adopting the point of view of the proponent or utterer and attempting to construct and enter into the sense of the utterance as it is (understood to be) intended. This is attempting to 'share' the constructor's meaning, rather than looking for grounds on which to dismiss it for failure of relevance, justification or form (this, taken to extremes, can pre-empt fully developed and elaborated sense-making). However, the role of proponent is not intrinsically reflective or higher order, for it can also be adopted at a lower, passively attentive level.

Secondly, there is the role of critic, in which a text, a sequence of signs, which could be an argument, a mathematical derivation, and so forth, is examined for weaknesses and flaws. This involves having appropriated and transformed into personal capabilities at least some of the context-specific criteria of acceptability manifested by others (primarily the teacher). These criteria typically pertain to the relevance, justifiability or rhetorical form of the text or sign utterance in question, and are meta-linguistic criteria when made explicit. Being able to adopt the role of critic to apply to others' or one's own texts is an intrinsically reflective and higher order capacity. It cannot be done meaningfully in an automatic or thoughtless way. This fits with the tradition in educational psychology that classifies evaluation, defined as making judgements using internal (i. e., textual) evidence and external criteria, as belonging to the highest level of intellectual functioning (Bloom 1956). It also evidently encompasses a dimension of agency since it constitutes the adoption of a specific agentic role.

In speaking/writing, higher level or autonomous functioning means constructing and elaborating signs or texts in a thoughtful and reflective way. Typically in school mathematics this involves the transformation of tasks presented as mathematical texts into further more manageable representations and in doing so applying a variety of textual and symbolic transformations to representations and their parts to complete the tasks. Different modes of representation can be employed singly or together in a school mathematics text, including any combination of symbols, written language, labelled diagrams, tables, sketches, models and arrayed objects (and even gestures where the text is spoken). It is common in school mathematics for problem solution processes to use more modes of representation than the starting text (task), or the final text (answer). The procedures of problem solving include the active processes of imagining, writing, drawing or making sequences of representations (not necessarily either monotonic or single branched sequences) progressing from the initial text (given task) to a final (in terms of fulfilling task demands) and permissible (derived by allowed transformations), often simple, textual representation (the potential task 'solution'). To carry through a multi-step process of this type successfully requires the student to be attentive to and in control of the purpose, direction and outcomes of subsidiary procedures and transformations. Where the construction and concatenation of the sequence of semiotic actions deployed is not automatic, that is has not been practised on similar tasks until it has become routinised for this particular student, it is appropriate to call it creative. It corresponds to non-routine problem solving and involves the student or person in constructing and combining in novel ways (new to herself, at least) different signs and procedures.

Carrying out tasks individually or in groups may be the most common higher level activity in speaking/writing in school mathematics. However, other activities can also occur such as the students writing mathematical questions and tasks, or posing mathematical problems themselves, with some sense of what the solution processes entail. Either way, speaking/writing at this level involves the most obvious and explicit manifestation of learner agency, since the activities are internally initiated and conducted. They are, of course, also texts uttered in response to antecedent texts in a conversation; but then so is all semiotic and communicative activity. Once again, the higher level agentic functioning involved in writing questions and tasks, and posing problems in the mathematical classroom is creative activity since it involves the construction of imaginative new texts.

Studies comparing novice and expert problem solvers in mathematics have shown that the latter successfully combine (and alternate between) the two higher level roles distinguished above, namely proponent and critic. Schoenfeld (1992), for example, found that novices typically spent most of their time in aimless exploration of problems, seeking to solve without any conscious design. This can be valuable for enriching understanding, but when persisted in, as in the study, it usually led to failure. The expert problem solvers and mathematicians cycled through a variety of activities directed at the problem, including reading, analysing, exploring, planning, implementing, and verifying. Furthermore, they repeatedly asked self-directed questions, typically at the points of transition between the different types of activity. These were higher level, critical and self-regulative questions asking what was being sought, what was being found, etc. This illustrates how higher level creative activity in mathematics needs to combine the roles of proponent and critic in an internalised, self-directed dialogue. Thus following the model shown in Figure 1 it is not just signs that become appropriated by persons, but the whole cyclic conversational process ultimately must become internalised for high level creative activity in school mathematics and in mathematics itself.

University of Exeter

NOTES

[1] I make this distinction, because as is well known the overt purpose of a classroom mathematical task and what the students come to learn is the teacher's actual focus of attention or emphasized outcome may differ (e. g., working an exercise vs. writing its solution in a certain style).

[2] In Ernest (1998) I utilize this model explicitly to account for the acquisition of language, mathematics and mental powers by young learners, as well as using a parallel model for the creation of shared mathematical knowledge in and by the mathematics research community. However, I view this model as showing the interplay between public vs. private and collective vs. individual in the role and meanings attributed to signs and texts (as well as in the construction of signs and texts) in conversation in general. This has particular relevance to the years of formal schooling, which I focus on here.

[3] Sign transformations do not always mean the replacement of just one (or more) part(s) of a compound sign by another part(s), with the retention of the unreplaced parts. It may involve replacement of the whole sign complex by another. For example, in a logical proof (a classic transformational sequence in advanced mathematics) some proof steps involve the insertion of a new sign with no components shared or overlapping with the previous step, e. g., in axiom use.

[4] The rhetoric of school mathematics concerns the standards, norms and rules (possibly tacit) of grammatical and expressional correctness, as well as stylistic and genre appropriateness, in the presentation and modes of expression of signs. These norms and rules are primarily applied to formal written texts (including symbols, diagrams, etc.), although spoken expressions are also rhetorically constrained, but usually more loosely. In contrast to logic the rhetoric of school mathematics is highly local and context-

bound, and for contingent and historical reasons varying rules and norms are applied across different institutions and locations (as well as at different ages).

REFERENCES

Bachelard, G. (1951). *L'activité rationaliste de la physique contemporaine.* Paris.

Bishop, E. (1967). *Foundations of Constructive Analysis.* New York: McGraw-Hill.

Bloom, B. S. et al. Eds (1956). *Taxonomy of Educational Objectives 1, Cognitive Domain.* New York: David McKay.

Chomsky, N. (1965). *Aspects of the Theory of Syntax.* Cambridge, Massachusetts: MIT Press.

Davis, C. (1974). Materialist philosophy of mathematics. In R. S. Cohen, J. Stachel & M. W. Wartofsky Eds., *For Dirk Struik.* Dordrecht: Reidel.

Dubinsky E (1988). On Helping Students Construct The Concept of Quantification. In A. Borbas (Ed.) *Proceedings of PME 12.* Veszprem, Hungary, Vol. 1, 255 – 262

Ernest, P. (1991). *The Philosophy of Mathematics Education.* London, Falmer Press.

Ernest, P. (1994). *Conversation as a Metaphor for Mathematics and Learning, Proceedings of BSRLM Conference, MMU 22 November 1993.* Nottingham: BSRLM, 58 – 63.

Ernest, P. (1998). *Social Constructivism as a Philosophy of Mathematics.* Albany, New York: SUNY Press.

Harré, R. (1983). *Personal Being.* Oxford: Blackwell.

Heyting, A. (1956). *Intuitionism: An Introduction.* Amsterdam: North-Holland.

Hughes, M. (1986). *Children and Number.* Oxford: Blackwell.

Machover, M. (1983). Towards a New Philosophy of Mathematics. *British Journal for the Philosophy of Science* 34, 1 – 11.

Mellin-Olsen, S. (1981). Instrumentalism as an Educational Concept. *Educational Studies in Mathematics* 12, 351 – 367.

Peirce, C. S. (1931 – 58). *Collected Papers* (8 Vols). Cambridge, Massachusetts: Harvard University Press.

Rotman, B. (1993). *Ad Infinitum.* Stanford California: Stanford University Press.

Schoenfeld, A. (1992). Learning to Think Mathematically, in Grouws, D. A. Ed., *Handbook of Research on Mathematics Teaching and Learning.* New York: Macmillan, 334 – 370.

Sfard, A. (1993). Reification as the birth of metaphor, *For the Learning of Mathematics* 14.1, 44 – 55.

Sierpinska, A. (1987). Humanities students and epistemological obstacles related to limits. Educational Studies in Mathematics, 18, 371–397.

Skemp, R. R. (1976). 'Relational Understanding and Instrumental Understanding' Mathematics Teaching 77, 20 – 26.

Skemp, R. R. (1979). Goals of Learning and Qualities of Understanding, Mathematics Teaching. 88, 44 – 49.

Troelstra, A. and van Dalen, D. (1988). *Constructivism in Mathematics: An Introduction,* Vol. 1. Amsterdam: North Holland.

Vygotsky, L. (1978). *Mind in Society, Cambridge.* Massachusetts: Harvard University Press.

Wittgenstein, L. (1953). *Philosophical Investigations,*.Oxford: Basil Blackwell.

SUSANNA MARIETTI

THE SEMIOTIC APPROACH TO MATHEMATICAL EVIDENCE AND GENERALIZATION

Abstract. The fundamentals of Peircean semiotics have been applied by Peirce himself to the main philosophical questions relating mathematics. Following Michael Otte's suggestion of resorting to a semiotic approach to mathematical epistemology in order to understand mathematical cognition, it is possible to account for the chief problem of generalization, going beyond the traditional explanations exemplified by Locke's use of abstract general ideas and Berkeley's criticism to it. Against the background of Peirce's main lines of departure from Kantian transcendentalism, the problem of the evidence obtained from proofs performed upon individual diagrams and laying claim to universality can be faced within a semiotic frame that focuses on the interplay of iconic, indexical and symbolic elements of signs.

Key words: diagram, evidence, generalization, Kant, mathematics, Otte, Peirce, semiotics.

Since the beginning of the history of philosophy, generalization has been the chief problem in the inquiry into the foundations of knowledge. There is no science of the individuals, said Aristotle. We have to apply our propositions to a general universe of things in order to obtain scientific knowledge. When we affirm that a falling stone has a definite acceleration, we are not referring to the particular stone that fell to the ground 5 min ago, but rather to each stone of a potentially unlimited set of stones that falls under certain conditions. In the same way, when we affirm that the sum of the internal angles of a triangle equals two right angles, we are not referring to the particular triangle drawn on the blackboard in front of us, but rather to an infinite number of triangles. Yet, we are not infinite creatures, and we are not able to cope with an infinity of cognitive acts. We cannot examine all possible triangles in order to verify the size of their angles. How can we then achieve that certainty which characterizes the geometrical theorem at issue? How can we rely on our limited cognitive faculties in order to gain a potentially unlimited knowledge?

Various have been the attempts to answer these questions, which in Michael Otte's proposal of a mathematical epistemology from a semiotic point of view take quite a peculiar form. "Epistemology is about the relationship between these types of entities, objects and signs," he says. "As all general phenomena are fundamentally semiotic entities, while singular phenomena are not intrinsically signs, we could also say that epistemology is concerned with the relation between the singular and the general. In this way generalization appears as a fundamental problem of epistemology and of education. ... to know implies, in any case, to relate a particular to a general, it means to generalize." (Otte in print, 3)

Charles Sanders Peirce, the father of modern semiotics, identifies mathematical reasoning with diagrammatic reasoning, describing the latter as follows:

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 35 – 43.

> By diagrammatic reasoning, I mean reasoning which constructs a diagram according to a precept expressed in general terms, performs experiments upon this diagram, notes their results, assures itself that similar experiments performed upon any diagram constructed according to the same precept would have the same results, and expresses this in general terms. This was a discovery of no little importance, showing, as it does, that all knowledge without exception comes from observation. (Peirce NEM IV, 47 f.)

This brief account runs through the different steps of what Peirce found out to be the common structure of all significant mathematical demonstrations, which he defined as *theorematic deductions*. In theorematic deductions, the mathematician has to perform experiments upon the original diagram constructed according to the premises. He must manipulate the diagram, because the major theorems of mathematics do not stand out immediately from it. Euclid, for instance, in proving the theorem on internal angles, adds new points and lines to the original triangle. These new elements are not suggested to him by any previous knowledge. In choosing them, he relies exclusively on his sagacity.

The original triangle and the enriched figure are both instances of a mathematical diagram. A diagram is always a single element, a *token* as opposed to a *type*. It is essential to its role that it should be an individual, in so far as the mathematician must experiment new strategies of demonstration upon it and observe the results of his or her experiments. "Thinking in general terms is not enough," Peirce says. "It is necessary that something should be *done*" (Peirce CP 4.233). All knowledge comes from observation, he can then conclude.

Although the presence of the diagram is particularly evident in geometry, all deductive reasoning is, according to Peirce, diagrammatic, which means that observation plays a role in the whole of mathematics. No matter how easy it may be, every deductive inference hinges upon observation. It is only by observation that we "recognize that because $y - y = 0$, therefore $x + y - y = x + 0$" (Peirce manuscript 16, MS in the following; numeration follows Robin 1967). Still more, "observation is required in the simplest syllogism. Thus, if we reason, 'All men are mortal, Enoch is a man, therefore Enoch is mortal,' we only do this by observing that the *man* of the first premise is the same predicate as the man of the second premise, etc." (Peirce MS 17).

I would like to leave Peirce for a while in order to read a brief passage from John Locke's *An essay concerning human understanding*. As is known, Locke solves the problem of generalization through the introduction of abstract general ideas. In order to be general, knowledge must address abstract ideas. Having defined knowledge as "the perception of the connection and agreement, or disagreement and repugnancy, of any of our ideas," (424) Locke notes that, in mathematics, the only kind of knowledge absolutely certain and universal, we are faced with two different degrees of evidence. The first one seems to be very close to that which characterizes Peirce's *corollarial deduction* (the simpler and more immediate kind of deduction, in which experiments upon diagrams are not required), while the other is so described:

> In this case then, when the mind cannot so bring its ideas together as, by their immediate comparison and, as it were, juxtaposition or application one to another, to perceive their agreement or disagreement, it is fain, by the intervention of other ideas (one or more, as it happens), to discover the agreement or disagreement which it searches; and this is that which we call "reasoning." Thus the mind, being willing to know the agree-

> ment or disagreement in bigness between the three angles of a triangle and two right ones, cannot by an immediate view and comparing them, do it ... In this case the mind is fain to find out some other angles, to which the three angles of a triangle have an equality; and finding those equal to two right ones, comes to know their equality to two right ones ... A quickness in the mind to find out these intermediate ideas (that shall discover the agreement or disagreement of any other), and to apply them right, is, I suppose, that which is called "sagacity." (Locke 1910, 434)

These words strongly recall the ones used by Peirce to affirm that mathematical theorems are not immediately evident as such, for the diagram must be modified in order to get to the conclusion. Like Locke, he also calls in question the sagacity of the mathematician engaged in choosing the suitable modifications. Furthermore, the use made everywhere by Locke of perception terms can create the impression that something similar to Peirce's insistency upon mathematical observation is at work here.

I am not directly interested in Lockean philosophy here. I mention it because it is representative of an approach to gnosiological issues that is very important in the history of thought. Having observed a certain number of individual things, Locke says, we are able to concentrate our attention on their common characters, thus gaining an abstract idea with which all the individuals agree. Locke thinks that this ability to create abstract general ideas is the distinctive mark of human being, as it allows us to communicate and to acquire true knowledge.

Now, though Locke explicitly says that only individuals exist, whereas the idea of an abstract triangle is imperfect and cannot exist, it seems to me that abstract ideas cannot be reducible to the ability just described. They cannot be merely the outcome of our temporary selective attention. Locke's explicit nominalism must necessarily have in itself a great deal of realism. If the general triangle has to fulfil its role of being the object of a general knowledge, I think it must possess a definite ontological status making it an autonomous entity independent from the single individuals from which it was abstracted. Indeed, if the whole matter had been really reducible to the psychological phenomenon of selective attention, abstract ideas would not have been discussed so much in the history of philosophy.

In order to explain general knowledge, Locke has to include a sort of Platonic form in his ontology, so that our mind, contemplating the general idea of a triangle, could perceive the agreement or disagreement of the angles at issue. This kind of perception, we can now see, is very different from that spoken of by Peirce in his description of mathematical inference, in which perception is something essential to the development of the demonstration. Mathematical reasoning is diagrammatic reasoning, and diagrammatic reasoning takes the form outlined in the quotation above. Everything takes place upon a diagram, a single individual semiotic entity, and the whole inference consists in nothing other than observing the material relations among the different parts of the diagram. In Locke's description, in contrast, observation plays a fully accidental role. Here we find a definite object of reference beyond the sign, that is to say, the abstract general idea, which is the only protagonist of the inference. Locke, it is true, has a propensity to conceive abstract ideas in terms of images, hence his use of a large amount of perception vocabulary. But it does not belong in an essential way to the description of mathematical reasoning as such. Ac-

cording to Locke, the only essential aspect of mathematical reasoning is binding together *in thought* ideas that have, in turn, an essential connection to each other.
It is well known how Locke's conception of abstract general ideas was opposed strongly by Berkeley, who maintained that the idea of a triangle that is neither equilateral nor isosceles nor scalene is intrinsically contradictory, so that the only ideas that can be formed in our mind are those of particular things. Between our knowledge and the ideas of individual objects, we cannot find the mediation of abstract ideas. General knowledge, in Berkeley's view, is supported by a quite different kind of process. According to him, what happens is that a single particular idea is used to represent all particular ones belonging to the same species. Euclid proves his theorem upon a singular triangle determined in all details, but "neither the right angle, nor the equality, nor determinate length of the sides are at all concerned in the demonstration. It is true the diagram I have in view includes all these particulars, but then there is not the least mention made of them in the proof of the proposition ... Which sufficiently shows that the right angle might have been oblique, and the sides unequal, and for all that the demonstration have held good." (Berkeley 1957, 15 f.)
As in the case of Locke, but from a different point of view, Berkeley's words also recall Peirce's description of deductive reasoning, in which demonstration refers to an individual diagram, and an analogous process seems to be responsible for generalization. In fact, Locke's abstract idea cannot support the experimentation on which Peirce bases his analysis of theorematic inference, which would be impracticable on a conceptual general level. Nevertheless, Berkeley's nominalism is far from Peirce's realist approach. A different conception of the individual must be at issue, a conception that has its root in Peirce's semiotic turn. According to Peirce, a fully determined individual image is inconceivable. To Berkeley's claim, he replies:

> No statement of Locke has been so scouted by all friends of images as his denial that the "idea" of a triangle must be either of an obtuse-angled, right-angled, or acute-angled triangle. In fact, the image of a triangle must be of one, each of whose angles is of a certain number of degrees, minutes, and seconds.
> This being so, it is apparent that no man has a true image of the road to his office, or of any other real thing. (Peirce CP 5.299-300)

Berkeley, however, is not dealing exclusively with images. According to him, even in thought it is not *logically* possible to possess an idea that is not fully determined. Peirce opposes resolutely the notion of such a logical atom, a "term not capable of logical division, ... one of which every predicate may be universally affirmed or denied ... Such a term can be realized neither in thought nor in sense" (CP 3.93).The semiotic bedrock beneath Peirce's gnosiology cannot recognize the absolute individual at all, for such an individual cannot be realized in a sign context.
I have sketched Locke's and Berkeley's conceptions of generalization because they are truly representative of our main philosophical inheritance. Against such a background, even with the apparent similarities that we may discern, Peirce's revolution in paradigm emerges in all its strength. As far as I can understand Peirce's semiotic explanation of mathematical knowledge, he no longer faces the question of the relationship between individual and general with respect to the kind of object at issue – general and abstract in itself or particular but interchangeable with other objects – but deals with it on the sole level of the diagrammatic sign. Locke's abstract idea,

not suitable for experimentation, and Berkeley's logical atom, not even conceivable are, according to Peirce, both dogmatic notions inadequate to explain mathematical knowledge. Mathematical knowledge is now seen in terms of the manipulation of signs, which have become triadic entities according to Peirce's system of categories.

Peirce's categorial analysis is the real foundation of his semiotics. Such analysis starts from a deep study of Kant's *Critique of pure reason*, but it departs from Kantian transcendentalism both in method and outcome. As to the former, the investigation carried on by Peirce does not resemble in any way a *critique*. Peirce's philosophy – to which he gives the name of *phaneroscopy* – is rather a positive knowledge, which differs from a special science only in that the latter "seeks such truth as can only be discovered from peculiar experiences sought out for the purpose," whereas the former "seeks such universal truth as can be discovered from everyman's hourly experience"(NEM IV, 228). Peirce defines the *phaneron* as "the collective whole of all that could ever be present to the mind in any way or in any sense" (ibid. 320). Now, as the main assumption of the semiotic approach is that "our cognitive access to reality is relative and mediated by signs," as Michael Otte (in print, 1) states it, whatever could be present to the mind always has a semiotic nature. Phaneroscopy thus turns into semiotics, which is conceived as the positive observation of all signs present to the mind. Such an observation is directed toward the ascertainment of the universal classes to which signs belong, which are found to be the three categories of *firstness*, *secondness,* and *thirdness*. Hence, the categories are found by Peirce through an *a posteriori* investigation. "My view is that there are three modes of being," he writes. "I hold that we can directly observe them in elements of whatever is at any time before the mind in any way." (CP 1.23)

It seems to me that this is to be considered as the main distinction between Peirce and Kant. From here, everything follows. This *a posteriori* method is responsible for the fact that not only the third category (mediation, symbol, generality), but also the first two already belong to sign interpretation. Otherwise, they could not be observed. This circumstance allows Peirce to get rid of whatever he maintained to be metaphysically dogmatic in Kant's conception in that it was not detectable through the direct observation of the phaneron: on the one hand, things in themselves, as what gives rise to the empiric manifold, on the other hand, the synthetic unity of consciousness, as what is responsible for the objectiveness of knowledge.

Translated into phaneroscopic terms, Kantian manifold becomes the category of firstness (Cf. CP 1.302), which, from a semiotic point of view, is considered by Peirce as the icon. It provides the "matter of consciousness," or "that which is immediately present in consciousness" (MS 16). In a mathematical diagram, the mere setting down of the relations among its parts constitutes that icon which is the matter of the diagram itself. But such a matter, unlike Kant's manifold, is already expressed in semiotic-categorial terms, because the relations are embodied in a material diagrammatic *token* that can be interpreted symbolically according to its formal structure.

As to the synthetic unity of consciousness, Peirce rejects from the core of his philosophy the Kantian conception of a mental synthesis preceding all analyses, maintaining that "something is presented which in itself has no parts, but which nevertheless is analyzed by the mind," and that it is only after such an analysis that "we are carried in spite of ourselves from one thought to another, and therein lies the first real synthesis. An earlier synthesis than that is a fiction"(CP 1.384). This means that the constitution of the object is no more dependent on self-consciousness, but rather on the semiotic chain that transforms signs into other signs that are equivalent under some respect. We are carried from one thought to another along a train of representations in which "[t]he object of representation can be nothing but a representation of which the first representation is the interpretant. But an endless series of representations, each representing the one behind it, may be conceived to have an absolute object at its limit." (CP 1.339)

The empirical sciences apply the three inferences – abduction, deduction, induction – in order to interpret a sign by another sign, according to a process that, in the long run, if research will be carried on indefinitely, will free our representations from errors and personal idiosyncrasies. Such a chain of interpretations is what now answers for the objectivity of knowledge, in which the absolute, fully determined object remains as a limit notion. This does not amount in any way to a weakening of the concept of real. Reality is defined sharply as "that which is such as it is whatever you or I or any generation of men may opine or otherwise think that it is" (Peirce, MS 498). Only, it is no more dependent on the synthesis of apperception that should bring to a conceptual unity a manifold placed beyond the system of categories. In other words, Peirce claimed to have overcome the residual dogmatism of transcendentalism by replacing the *a priori* principles of understanding through the unlimited community of researchers. Unlike Kant's, his concern is not, in Michael Otte's words, "with the unity of ideas (*Vorstellung*) in a self-consciousness, but rather with the socially effective unity represented by signs." (1997, 337)

An analogous approach, it seems to me, holds for mathematical knowledge. If I am correct, that circumstance is responsible for the sharp difference between Peirce's answer to the generalization problem and the traditional ones exemplified in the words of Locke and Berkeley. Like the empirical case, mathematical reasoning also has to be conceived as a chain of interpreting signs, namely, diagrams in their permitted transformations. It is on this semiotic level that the constitution of the object is grounded. In his system, Peirce can deal neither with Berkeley's pre-semiotic object nor with Locke's wholly conceptual abstract idea. The diagram already partakes of all three categorial universes, which become the iconic, indexical (required in order to link the diagram to the mathematical hypothesis at issue), and symbolic sign.

The diagram is in itself already particular and general. The mere *suchness* of firstness – namely, the simple presentation of concrete elements in mutual relation – and the general mediation of thirdness are both required in its construction. In Kant's terminology, it is a *schema*. This is, in a certain sense, the only element of Kantian transcendental analysis to survive in Peirce's system. But here its role is no longer that of mediating between two different faculties and their heterogeneous contributions to knowledge, but rather it constitutes the very starting point of phaneroscopic investigation. There is nothing before the triadic sign. Peirce charges Kant with the fault of having drawn "too hard a line between the operations of observation and of ratiocination. He allows himself to fall into the habit of thinking that the latter only begins after the former is complete." And he adds: "His doctrine of the *schemata* can only have been an afterthought, an addition to his system after it was substantially complete. For if the *schemata* had been considered early enough, they would have overgrown his whole work" (CP 1.35). Peirce's intent was to avoid every dogmatic element not detectable through the direct inspection of what is present to the mind. In his *a posteriori* methodology, the Kantian distinction between two different faculties – sensibility and understanding – becomes vacuous. We can only deal with already formed schemata.

Now, a mathematical diagram is a set of elements among which some mutual relations hold. By manipulating these elements, we find evidence of new hidden relations among them. Indeed, "necessary reasoning makes its conclusion *evident*." But here Peirce asks:

> What is this "Evidence"? It consists in the fact that the truth of the conclusion is *perceived*, in all its generality, and in the generality the how and why of the truth is perceived. What sort of a Sign can communicate this Evidence? (Peirce NEM IV, 317)

It cannot be an index, because of the *brute force* through which it signifies its object. It cannot be a symbol either, because a symbol only rests on habits, and habits are not evidence. It must then be an icon, the only kind of sign that can communicate evidence through the perception of it. An icon, in fact, is always an individual object, and thus it is capable of being observed. It signifies thanks to the concreteness of each single replica: "Such a sign whose significance lies in the qualities of its replicas in themselves is an icon, image, analogue, or copy." (Peirce MS 7)

The mathematician can perceive the iconical sign, "the only sign which directly brings the interpretant to close quarters with the meaning; and for that reason it is the kind of sign with which the mathematician works" (ibid.). But, in the new triadic conception of the sign, the diagram presupposes a symbolic interpretant, corresponding to the third category. The iconic diagram and its symbolic interpretant "constitute what we shall not too much wrench Kant's term in calling a *Schema*, which is on the one side an object capable of being observed while on the other side it is General." (NEM IV, 318)

Peirce explains:

> It is, therefore, a very extraordinary feature of Diagrams that they *show*, – as literally *show* as a Percept shows the Perceptual Judgment to be true, – that a consequence does follow, and more marvellous yet, that it *would* follow under all varieties of circumstances accompanying the premisses. It is not, however, the statical Diagram-icon that directly shows this; but the Diagram-icon having been constructed with an Intention, involving a Symbol of which it is the Interpretant (as Euclid, for example, first enounces in general terms the proposition he intends to prove, and then proceeds to draw a diagram, usually a figure, to exhibit the antecedent condition thereof) which Intention, like every other, is General as to its Object, in the light of this Intention determines an Initial Symbolic Interpretant. Meantime, the Diagram remains in the field of perception or imagination. (Peirce NEM IV, 318)

There is a sort of vertical interpretation in which the icon is a transitory element between two symbolic ones. The diagram has been constructed according to a general symbolic *intention* – an isomorphism between the symbolic proposition and the icon – and it must, in turn, determine a symbolic interpretant. When we contemplate the diagram, we at once prescind "from the accidental characters that have no significance. They disappear altogether from one's understanding of the Diagram." Nevertheless, this "is only an understood disappearance and does not prevent the features of the Diagram, now become a Schema, from being subjected to the scrutiny of observation" (ibid. 317), so meeting the requirement of experimentation and being responsible for the evidence that only observation can provide.

Now, it is worth noting that such a process is not a mental one grounded on our human psychology, but rather it is the very way in which diagrammatic signs signify. The same would happen for every *quasi-mind* engaged in a deductive inference. For the diagram itself "is an *icon* or schematic image embodying the meaning of a general predicate" (ibid. 238), and the relation it represents – the general predicate – is a *rational* one, an intrinsically general formal relation. It is "not merely one of those relations which we know by experience, but know not how to comprehend, but one of those relations which anybody who reasons at all must have an inward acquaintance with" (ibid. 316). For instance, it would be impossible, Peirce says, to represent in a diagram the mere relation of killer to killed, because it is something that is not intelligible. It is simply known as a fact.

As to the empirical subject, Peirce does not inquire:

> That step of thought which consists in interpreting an image by a symbol, is one of which logic neither need nor can give any account, since it is subconscious, uncontrollable and not subject to criticism. Whatever account there is to be given of it is the psychologist's affair. (Peirce CP 4.479)

At any rate, the process described above in Peirce's words is half of the matter. Beside the vertical interpretation outlined, another one takes place that is, so to say, horizontal. This second interpretation amounts to the permitted transformations of the diagram in which a diagrammatic sign is interpreted by a new one. Both interpretations are required in order to gain the general evidence of the conclusion:

> The Schema *sees*, as we may say, that the transformate Diagram is substantially contained in the transformand Diagram, and in the significant features to it, regardless of the accidents ... The transformate Diagram is the Eventual, or Rational, Interpretant of the transformand Diagram, at the same time being a new Diagram of which the Initial Interpretant, or signification, is the Symbolic statement, or statement in general terms, of the Conclusion. By this labyrinthine path, and by no other, is it possible to attain to Evidence; and Evidence belongs to every Necessary Conclusion. (Peirce NEM IV, 318 f.)

This *labyrinthine path* is the answer that Peirce proposes to the problem of mathematical evidence and generalization. It is a path grounded on the whole of his semiotic system of categories. I would like to stress once again that it does not amount in any way to Berkeley's solution. Berkeley was concerned with an absolute object beyond the sign, while, according to Peirce, it is on the level of the sign itself that we must work. There is a wholly different ontological approach. Nor does Peirce have to commit himself to the ontological assumption of Platonic forms, as Locke's abstract ideas appear to do.

Peirce's philosophy of *synechism*, based on the concept of continuity, brings a completely different ontological commitment. Signs confer reality to the relations studied by mathematics. "I believe I may venture to affirm," Peirce writes, "that an intelligible relation, that is, a relation of thought, is created only by the act of representing it" (ibid. 316). As Michael Otte puts it, "the diagram in mathematics is a machine which permits us to confer reality to certain relations ... From a continuum of real possibilities, some of these are being actualized by means of distinctions" (1997, 362). This being the case, we can see that a great gulf separates Peirce's understanding of mathematical cognition from all previous philosophies, a gulf that is worth exploring in order to clarify all features of his semiotic approach.

Università di Milano

REFERENCES

Berkeley, G. (1957). *A treatise concerning the principles of human knowledge*. Indianapolis-New York: The Bobbs-Merrill Co.

Locke, J. (1910). *An essay concerning human understanding*. London: Routledge.

Otte, M. (1997). Analysis and synthesis in mathematics from the perspective of Charles S. Peirce's philosophy. In M. Otte and M. Panza (Eds.), *Analysis and synthesis in mathematics* (327-362). Dordrecht-Boston-London: Kluwer Academic Publishers.

Otte, Michael (in print). Mathematical epistemology from a semiotic point of view, *Educational Studies in Mathematics*. Special issue on "The teaching and learning of mathematics: Semiotic and epistemological perspectives" (ed. by Adalira Sáenz-Ludlow and Norma Presmeg).

Peirce, C. S. (CP) (1931 – 1935, 1958). *Collected papers of Charles Sanders Peirce*. Cambridge, MA: Harvard UP.

Peirce, C. S. (NEM) (1976). *The new elements of mathematics by Charles S. Peirce* (Vol. I – IV). The Hague-Paris/Atlantic Highlands, N.J. Mouton/Humanities Press.

Robin, R. (1967). *Annotated catalogue of the papers of Charles S. Peirce*. Amherst: University of Massachusetts Press.

MICHAEL H. G. HOFFMANN

SIGNS AS MEANS FOR DISCOVERIES:

Peirce and His Concepts of "Diagrammatic Reasoning," "Theorematic Deduction," "Hypostatic Abstraction," and "Theoric Transformation"

Abstract. The paper aims to show how by elaborating the Peircean terms used in the title *creativity* in learning processes and in scientific discoveries can be explained within a semiotic framework. The essential idea is to emphasize both the role of external representations and of experimenting with those representations ("diagrammatic reasoning"), and to describe a process consisting of three steps: First, looking at diagrams "from a novel point of view" ("theoric transformation") offers opportunities to synthesize elements of these diagrams which have never been perceived as connected before. Second, by forming those observed syntheses to "new objects" of thinking, and by signifying these objects through new signs ("hypostatic abstraction"), new *means of thinking and acting* are created (to be used for "theorematic deductions"). And finally, by applying these new means – in proofs, for instance – the "intelligibility" of new discoveries and their power to explain problematic facts must be tested.

Key words: diagrammatic reasoning, hypostatic abstraction, mathematics, Peirce, semiotics, synthesis, theorematic deduction, theoric transformation.

Semiotic theories agree that "signs" are basically means of signifying an object or means of representing something for somebody. The crucial point of Charles S. Peirce's epistemologically based semiotics, however, is his emphasis on a second, more fundamental function of signs, namely, signs as *means of thought, of understanding, of reasoning, and of learning*. From this point of view, signs are, as in Kant, conditions of these activities. Unlike Kant, however, signs are no transcendental, subjective, or mental conditions of possible experience for Peirce, but conditions for which the difference between internal and external plays no role: "All our thinking is performed upon signs of some kind or other, either imagined or actually perceived. The best thinking, especially on mathematical subjects, is done by experimenting in the imagination upon a diagram or other *scheme*, and it facilitates the thought to have it before one's eyes." Thus, for any "concept" or mental state, "external signs answer every purpose, and there is no need at all of considering what passes in one's mind" (Peirce, NEM I 122).

In this paper, I shall elaborate some details of this semiotic approach. The goal is to clarify, firstly, the role of *external representations* in processes of learning and of scientific discoveries; and secondly, the problem of *interpreting* those representations.[1] The starting point is Peirce's concept of "diagrammatic reasoning," by which one can explain the development of knowledge on the basis of a three-step activity: *constructing* representations, *experimenting* with them, and *observing* the results.[2] The idea is that by representing a problem in a diagram, we can experiment with our own cognitive means, and thus develop them. "The diagram becomes the something

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 45 – 56.

(non-ego) that stands up against our consciousness," as Kathleen Hull puts it; "reasoning unfolds when we inhibit the active side of our consciousness and allow things to act on us" (1994, 287).

In order to explain, however, the genuine *creativity* necessary for each of these three steps, one has to go beyond this concept of diagrammatic reasoning. The thesis of this article is that there are two further concepts developed by Peirce that can offer deeper insights into the possibility of learning and of discovering: Firstly, the concept of "theorematic deduction" by which the "hypostatization" of new proof elements is described; and secondly, the concept of "theoric transformation" by which Peirce described changing the point of view on a problem or a representation. Peirce did not distinguish either concept until after 1907, as far as I can see, a few years before he died (in 1914). In spite of its obvious relevance for understanding creativity in mathematics, the difference between "theorematic deduction" and "theoric transformation" has not been noticed in Peirce scholarship before my recent study on "Knowledge Development."[3]

DIAGRAMMATIC REASONING

Understanding Peirce's notion of "diagrammatic reasoning" presupposes knowing something about what he called a "diagram." The use of this concept is in no ways restricted to "images" or "graphical representations." Based on his highly differentiated semiotic terminology,[4] Peirce defines a "diagram" as "a representamen which is predominantly an icon of relations and is aided to be so by conventions. Indices are also more or less used. It should be carried out upon a perfectly consistent system of representation, founded upon a simple and easily intelligible basic idea" (Peirce, CP 4.418, 1903).

Thus, a diagram is a complex sign in which "indices," whose function is to direct attention to something (also variables in equations are "indices"), and "conventional" signs also play a role (i. e., "symbols," the only signs that have a "meaning" for Peirce; i. e., a law-like relation between the symbol's object and its interpretation). Most important, however, is the *iconic* character of diagrams. An "icon" is defined as a sign that – based on a certain "likeness" to its object – "excites an idea naturally allied to the idea that object would excite" (Peirce, EP II 13). Its function is most of all to represent *relations*, so that not only photos and footprints are icons, but also for example, sentences and algebraic equations. The latter however, belong to a certain subgroup of icons, namely, the *diagrams* we are looking for. The specific difference of "diagrams" in relation to other icons can be seen in the fact that they are "carried out upon a perfectly consistent system of representation," as quoted above. If we are confronted, for example, with the complex sign "Theaetetus-Socrates-stands-sits-and," we could interpret this sign as an *icon*, because it represents a certain relation. But if we read "Theaetetus stands and Socrates sits," we have a *diagram*, because this sign represents a relation that is carried out upon our *grammatical* "system of representation" as defined by syntax.

This definition of "diagram" has some essential consequences for each of the three steps of diagrammatic reasoning mentioned above. Let us begin with the first two steps: Any *construction* of a diagram is carried out by the means of a given representational system, and any *experiment* we perform on it is determined by the rules of that system. Thus, the representational systems of our natural or artificial languages offer all kinds of words and a certain syntax; axiomatic systems like Euclidean geometry formulate definitions, postulates and axioms; neural systems like our brain are defined by electrochemical states and a complex order of connectivity; and so forth.

Representational systems are more or less "consistent," ranging from axiomatic systems in mathematics to the stylistic means of art, whose history for itself might be conceived as a development from very strict regulatory rules to plurality and freedom. Nonetheless, even if representational systems are only partly consistent, their *rational* and *normative* character is essential for what happens in the experiments we perform on diagrams. The inference rules and conventions the mathematician submits to in acting on diagrams define the *limits* of possible transformations, and they define *constraints* that determine – sometimes within a range of possibilities – the outcome of experiments.

These are the first two steps of diagrammatic reasoning. However, the consistency of representational systems plays its most fundamental role in the third step: in *observing* what happens in diagrammatization. In a manuscript titled "Pragmatism" written circa 1905, Peirce highlights as a core idea of this philosophy that all reasonings – and especially mathematical reasonings – "turn upon the idea that if one exerts certain kinds of volition, one will undergo in return certain compulsory perceptions. Now this sort of consideration, namely, that certain lines of conduct will entail certain kinds of inevitable experiences is what is called a 'practical consideration'" (CP 5.9, c. 1905). Such an "inevitableness" depends obviously on the given rules and conventions of the representational system in which such reasoning is performed, as is evident from mathematics: That 2 plus 2 equals *4* results from the rules and conventions of arithmetic as the chosen representational system. The point, however, is that one needs to have really "internalized" the normativeness of representational systems in order to *experience* this inevitableness. We have to be quite sure about the rules and conventions of a chosen representational system to *feel* inconsistencies, for example, or to be *surprised* by what our experimentation with diagrams generates.

This "inevitable experience" resulting from rule-driven activity is the most important precondition of discovering something *new* by diagrammatic reasoning. To demonstrate this, we can distinguish, first of all, two possible cases: On the one hand, we can gain something new by unfolding new *implications* of constructions within a *given* system of representation; and, on the other hand, there is the process of *developing* representational systems *themselves* that can open up new horizons and possibilities.

The first case is based on the consideration that we never can have a complete overview of all the implications of what we know already. Only experimentation with representations in concrete situations reveals what might already be given implicitly in our own systems of knowledge. Peirce described this case of discovering something new by saying that a diagram constructed by a mathematician "puts before him an icon by the observation of which he detects relations between the parts of the diagram other than those which were used in its construction" (NEM III 749, 1901-2). By experimenting upon the diagram and by observing the results thereof, it is possible, as Peirce says, "to discover unnoticed and hidden relations among the parts" (CP 3.363, 1885).

Even more important are "inevitable experiences" that result from the rationality of our representational systems with regard to the second form of discovering something new. It might be that a rule-driven experimentation with diagrams brings to light inconsistencies or undecidable situations within our chosen representational system. In that case, we are forced either to doubt the correctness of the diagrammatic transformations we have performed, or to question the representational system we have used. If we have reasons to follow the latter course of thought, a genuine creativity is demanded.

In this situation, the *compelling* character of diagrams and the "inevitable experience" we make in diagrammatic reasoning are decisive. The results of experiments have to "stand up against our consciousness," as Kathleen Hull (282) puts it, because only in that case can a diagram "compel us to think quite differently," as Peirce said (CP 1.324, 1903). It is only if we have already certain expectations concerning what should happen in processes of diagrammatization that there is a need to develop something new when those expectations are frustrated. This form of "resistance" of diagrams to what we think about them is the most important difference distinguishing the *semiotic* approach of learning from all sorts of "constructivism."[5] Peirce called the necessary experience of "resisting objects" an experience of "secondness."[6] This secondness is responsible for the fact that, in perception, there is an "effect" of "other things" outside ourselves on us that "is overwhelmingly greater than our effect on them" (CP 1.324, 1903). It guarantees the possibility of learning, because it is only if the "realities compel us" (CP 1.383) that we can *transcend* what we already know and what was in the past an undoubted fundament of our "constructions" of the world.

But precisely how can we react on those compulsions? In which way might learning and new discoveries be possible in a situation in which we are "bumping up against hard fact;" in which we "expected one thing, or passively took it for granted, and had the image of it in our minds, but experience forces that idea into the background, and compels us to think quite differently" (CP 1.324, 1903)?

THE HIGHEST KIND OF SYNTHESIS

While the previous considerations have emphasized the *normative* and *compelling* character of the logic of representational systems as a condition to overcome our prevailing and undoubted expectations, it seems obvious that we need a certain *freedom* for the genuine *creative* aspects of learning and discovering. Peirce called this "the highest kind of synthesis," when the mind – "in the interest of intelligibility" – introduces "an idea not contained in the data, which gives connections which they would not otherwise have had." He compares such acts of synthesizing with the creativity of an artist:

> The work of the poet or novelist is not so utterly different from that of the scientific man. The artist introduces a fiction; but it is not an arbitrary one; it exhibits affinities to which the mind accords a certain approval in pronouncing them beautiful, which if it is

> not exactly the same as saying that the synthesis is true, is something of the same general kind. The geometer draws a diagram, which if not exactly a fiction, is at least a creation, and by means of observation of that diagram he is able to synthesize and show relations between elements which before seemed to have no necessary connection. The realities compel us to put some things into very close relation and others less so, in a highly complicated, and in to[7] sense itself unintelligible manner; but it is the genius of the mind, that takes up all these hints of sense, adds immensely to them, makes them precise, and shows them in intelligible form in the intuitions of space and time. Intuition is the regarding of the abstract in a concrete form, by the realistic hypostatization of relations; that is the one sole method of valuable thought. (CP 1.383, 1888)

In this quote we find a very dense sequence of crucial points regarding what happens in discovery processes. Thus, to begin with a first point, the "intelligibility" of synthesis hints at a certain *teleological* moment. As it is clear in mathematics that any creative act of developing the plan of a demonstration, of transforming diagrams, or of formulating adequate lemmas is successful only when, "in the end," a strict proof is possible, the success of any discovery has to be measured by its power to *explain* problematic facts within a holistic system of beliefs. The "interest of intelligibility," thus, is fulfilled when the synthesis of a creative mind is successful in that sense.

A second important point concerns Peirce's identification of the act "to synthesize" with the act to "show relations between elements which before seemed to have no necessary connection." The central point of constructing diagrams and of experimenting with them is that these are the only activities through which we gain "elements" and "relations" to observe. We need something before our eyes in order to discover something "new" in the constructions of what we already presume to know. Any construction of a diagram is, in itself, a creative act in which we express a certain interpretation of the problem at hand. And any such representation of a problem puts signs before our eyes that, once again, can be interpreted in different ways. Thus, we might observe "new" relations in this representation that had played no explicit role in its construction.

The third point I would like to highlight following the quote above concerns Peirce's talk about the creative mind's showing interpretations "in intelligible form in the intuitions of space and time." This seems hard to understand, even when Peirce explains his use of "intuition" here as "the regarding of the abstract in a concrete form, by the realistic hypostatization of relations." These few words, however, touch a central point in Peirce's theory of knowledge development: "Hypostatization" (from the Greek ὑπόστασις) or "*reification*" (in Latin) consists in creating a thing out of what is not a thing; an entity out of an abstraction. This process (Peirce discussed it mostly under the heading of "hypostatic abstraction") is indeed the core of any abstract science like mathematics. Our "natural" numbers are already products of hypostatization, for you do not find "twoness" anywhere but only *pairs of things*, and coining the term "two" is nothing other than *perceiving* pairs of things only with regard to their *being* two or pairs (for more details see below).

Based on this analysis of the quotation above, we can identify three essential aspects within the creative act of *synthesizing* disconnected elements in a diagram that still require further attention. To put them in chronological order, the first step is *observing* new and previously unnoticed "relations" between those elements; the second step is "the regarding of the abstract in a concrete form," that is the *creation* *of a new sign* – an "intelligible form in the intuitions of space and time"[8] – that can *represent* now what we have observed so far; and the third step is testing whether the newly created hypostatic abstraction is really an "intelligible" one, in the sense of the *teleological* element mentioned above.

In the remainder of this paper, I shall discuss the first two of these three points, because they seem to be central for discovering something new. Following a very late development in Peirce's terminology, I shall discuss them under the heading of *"theoric transformation"* and *"abstractional theorematic deduction"* respectively. Before this, however, we have to reconstruct Peirce's differentiation of four kinds of deduction as the framework in which he locates both concepts.

KINDS OF DEDUCTION

A first, and most famous distinction formulated by Peirce is the distinction between "theorematic" and "corollarial" deduction. He defined both, for example, as follows:

> A Necessary Deduction is a method of producing Dicent Symbols[9] by the study of a diagram. It is either *Corollarial* or *Theorematic*. A Corollarial Deduction is one which represents the conditions of the conclusion in a diagram and finds from the observation of this diagram, as it is, the truth of the conclusion. A Theorematic Deduction is one which, having represented the conditions of the conclusion in a diagram, performs an ingenious experiment upon the diagram, and by the observation of the diagram, so modified, ascertains the truth of the conclusion. (Peirce CP 2.267, c. 1903; cf. CP 7.204, 1901)

Besides the fact that "theorematic deduction" seems to be more creative for Peirce than "corollarial deduction," it is not really clear what the essential difference between both kinds of deduction should be. In this situation, Hintikka formulated a very clear interpretation suggesting that "theorematic inference is characterized by the introduction of auxiliary individuals into the argument" (Hintikka, 1983 <1980>, 113, cf. 109 f.). Peirce himself hints, for example, at "subsidiary lines or surfaces, that are not mentioned either in the proposition to be proved nor in previously proved propositions" (NEM III 172, 1911), or at the need to formulate a *lemma* "when it comes to proving a major theorem" (EP, II 96, 1901). A most general definition speaks of introducing "something not implied at all in the conceptions so far gained, which neither the definition of the object of research nor anything yet known about could of themselves suggest, although they give room for it."[10] Thus, there are good reasons to follow Hintikka when he uses modern quantification theory in order to gain a precise criterion for Peirce's theorematic-corollarial distinction: "What makes a deduction theorematic according to Peirce is that in it we must envisage other individuals than those needed to instantiate the premise of the argument. ... a valid deductive step is theorematic if it increases the number of layers of quantifiers in the propositions in questions" (110).

This interpretation, however, was criticized by Kenneth L. Ketner (1985) who claimed Hintikka had neglected the importance of the "visual observation" emphasized by Peirce for theorematic deductions. This is true, but the point Ketner missed is that "visual observation" is just as important for corollarial reasoning, so it cannot be a criterion for distinguishing the two. On the other hand, observation is, according to Peirce, indeed of greatest relevance when discussing the problem of learning and discovering. But neither Hintikka nor Ketner has seen that Peirce attempted to clarify just this point a little further in 1907. I, at any rate, would suggest to use these later writings and to sharpen the distinction between "theorematic deduction" and "theoric transformations" that Pierce formulates only in these writings. This can justify, on the one hand, Hintikka's argument concerning "new individuals," and, on the other hand, that of Ketner regarding the relevance of observation.

Based on a more careful study, I shall distinguish not just "corollarial" and "theorematic deduction" as Hintikka did, but *four* kinds of deductive inference in all:[11]

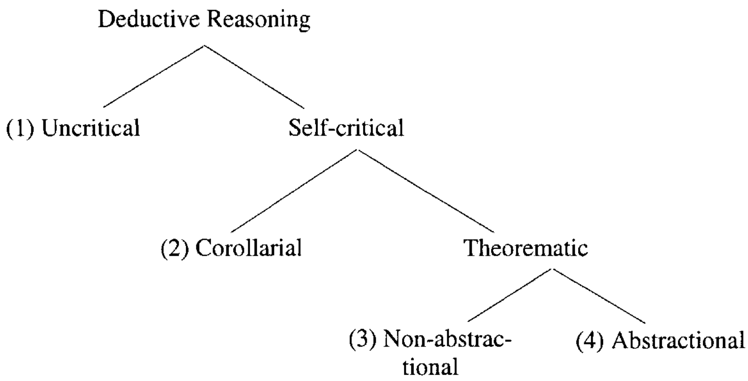


Figure 1. Four kinds of deductive inference

The standard case of "uncritical" deduction is the operation of what Peirce called a "Logical Machine," that is, a machine that processes input-output transformations only on the basis of fixed rules without any freedom to decide between different possible procedures. "Corollarial reasoning" is based only on the definitions of the terms in a proposition to be proved, and uses, besides those definitions, only general principles of logic. In contrast to "uncritical" reasoning, "corollarial" deduction involves a careful analysis of the propositions from which something should be deduced. Therefore, it is possible only as a kind of "self-critical" deduction. But there is no "use of any other construction" than that is implied already in the proposition to be proved (NEM IV 288 f., 1903). "Theorematic deduction," now, goes beyond this limitation of corollarial deduction by introducing "other individuals than those needed to instantiate the premise of the argument," as Hintikka puts the point.

In his "Carnegie Application," Peirce called his distinction between corollarial and theorematic reasoning his "first real discovery about mathematical procedure ... I show that no considerable advance can be made in thought of any kind without theorematic reasoning. When we come to consider the heuretic part of mathematical procedure, the question how such suggestions are obtained will be the central point of the discussion" (NEM IV 49, 1902).

To grasp this "central point" he hints in a few further sentences at the relevance of what he usually called "hypostatic abstraction:"

> Passing over smaller discoveries, the principal result of my closer studies of it [i. e., "the central point"] has been the very great part which an operation plays in it which throughout modern times has been taken for nothing better than a proper butt of ridicule. It is the operation of *abstraction*, in the proper sense of the term, which, for example, converts the proposition 'Opium puts people to sleep' into 'Opium has a dormitive virtue.' This turns out to be so essential to the greater strides of mathematical demonstration that it is proper to divide all Theorematic reasoning into the Non-abstractional and the abstractional. I am able to prove that the most practically important results of mathematics could not in any way be attained without this operation of abstraction. (NEM IV 49, 1902)

Here we have the last distinction I mentioned in Figure 1. Before one can understand its relevance, however, it is necessary to discuss what Peirce meant by the operation of "hypostatic abstraction." This concept is the core of what Peirce calls here "abstractional" theorematic deduction.

HYPOSTATIC ABSTRACTION

Whatever we discover or learn as something "new" can become a *subject* of our considerations only in the form of "hypostatic abstraction" – to use Peirce's terminology. In the operation of hypostatic abstraction we generate new signs signifying objects that were never mentioned before as objects. Thus, all concepts in our languages are outcomes of hypostatic abstraction performed at some time in the long history of our cultures. When we teach our children concepts, we usually try to give them opportunities to repeat this creative operation of hypostatic abstraction; they discover the world by generating their own hypostatic abstractions out of experiences and observations for which they did not have adequate concepts before.

Operations of hypostatic abstractions are particularly important for us in the genesis of mathematical knowledge:

> In order to get an inkling – though a very slight one – of the importance of this operation in mathematics, it will suffice to remember that a *collection* is an hypostatic abstraction, or *ens rationis*, that *multitude* is the hypostatic abstraction derived from a predicate of a collection, and that a *cardinal number* is an abstraction attached to a multitude. So an *ordinal number* is an abstraction attached to a *place*, which in its turn is a hypostatic abstraction from a relative character of a unit of a *series*, itself an abstraction again. Now, ... what you mean by a *concept* is a predicate considered by itself, except for its connection with the word or other symbol expressing it, and now regarded as denotative of the concept. Such a concept is not merely prescissively abstracted, but, as being made a subject of thought, is hypostatically abstract."[12]

The most important point of hypostatic abstraction is that a sign that is generated as a *new object*, and as *referring* to a new object, can be used, in turn, as a means for further operations – also in different contexts. Michael Otte discusses this point in terms of the "complementarity," or the "dialectic of means and objects" as the essence of mathematical activity (1997, 360). He understands "by 'object' any problem or any kind of resistance of reality against the subject's activity, and by 'means' anything which seems appropriate to achieve mediation between the subject and the object of cognition" (ibid.).

Learning, from this point of view, means creating new *objects* by hypostatic abstraction, and using them as new means for mediating between the subject and the object of cognition. These new objects – as "independent" from the subject's activity – are the starting point for further hypostatic abstractions, so that we gain a process that can be grasped as a process of *generalizing* our representational means. A generalization that "echoes" the history of our cultures, as we can learn from Salomon Bochner's considerations concerning what he called "full-scale symbolization" in contrast to mere "idealization":

> ... full-scale symbolization is much more than mere idealization. It involves, in particular, untrammeled escalation of abstraction, that is, abstraction from abstraction, abstraction from abstraction from abstraction, and so forth; and, all-importantly, the general abstract objects thus arising, if viewed as instances of symbols, must be eligible for the exercise of certain productive manipulations and operations, if they are to be mathematically meaningful. (Bochner 1966, 18)

With regard to the distinction between "abstractional" and "non-abstractional" theorematic deduction mentioned in Figure 1, we can conclude that any theorematic deduction in which a new hypostatic abstraction is created in order to formulate a necessary inference can be called "abstractional," whereas a theorematic deduction that uses hypostatic abstractions already given in other contexts is "non-abstractional."

The question that arises at this point, however, is how to explain the possibility of finding adequate hypostatic abstractions, or of creating new ones. What are the conditions of that very creative act? One most important condition I shall discuss now was "hypostatized" by Peirce in his concept of a "theoric transformation of a problem."

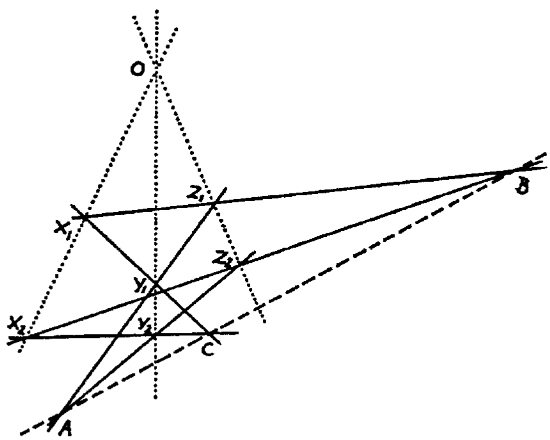


Figure 2. Peirce's diagram of Desargues' theorem (NEM II 212)

THEORIC TRANSFORMATION

Peirce takes the term "theoric" from the Greek "ϑεωρία" (our "theory," original meaning: "vision"), which he translates as "the power of looking at facts from a novel point of view" (MS 318: CSP 50 = ISP 42, 1907). "Theoric" reasoning consists "in the transformation of the problem, – or its statement, – due to viewing it from another point of view" (ibid., CSP 68 = ISP 225). Thus, a "theoric transformation," or a "theoric step" in a deductive argument, means changing the perspective. The relevance of this transformation for all kinds of creativity in mathematics can be seen in the fact that it is the precondition for *perceiving* something new in a certain well-known representation. Peirce hints, for example, at the moment when the "plan of a demonstration ... spring up in the mind ... . The thought of the plan begins with an act of ἀγχίνοια [ready wit] which, in consequence of pre-existent associations, brings out the idea of a possible object, this idea not being itself involved in the proposition to be proved" (CP 4.612, 1908).

As an example, he often hints at the famous proof of Desargues' theorem about two triangles in a projection. According to Figure 2, the theorem can be formulated as follows: Given two triangles $X_1Y_1Z_1$ and $X_2Y_2Z_2$, if the straight lines X_1X_2, Y_1Y_2 and Z_1Z_2 meet in O, then the intersection points C of X_1Y_1 and X_2Y_2, B of X_1Z_1 and X_2Z_2, and A of Z_1Y_1 and Z_2Y_2 lie on the same line ABC.

It seems to be remarkable that these points belong to the same line, but if you *change the point of view*, and perceive the triangles as *planes* intersecting a pyramid with O as apex, then the situation is quite clear, as Figure 3 shows. At least in an intuitive manner, Desargues' theorem can be proved easily by saying that any two planes in space intersect in a line (in the case of parallel planes, the situation is different, of course). In this proof, as Peirce said, everything "is corollarial except the single idea that the plane figure is a projection of a figure in three dimensional space. That is certainly not corollarial, since there is nothing in the problem to suggest it, – no reference to a third dimension" (MS 318: CSP 53 = ISP 45, 1907).

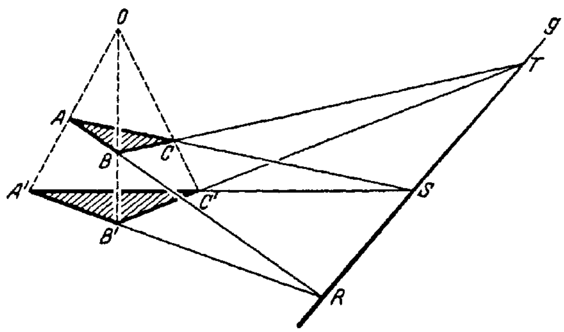


Figure 3. A three-dimensional representation of Desargues' theorem (Hilbert and Cohn-Vossen, 1973 <1932>, 107)

Theoric transformations are the key elements in all kinds of *creative* reasoning. New ideas do not emerge from nowhere in our minds, nor can we generate them out of nothing in acts of pure contemplation. The only way we can discover something new is by *constructing* a diagram to represent a problem, *experimenting* with it according to the rules of the chosen representational system, and *observing* what happens. A constructed diagram might be well-known, but by changing the point of view, and by interpreting it in a new way we can gain new insights. If this is the case, a hypostatic abstraction of what we can see from such a new perspective produces a new object, and thus a new means for further discoveries.

Georgia Institute of Technology, Atlanta, Georgia

NOTES

[1] The background is a more detailed study in German, Hoffmann, 2003a.
[2] Peirce, NEM IV 47 f., 1902. Cf. Stjernfelt, 2000, Hoffmann, 2003b, and in print. Attempts to apply this concept to problems of mathematics education are also formulated by Dörfler, 2004, and by Bakker, in preparation.
[3] Hoffmann, 2003a, Chapter 6.4.
[4] Cf. Hoffmann, 2001a, b, and 2003c, 48-69.
[5] Cf. also my criticism of constructivist approaches in Hoffmann, 2001c, 247 f.
[6] He distinguished three fundamental modes of how something is present to us and called them "firstness," "secondness," and "thirdness." A fuller description is given in Hoffmann, 2001a, b, and 2003c, 59-62.
[7] "to" according to the editors of CP (instead of "the").
[8] In spite of this formulation's similarity to Kant's "a priori forms of pure intuition" (space and time), Peirce obviously means just the opposite here: the concretization of pure or "intelligible forms" *in* a certain (empirical) space and time.
[9] According to Peirce's 1903 classification of signs that means that the conclusion of a deduction is a "Dicisign," i. e., a proposition, that is read on the basis of certain conventions; cf. CP 2.250 f., c. 1903.
[10] NEM IV 42, 49, 1902. Cf. NEM IV 289 f., 1903.
[11] Cf. Hoffmann, 2003a, chap. 6.2 – 6.4 for further argumentation and references. A different reconstruction was developed by Levy, 1997, 103.
[12] Peirce CP 5.534, c.1905. In the last words, Peirce hints at his distinction of two forms of "abstraction": "In geometry, for example, we 'prescind' shape from color," while hypostatic abstraction means "the creation of *ens rationis* out of an *επος πτεροεν* a [winged word] – to filch the phrase to furnish a name for an expression of non-substantive thought" (CP 5.449).

REFERENCES

Bakker, A. (in preparation). *Design research in statistics education. On computer tools and diagrammatic reasoning*. Utrecht, the Netherlands: CDBeta Press.

Bochner, S. (1966). *The role of mathematics in the rise of science*. Princeton: Princeton University Press.

Dörfler, W. (2004). Diagrammatic thinking: Affordances and constraints. *This volume*.

Hilbert, D., & Cohn-Vossen, S. (1973 <1932>). *Anschauliche Geometrie*. Darmstadt: Wiss. Buchges.

Hintikka, J. (1983 <1980>). C. S. Peirce's 'First Real Discovery' and its Contemporary Relevance. In E. Freeman (Ed.), *The relevance of Charles Peirce*. La Salle, Ill.: Hegeler Institut, 107-118.

Hoffmann, M. H. G. (2001a). The 1903 Classification of Triadic Sign-Relations. In J. Queiroz (Ed.), *Digital encyclopedia of Charles S. Peirce*.
Online: http://www.digitalpeirce.org/hoffmann/sighof.htm.

Hoffmann, M. H. G. (2001b). *Peirces Zeichenbegriff: seine Funktionen, seine phänomenologische Grundlegung und seine Differenzierung*.
Online: http://www.uni-bielefeld.de/idm/semiotik/Peirces_Zeichen.html.

Hoffmann, M. H. G. (2001c). Skizze einer semiotischen Theorie des Lernens. *Journal für Mathematik-Didaktik 22(3/4)*, 231-251.

Hoffmann, M. H. G. (2003a). *Erkenntnisentwicklung. Ein semiotisch-pragmatischer Ansatz*. Dresden: Philosophische Fakultät der Technischen Universität (Habilitationsschrift 2002).

Hoffmann, M. H. G. (2003b). Peirce's "Diagrammatic Reasoning" as a Solution of the Learning Paradox. In G. Debrock (Ed.), *Process Pragmatism: Essays on a Quiet Philosophical Revolution*. Amsterdam Rodopi Press, 121-143.

Hoffmann, M. H. G. (2003c). Semiotik als Analyse-Instrument. In M. H. G. Hoffmann (Ed.), *Mathematik verstehen – Semiotische Perspektiven*. Hildesheim: Franzbecker, 34-77.

Hoffmann, M. H. G. (in print). How to get it. Diagrammatic reasoning as a tool of knowledge development and its pragmatic dimension. *Foundations of Science*.

Hull, K. (1994). Why hanker after logic? Mathematical imagination, creativity and perception in Peirce's systematic philosophy. *Transactions of the Charles S. Peirce Society 30*, 271-295.

Ketner, K. L. (1985). How Hintikka Misunderstood Peirce's Account of Theorematic Reasoning. *Transactions of the Charles S. Peirce Society 21*, 407-418.

Levy, S. H. (1997). Peirce's Theoremic/Corollarial Distinction and the Interconnections Between Mathematics and Logic. In N. Houser & D. D. Roberts, & J. van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce*. Bloomington and Indianapolis: Indiana University Press, 85-110.

Otte, M. (1997). Analysis and Synthesis in Mathematics from the Perspective of Charles S. Peirce's Philosophy. In M. Otte, & M. Panza (Eds.), *Analysis and synthesis in mathematics. History and philosophy*. Dordrecht, Boston, London: Kluwer, 327-364.

Peirce. (CP). *Collected papers of Charles Sanders Peirce*. 1931 – 1935, 1958. Cambridge, MA: Harvard UP.

Peirce. (EP). *The essential Peirce. Selected philosophical writings*. Vol. 1 (1867 – 1893), Vol. 2 (1893 – 1913). Bloomington and Indianapolis 1992 + 1998: Indiana University Press.

Peirce. (NEM). *The new elements of mathematics by Charles S. Peirce* (Vol. I – IV). The Hague-Paris/Atlantic Highlands, N. J., 1976: Mouton/Humanities Press.

Stjernfelt, F. (2000). Diagrams as centerpiece of a Peircean epistemology. *Transactions of the Charles S. Peirce Society 36*, 357-384.

WILLIBALD DÖRFLER

DIAGRAMMATIC THINKING

Affordances and Constraints

Abstract. For arriving at a better understanding of the Peircean notion of diagrammatic reasoning there appear to be two complementary ways. One way is to substantiate its impact and relevance by interpreting actual mathematical reasoning as being diagrammatic. This in fact can be done in a great variety of cases. Another way is to exhibit cases of mathematical notions, concepts and arguments which inherently do not lend themselves in a direct way to diagrammatic reasoning. Analyzing those examples will again sharpen and refine the notions of diagram and diagrammatic reasoning. Or, it might possibly point to the necessity of widening those notions to comprise also the manipulation of words and linguistic terms according to specific rules. The latter two topics are the main issues treated in the paper.

Key words: actual and potential infinite, diagram, diagrammatic reasoning, mathematical existence, representation, visualization.

INTRODUCTION

The American philosopher Ch. S. Peirce has proposed that mathematical thinking, reasoning, and argumentation consist widely in the manipulation of and operation with various kinds of diagrams. A number of his statements on this topic can be found in Hoffmann (2002) and, among others, he says:

> It has long been a puzzle how it could be that, on the one hand, mathematics is purely deductive in its nature, and draws its conclusions apodictically, while on the other hand, it presents as rich and apparently unending a series of surprising discoveries as any observational science. Various have been the attempts to solve the paradox by breaking down one or other of these assertions, but without success. The truth, however, appears to be that all deductive reasoning, even simple syllogism, involves an element of observation; namely, deduction consists in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of experimenting upon this image in the imagination, and of observing the result so as to discover unnoticed and hidden relations among the parts. . . . As for algebra, the very idea of the art is that it presents formulae, which can be manipulated and that by observing the effects of such manipulation we find properties not to be otherwise discerned. In such manipulation, we are guided by previous discoveries, which are embodied in general formulae. These are patterns, which we have the right to imitate in our procedure, and are the icons par excellence of algebra. (Collected Papers 3.363)

This far-reaching thesis about diagrammatic reasoning should be analyzed in two ways: The positive one is to look for and study within mathematics instances that can be interpreted as substantiating and explicating Peirce's claim. The negative one is to scrutinize this view by exhibiting kinds and means of mathematical reasoning that cannot be categorized readily as being of that diagrammatic character. The

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 57 – 66.

power of diagrammatic reasoning is supported by the study of specific examples in Dörfler (2003). The present chapter is devoted to finding mathematical concepts, theorems, and proofs that do not lend themselves in a direct way to the use of diagrammatic means. Or, to put it differently, the notion of diagrammatic reasoning is used to analyze different forms and ways of representation in mathematics.

For both research projects, it will be necessary to have at hand a good understanding and conception of what is or can be taken as a diagram in the sense of Peirce. One kind of clarification is offered in the above-mentioned paper from Dörfler (2003). It is very likely the most important point thereby that a diagram lends itself to highly specific operations (like transformations, combinations, constructions) according to conventional rules. Thus, in this context, diagrams are not just static structures that only have to be perceived, but the objects of operations. And it is the latter that determine the meaning of the diagrams. One might therefore think of a diagram as a (type of) inscription together with a system of operations on it. Inscriptions plus system of operations present the core of the respective mathematical concept. An example is the concept of a matrix that is essentially a rectangular schema (of some kind of numbers) together with a system of conventional operations (addition, multiplication, scalar multiplication, transposition, etc.).

For our purposes here, it is very important to make a clear distinction between "diagrams" and all kinds of representations, visualizations, drawings, graphs, sketches, and illustrations as widely used in professional mathematics and in mathematics education as well. Although these might be diagrams in the specific sense used here, this is mostly not the case. This is due to the lack of the constituting operations by which an inscription or visualization becomes only a diagram. These kinds of visual means might be useful in suggesting certain properties of an intended mathematical concept. But mostly this calls for a metaphorical interpretation of the visualization and its complementation by some kind of idealization. Diagrams, in contrast, have to be taken verbally in a strict sense; they are not open to interpretation or metaphoric use as long as one wants to stay with the mathematical concept or method under direct consideration. For instance, in this view, the graphic visualization of notions like continuity or differentiability by Cartesian graphs of functions cannot have the quality of a diagram (just try to carry through a proof of a theorem on these notions by using graphs alone): Neither of these notions is a graphic or operative property of the function graphs as such. There are no operations defined on or for Cartesian graphs that truly express continuity or differentiability. Thus, in this sense, recourse to the usual arithmetic definitions is unavoidable. These $\varepsilon-\delta$ definitions, by the way, can be translated on their part into graphic diagrams by using appropriate straight lines enclosing the graph with arbitrarily close approximation. At many places in mathematics, one takes recourse to verbal illustrations instead of graphic or visual ones, like "... approaches the limit indefinitely." These texts, of course, also cannot play the role Peirce ascribes to diagrams, because they again lack those operations or transformations of which the observation of their outcome and regularities appears to be constitutive for Peircean diagrammaticity.

A certain kind of limitation to diagrammatic reasoning is the complexity of the diagrammatic inscriptions and the operations with and/or on them. Nowadays, this can be overcome in part through the use of computers (CAS). Traditionally, it is handled in mathematics by introducing notions presenting certain regularities in the diagrams and their transformations. Afterwards, one no longer argues directly with the diagrams but with their verbally described properties (for which one might invent another diagrammatic expression). This leads to the development of a language (or theory) in which it is possible to express the results of (observing) the diagrammatic operations and deduce on a conceptual level new properties of the latter. Take, for example, the section on polynomials in any algebra text-book: Basic properties of polynomials (= diagrams) are obtained diagrammatically in the sense of Peirce. But, sooner rather than later, the definitions and proofs make direct use of those properties without referring explicitly to the diagrams. Thus, one obtains that $P(x)-P(a)$ is divisible by $(x-a)$ through diagrammatic reasoning (i. e. calculations) and directly from this: If a is root of $P(x)$, then $P(x)$ is divisible by $(x-a)$, and, therefore, $P(x)$ has a maximum of n roots ($n = \deg P$).

For another example, consider the theory of (combinatorial) graphs (see Bondy & Murty 1976). Here, many concepts stand for possible diagrammatic properties of graphs (like being connected, regular, Hamiltonian, Eulerian, etc.), and theorems state (general) relationships between them. Again, proofs of basic theorems will depend almost exclusively on diagrammatic reasoning; but, in due course, the concepts and their established relationships themselves will be used directly by arguing verbally and conceptually. I will no longer consider this as diagrammatic reasoning, but as reasoning about diagrams and their properties and operations. Other kinds of diagrams might be invented for this purpose (like $Q(x) \mid P(x)$) that can then be used for another level of diagrammatic reasoning. The basic rules for operating with these new diagrams result from and express regularities in the operations with the lower-level diagrams.

But this deliberate shift from diagrammatic to conceptual-verbal reasoning is not our topic here. This is because it does not reflect an inherent limitation to diagrammatic reasoning, but rather an economic substitute for it, which, at least in principle, could be dispensed with. In the following, I shall consider instances in which, in my view, there is an inherent impossibility of diagrammatic reasoning as understood here.

In a sense, the article endeavors in this way to offer a specific interpretation of the notion of diagram, and it is open to debate to which degree this interpretation is in accordance with the stance taken by Peirce. What comes clear, I think, is that different kinds of signs are used in mathematics; diagrams being but one very important one of them. It is not the intent of this article to study the use of other kinds of signs, such as indices, in mathematics. But I shall point to many instances of mathematical reasoning based on linguistic signs (words) and the highly specific discourse with them that I call conceptual-verbal (in contrast to diagrammatic). The study of conceptual-verbal reasoning; its operations, rules, and strategies (in mathematics); and its specific semiotic character and usage of words as signs will be the objective of further research.

To repeat the objective of this article, I should like to point out again that the operative aspect of diagrammatic reasoning is considered to be of central relevance here. This means the transformations of diagrams within a system of conventional rules that also include the inventive construction of new diagrams or of parts of them (see, again, Dörfler 2003). When Peirce emphasizes that all thinking occurs in terms of signs, he nevertheless explicitly warns against conceptions that are not amenable to diagrammatic presentations. Thus, diagrams are rather special signs that permit conclusive and apodictic reasoning. As long as conceptual-verbal reasoning refers to diagrams and their general (diagrammatic) properties, this quality appears to be maintainable. But it might be lost in at least some of the cases analyzed in the following. In these cases, it is the chosen perspective on and the interpretation of the diagrammatic observations that can no longer be based on diagrammatic reasoning and presentations. I pose this positively as the research problem of analyzing these transitions between very different sign systems (like diagrams in a strictly operative sense and verbal-conceptual presentations) and their impact on learning and understanding in mathematics.

For a comprehensive presentation and analysis of the writings of Peirce on the topic of diagrammatic reasoning, the reader is again referred to Hoffmann (2002). Michael Otte deserves the credit for having brought the ideas of Peirce to the attention of mathematics educators through papers like Otte (1997a, 1997b, 1998).

IMPOSSIBILITIES

There is a standard proof that the square of no fraction can be equal to 2. One assumes to the contrary that $(p/q)^2 = 2$ where $(p,q)=1$. Then one observes a series of diagrammatic transformations: $p^2 = 2q^2$ implies $2 \mid p^2$, which implies $2 \mid p$ (this was a result of earlier diagrammatic reasoning); thus $p = 2p_1$, and so $2 \cdot 2 \cdot p_1 \cdot p_1 = 2q^2$, and therefore $2 \mid q^2$, which implies $2 \mid q$ and thus $(p,q)>1$. This is a contradiction to what was assumed. In other words, among all diagrams p/q with the usual operation rules, there is no such diagram the square p^2/q^2 of which is equal to 2 (again according to the agreed rules of equality of these diagrams). I think it is perfectly in accordance with Peirce's statement to consider fractions as diagrams that can be subjected to certain operations (manipulations in Peirce). Thus, this theorem first of all expresses a property of these diagrams. But for that very impossibility, we do not possess a diagram on which we could perform operations that would reflect that impossibility. There are names for that impossibility like $\sqrt{2}$ or, more generally "irrational number." These names, that is, the symbols for specific irrational numbers like $\sqrt{2}, e,$ or π, are not diagrams in the sense used here, because they do not permit manipulations and transformations that lead to an exploration of the notion of irrationality. Observe the difference to the fraction-diagrams with which one can derive a great many properties. There are also other expressions of this impossibility: There is no finite or periodic decimal fraction d with $d^2 = 2$; the lengths s,d for the side and diagonal of a square have no common measure, that is, there are no natural numbers m,n with $ms = nd$. Again, this states the impossibility of certain diagrams as long as one obeys specific transformation rules. But here as well, no diagram is available for the stated impossibility as such. Infinite decimals are clearly not diagrams in our sense, because they do not admit operations and manipulations that can be observed even in imagination in a virtual sense (contrary to finite decimals of any length or periodic decimals). This lack of (specific and generic) diagrams for irrational numbers or relations might be a cause for the fact that comparatively few notions and theorems have been devised in mathematics for these numbers. Perhaps this is not surprising when one takes into account that one can speak and think about mathematical objects mainly through their diagrams (or more generally through their representations). Much of what has been said about irrationals is about "approximation" by rationals; sometimes in a special form (continued fractions), but not directly about irrationals. The latter could be considered as reified or hypostatized impossibilities. Having no diagrams available also means that one cannot calculate with irrationals in the way we do with whole and fractional numbers. One can write $\sqrt{2}+\sqrt{3}$, but we cannot calculate its decimal expansion. Here $\sqrt{2}+\sqrt{3}$ as a sum receives meaning from the "representation" on the number line (which is highly metaphoric and virtual and certainly not diagrammatic) or again via approximating rational sequences $(a_n),(b_n)$ for $\sqrt{2}$ and $\sqrt{3}$ as $\sqrt{2}+\sqrt{3} = \lim(a_n + b_n)$. Thus one takes recourse either to fractions and thus to diagrams or to conceptual-verbal language. The lack of diagrams is also related to a lack of guidance for finding or devising interesting properties and problems about the respective mathematical objects. In a different context, inscriptions like $\sqrt{2}$ can be ascribed a diagrammatic quality. This occurs when they are subjected to the rule system of arbitrary exponentiation whence relations like $\sqrt{2}\sqrt{3} = \sqrt{6}$ result. In any case, such "formulae" are of a diagrammatic quality but not the "irrationality" of, say, $\sqrt{2}$.

A similar situation occurs in the case of transcendental numbers. These irrationals are defined as not being roots of algebraic equations with rational coefficients. This is again the (postulated) impossibility (or in a more ontological jargon, the non-existence) of diagrams of a certain kind (polynomials). For algebraic numbers, the respective algebraic equation presents a diagram for that number which corresponds to the diagrams m/n in the case of fractions (or to the polynomial $nx - m = P(x)$). Manipulations and transformations of these diagrams lead to invariants and regularities that are considered to be properties of the algebraic numbers. For instance, sum and product of algebraic numbers are algebraic numbers. For transcendental numbers, there are no diagrams in this sense, and they can only be "studied" by using that lack. Here again, one has no intuitive or empirical-observational source and guidance for devising properties or qualities of transcendental numbers.

Another case of impossibility is the notion of linear independence in vector spaces. Linear dependence of vectors $v_1,\dots,v_k$ is defined by the possibility of a diagram of the form $a_1v_1+\dots+a_kv_k$ (linear combination) which equals the zero vector and in which not all coefficients a_i are zero. And it is by manipulating these diagrams that theorems about linearly dependent vectors are derived. Linear independence, in contrast, as the impossibility of such a diagram, has no direct diagrammatic expression, although it is related (negatively) to diagrams. And it can only be investigated by studying this negation. In other words, one can show or point to linearly dependent vectors but not (directly) to linearly independent ones. Linear dependence is observable (directly); linear independence is not. This might be (one) reason for the documented difficulties students have with linear independence, for example, with proofs in linear algebra. It is illuminating that diagrammatic expressions for linear independence could also be found successfully in linear algebra, for instance, through the Gauss algorithm or determinants. It is through manipulating the respective diagrams that one can decide about linear independence and turn it into a visible and observable property of diagrams.

This striving for diagrams that reflect nondiagrammatic impossibilities occurs in many places in mathematics. The invention of the fraction symbols is already of this kind, because a (proper) fraction denotes the impossibility of solving $mx=n$ by a whole number x (m,n whole). Just one more example is the famous Kuratowski Theorem in graph theory that characterizes nonplanar graphs by the existence of subgraphs of a certain kind (essentially K_5 and $K_{3,3}$). Here again, the impossibility of diagrams of a given type is shown to be equivalent to the occurrence of specific diagrams that at least in principle, is observable.

ALL AND EVERY

In the mathematical discourse, it is a common practice to speak of all possible or conceivable instances of a mathematical concept and to consider this virtual or ideal totality as a new entity or structure. This is expressed by the common metamathematical terminology that uses verbs such as: "we form," "we construct," "we build," and so forth. The set or structure N of all natural numbers already furnishes an example of such a speech act. There are generic diagrams like lists of strokes or tokens or decimal numbers by which either an arbitrary natural number or the formation of the successor to a number can be presented diagrammatically. This is done by giving a rule that leads from a given diagram (for n say) to the diagram for $n+1$. But the notion of "all natural numbers" cannot be expressed in a diagrammatic way: N has no diagram. Or, in other words, infinity as a never-ending process (of counting on) is amenable to diagrammatic thinking and reasoning but not the actual infinity of all the steps of that process. The potential infinite has a diagram; the actual accomplished infinite does not. The latter can be conceived only in a conceptual-verbal way. One could think of the number line as a diagram of infinity. Yet, here again, only the potential continuation or extension of any finite part has diagrammatic quality, but not the totality of the infinitely long number line.

Another example of this kind is the well-known statement that there are infinitely many prime numbers. Let us first look at Euclid's proof, which runs as follows. If $p_1, p_2, \ldots, p_r$ are all prime numbers, then calculate $p_1 p_2 \ldots p_r + 1$. This number is either prime or divisible by a prime number. In both cases, we find a prime number p that differs from $p_1, \ldots, p_r$. This argument can be interpreted perfectly well to be of a diagrammatic character, because all its steps can be justified by manipulations of appropriate diagrams. For instance, that each number, viewed as a list of strokes, has a prime divisor, is a diagrammatic property of arranging the strokes (or better the dots for that matter) in a rectangular array: It is the smallest side ($\neq 1$) in any such array. Thus, the statement that there are arbitrarily (finitely) many prime numbers is accessible to diagrammatic reasoning. Yet, jumping to the actually infinite totality of "all" prime numbers has no diagrammatic expression and can be expressed only in conceptual language. This is a qualitative leap that might or might not be accepted, as shown by the widespread discussion over the acceptability of the actual infinite in mathematics (see, also, Dörfler 2002). On the other hand, the conceptual-verbal (discursive) formation of N or the set of all primes is of little consequence in mathematics. Possibly because of the lack of appropriate diagrams, there are no theorems about these totalities (as totalities). Apparently, they have no properties. Essentially, all pertinent theorems are limit theorems like the Prime Number Theorem, which could also be formulated and proved in a framework that restricts itself to the potential infinite. The totality of all natural or prime numbers is not needed anywhere as a whole but only their unlimited succession.

In a similar way, discourse about arbitrary and arbitrarily finitely many fractions and all operations with them is backed up by and based on diagrammatic reasoning, because we have generic diagrams and rules for all that. A qualitatively very different and nondiagrammatic level of discourse is introduced by speaking of Q as a ready-made and completed structure. I do not think that this is more than a (feasible!) way of speaking that is still monitored and regulated by the diagrams it purports to speak about.

As we have seen, the conceptual-verbal "formation" of entities comprising all exemplars of objects of a given kind cannot be supported directly by diagrammatic reasoning. We find cases in which one speaks of the totality of all diagrams of a specified sort, as in the case of all natural numbers, whole numbers, fractions or also all triangles, all connected graphs, and so forth. In these cases, it is typical to have a generic diagram (decimal numerals, p/q, etc.) or a generic description of how to produce any diagram of the given sort (as for graphs or triangles). But in mathematics, the discourse about "building" new totalities goes much further. Already in the case of real numbers, we have, strictly speaking, no diagrams for the mathematical objects. Although decimal expansions are a substitute, they lack many important characteristics of diagrams because we can never determine them and use them in an operative way. Considering how irrationals are defined in mathematics (Dedekind cuts, Cauchy sequences, etc.), one could even say that irrationals already depend on actually infinite totalities of fractions and thus of diagrams. Even farther away from diagrammatic thinking are those discursive formations in which there is no diagram of any kind for the objects lumped together. This, in my view, occurs when speaking about, for example, the "space" $C[a,b]$ of all real functions continuous on the interval $[a,b]$. Compare this situation with, for example, the vector space of all polynomials over Q.

This leads us to analyze the concepts of limit and continuity with regard to diagrammaticity. Consider the standard definition: f is continuous at x_0 if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ for all x with $|x - x_0| < \delta$. Here we find diagrams like $|f(x) - f(x_0)| < \varepsilon$ that also play a pivotal role in all pertinent proofs in that these proofs consist to a large extent of manipulations and transformations of inequalities of this kind. Just inspect any textbook of basic calculus. This constitutes the diagrammatic part of the definition and the proofs and theorems depending on it. But the essential feature of this definition is how to interpret the phrases "for every $\varepsilon > 0$" and "there exists a $\delta > 0$." They refer to all (arbitrarily small) positive reals, and I take the view that this reference is highly nondiagrammatic. The arbitrariness of $\varepsilon > 0$ has no diagram in the sense of Peirce that is amenable to manipulations. The intervals on the number line are, at best, vague visual hints to the intended meaning. It is an idealizing discourse (a language game according to Wittgenstein) that has to be added to the diagrams in the definition of continuity and that, in this respect, consists in the nondiagrammatic interpretation of diagrams. The nondiagrammatic character of this definition is underscored by the postulate that ε varies over an actual-infinite set of even uncountably many values. Constructive approaches to real analysis try to avoid this by sticking to the construction of diagrams of a specific kind. But classic analysis is a very complex interplay of diagrammatic operations (with the inequalities) according to established rules and ideal, nondiagrammatic interpretations in purely verbal formulations (speech acts). It is possibly this that makes it so difficult in the learning process. The usual visualizations (e.g. through Cartesian graphs) are not diagrammatic and do not reflect the very essence of continuity: Continuity and differentiability do not have diagrams. Just try to prove that differentiable implies continuous by using graphs alone. It will be impossible, because we do not have a generic "arbitrary" differentiable or continuous graph as a generic diagram for these concepts. What the visualization can offer is, at best, to guide the interpretations of the $\varepsilon - \delta$ definition. More important for mathematical practice is the availability of a calculus that operates on diagrams (function terms) and permits the evaluation of derivatives, antiderivatives, and integrals according to established diagrammatic operation rules (like $\sin' = \cos$, $(x^n)' = nx^{n-1}$, etc). Here again, we find the striving for manipulable diagrams that can be taken to be an accurate reflection of the related nondiagrammatic structures and processes.

The construction of a (formal) axiomatization in the sense of Hilbert's formalist program can be considered as another method of translating a mathematical notion into diagrams. Thus, for instance, there is an axiom system available for the whole structure of all real numbers that, of course, consists of finite formulas together with a logical formalism (first-order predicate logic). These can be viewed as diagrams in the sense intended by Peirce; and proofs, arguments, and theorems are then obtained by manipulating these diagrams and observing the outcomes of the manipulations (the logical deductions). One could therefore interpret (formal) axiomatization as a kind of diagrammatization.

A very informative example of that sort is axiomatic set theory (see Klaua 1979). For the intended objects, the infinite sets, of course, no diagrams are available (not even any visualization or the like) whose analysis could be taken as the study of these objects, as conceivable, at least in principle, for finite sets. The terms and formulas of, say, ZFC then constitute diagrams that do not correspond to sets but to certain relations between them and their basic (postulated) properties. Thus it is a deception to consider that mathematics in any way studies the infinite directly; it only interprets the study of certain diagrams as research into the properties of infinite sets.

THERE EXISTS

In mathematics, one finds very different proofs of the so-called existence of some mathematical object. There are constructive proofs that exhibit a diagram with the postulated properties or show the possibility of constructing such a diagram. Such a construction can be found within a given collection of already available diagrams: For instance, exhibiting a fraction (rational number) that lies between two given fractions is a construction of this kind. A different form of diagrammatic construction starts from a given collection of diagrams and goes on to construct new diagrams with a desired property. The formation of fractions out of natural numbers is of this kind, or, more generally, the construction of a field K_1 containing K in which one can exhibit a root for a given polynomial over the field K (see Dörfler 2003). A special case of this construction is the complex numbers. In constructive mathematics, these are the only existence proofs that are considered to be feasible or reasonable. Generally speaking, all existence proofs based on proofs by contradiction are, in my view, of a nondiagrammatic character, this may be why they pose problems to a student's understanding. In the case of infinite sets, this situation is exacerbated dramatically. The countability of the algebraic numbers is still based on diagrammatic reasoning (as long as it is not viewed as talking about an actual infinity). But to conclude from this that there are infinitely many (or even uncountably many) transcendental numbers cannot be based on diagrammatic reasoning, because it presupposes the actual infinity of all real numbers (and their not being countable). This is not to refuse those arguments, but to point out essential differences and features from the point of view of diagrammatic reasoning. Thus, both of Cantor's "diagonal" methods can be viewed as a kind of diagrammatic reasoning as long as one interprets them within a potentially-infinite stance. The first one shows in a diagrammatic way that any fraction can be reached by this specific counting process. For this, the finite inscriptions (there are no other ones of course) suffice; everything else is discursive interpretation or conventional speech. Similary, the second diagonal argument can be interpreted in a processual way as successively constructing a decimal expansion different from the given ones.

CONCLUSION

Peirce has made it clear that a great and important part of mathematical reasoning is of a specific empirical form based on observing the behavior and the properties of inscriptions considered as mathematical diagrams. This shifts mathematics from an esoteric, abstract, or purely mental activity "down" to a material activity on perceivable and therefore also communicable objects, that is, the diagrammatic inscriptions. Learning mathematics therefore must consist to a great extent in becoming intimately acquainted with these diagrams and their manipulations ("calculation"). Yet, as I have tried to show, there are limitations of different kinds to these activities, and diagrammatic thinking has to be substituted and complemented by conceptual-verbal reasoning. But again, this does not consist in studying abstract objects directly, but in arguing verbally according to agreed upon ways. A main thesis is that an intrinsic lack of diagrams related to a mathematical notion poses what can be called an epistemological obstacle to learning that notion.

Universität Klagenfurt

REFERENCES

Bondy, J. A., Murty, U. S. R. (1976). *Graph Theory with Applications.* New York: North Holland.

Dörfler, W. (2002). Formation of Mathematical Objects as Decision Making. *Mathematical Thinking and Learning* 4.4, 337-350.

Dörfler, W. (2003). Diagrams as Means and Objects of Mathematical Reasoning. To appear in: *Developments in Mathematics Education in German-speaking Countries. Selected Papers from the Annual Conference on Didactics of Mathematics.*

Hoffmann, M. H. G. (2002). *Erkenntnisentwicklung. Ein semiotisch-pragmatischer Ansatz.* Dresden: Habilitationsschrift.

Klaua, D. (1979). *Mengenlehre.* Berlin-New York: Walter de Gruyter.

Otte, M. (1997a). Analysis und Synthesis in Mathematics from the Perspective of Charles S. Peirce's Philosophy. In M. Otte & M. Panza (Eds.), *Analysis and Synthesis in Mathematics. History and Philosophy.* (Boston Studies in the Philosophy of Science 196. Dordrecht, Boston, London: Kluwer, 327-364.

Otte, M. (1997b). Mathematik und Verallgemeinerung – Peirce' semiotisch-pragmatische Sicht. *Philosophia naturalis* 34, 175-222.

Otte, M. (1998). Limits of Constructivism: Kant, Piaget and Peirce. *Science Education,* 7, 425-450.

Peirce, Ch. S. (1931 – 1958). *Collected Papers* I-VIII. Cambridge, MA: Harvard University Press.

FALK SEEGER

NOTES ON A SEMIOTICALLY INSPIRED THEORY OF TEACHING AND LEARNING

Abstract. The present text starts from the assumption that "mediating" and "weaving" are two core concepts that can grasp the still young relationship of educational theory and semiotics. The distinction of these two concepts can be seen as an elaboration of the idea that relations within a signs differ from relations between signs, an idea that runs parallel to the distinction of the coherence and correspondence of sign systems. The thrust of the present paper is to overcome the confrontation of "learning as acquisition" vs. "learning as participation." The complementarity of mediating and weaving can be helpful in formulating roads out of this fruitless confrontation. These concepts may also be helpful to elucidate how the semiotic-psychological approach of L. S. Vygotskij is related to the semiotics of C. S. Peirce

Key words: abduction, acquisition, learning, mediating, participation

Le signe est une fracture qui ne s'ouvre jamais que sur le visage d'un autre signe (Barthes)[1]

In this chapter, I would like to take a simple exploratory idea and see how far one can go with it in the theoretical reconstruction of certain pervasive problems in mathematics teaching and learning. I shall not get very involved in this math-education-related discussion here, as I have expanded on this elsewhere (see, e. g., Seeger 2003). For reasons of brevity, I shall focus on some speculative thoughts regarding the relative positions of Peirce and Vygotskij. The idea is that Peirce and Vygotskij each have delivered an approach based on the central role of sign processes. Although there are lots of commonalities between the two approaches, one can also find fundamental differences. These will be explored on dimensions of a theoretical perspective on learning that are, admittedly, relatively well known. However, it is hoped that well-known things may appear in a new light.

Talking about Peirce and Vygotskij will lead us to *mediating* as opposed to *weaving* as core concepts from semiotic. These two concepts are more or less loosely related to some fundamental positions of Peirce and Vygotskij. I am discriminating here between a level that is intra-semiotic, that is, concerns processes within the very core of the sign, and a level that is inter-semiotic, that is, concerns processes between signs.

This happens to be a distinction that I found could serve as a starting point to elaborate the differences between the semiotic approaches of Lev Semenovich Vygotskij and Charles Sanders Peirce. In reverse, when trying to improve my understanding of the different orientation of their respective approaches, I found it helpful to give a certain substance to mediating and weaving as semiotically inspired con-

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 67 – 76.

cepts. Ultimately, this discussion should hopefully be fruitful for a new perspective on the old discussion on the nature of learning known as the learning paradox and discussed in more modern terms as the contradiction between a view of learning as appropriation versus participation.

From the perspective of a complementarity of contradictory terms, often advocated by Michael Otte (see, e. g., 1984, 1990, 1994), it strikes me that appropriation and participation cannot be separated but seem to form the basic terms of a theory of learning. If we look at the philosophical "schools" or orientations to which appropriation and participation are usually attributed, we find, roughly speaking, that appropriation is typically attributed to a cultural-historical approach, whereas participation is attributed to constructivism, at least as it is related to the construction of taken-as-shared meaning. Whereas, under close scrutiny, this sketch certainly will appear to be much more differentiated it points our attention to the two major approaches in the philosophy of signs: Plato's instrumentalism and Aristotle's representationalism. I shall come to the complementarity of instrumentalism and representationalism at the end of this chapter, pointing out that this complementarity could only be of benefit to the theory of teaching and learning.

What I shall do is to take a number of questions that make it possible to compare the Vygotskian and the Peirceian approach. I shall begin with a discussion on what might be central concepts for Peirce and for Vygotskij. In this discussion, I am not striving to cover all the important topics and grounding concepts of Vygotskij and Peirce.

One important topic is related to signs as means and the ubiquity of signs.

Veresov (1999) argues that Vygotskij's understanding of the sign is essentially connected to his work on children with learning and developmental handicaps. Because Vygotskij understood being handicapped in learning and development as a collapse of the structure of behavior, the goal of helping and therapeutic intervention was the reconstruction of that behavior. Vygotskij's approach was characterised by the idea that the relation of stimulus and response[2] has to be re-mediated, that is, new means have to be found for mediating between the social and physical environment and the activity of the subject. These means are signs. If new signs mediating stimulus and response can be integrated successfully into the structure of behavior, the handicap becomes more or less obsolete – and simultaneously this means that "higher functions," the specifically human functions, are back in operation again. In a sense, it is also the attempt to give back to the handicapped their self-image as humans.

> For higher functions, the central feature is self-generated stimulation, that is, the creation and use of artificial stimuli (*Vygotskij* 1978, 39)

Vygotskij's approach to the role of signs could not be more mean-related, more instrumentalist. Culture, in a sense, is also the aggregate of means, means to develop and to foster higher psychological functions. In a sense, even the social side of Vygotskij's main metaphor for appropriation, the "zone of proximal development" can be seen as a means for appropriation: Adults or "more capable peers" form the zone of proximal development in which learning and development meet (see Vygotskij 1987). In this light, adults and peers appear only as instrumental for appropriation, because they display the next step in the development of an ability and of knowledge. The "power of the sign" for Vygotskij does not so much spring from the sign itself but from using it as a means, from creating new signs arbitrarily to be used as means in novel situations.[3]

Contrary to Vygotskij, Peirce was not primarily interested in signs as means – neither was he overconcerned with the construction of signs. Sign processes for him are the ground of being human expressed in the words "Man is a sign":

> ... it is sufficient to say that there is no element whatever of man's consciousness which has not something corresponding to it in the word; and the reason is obvious. It is that the word or sign which man uses is the man himself. For, as the fact that every thought is a sign, taken in conjunction with the fact that life is a train of thought, proves that man is a sign; so, that every thought is an external sign, proves that man is an external sign. That is to say, the man and the external sign are identical, in the same sense in which the words homo and man are identical. Thus my language is the sum total of myself; for the man is the thought (Peirce CP 5.314)

It is claimed that the semiotic of Peirce is primarily a theory of reading signs (Trabant 1996). In the process of reading and understanding signs, abductive processes are essential.

Peirce is dealing with the "new" in development and in thinking primarily as it appears in connection with a form of logical conjecturing. This logical form does not codify the successful past of human thinking and observation but is directed toward the anticipation and development of the new. But abduction is not only directed toward the discovery of totally new entities. It is the essence of language that abductions are necessary for understanding, just because language and speaking are and remain ambiguous.

What is an abduction? Peirce introduces such a form of logical conjecture as abduction, a third form in addition to deduction and induction (cf., e. g., Shank 1998, Hoffmann 1999). In an abduction, the relation between the facts in the premise and the conclusion is not necessarily only singular; but manifold relations are imaginable:

> An originary Argument, or Abduction, is an argument which presents facts in its Premiss which present a similarity to the fact stated in the Conclusion, but which could perfectly well be true without the latter being so, much more without its being recognized; ... For example, at a certain stage of Kepler's eternal exemplar of scientific reasoning, he found that the observed longitudes of Mars, which he had long tried in vain to get fitted with an orbit, were (within the possible limits of error of the observations) such as they would be if Mars moved in an ellipse. The facts were thus, in so far, a likeness of those of motion in an elliptic orbit. Kepler did not conclude from this that the orbit really was an ellipse; but it did incline him to that idea so much as to decide him to undertake to ascertain whether virtual predictions about the latitudes and parallaxes based on this hypothesis would be verified or not. This probational adoption of the hypothesis was an Abduction. (CP 2. 96).

While induction shows the factual given-ness and deduction the logical necessity, abduction expresses only a possibility:

> Deduction proves that something must be; Induction shows that something actually is operative; Abduction merely suggests that something may be. ... Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea; for induction does nothing but determine a value, and deduction merely evolves the necessary consequences of a pure hypothesis. (CP 5.171)

Now, it is interesting to ask how the idea of an abduction is connected to another aspect of Peirce's thinking characterized by Roman Jakobson in admirable simplicity as follows:

> One of the most felicitous, brilliant ideas which general linguistics and semiotics gained from the American thinker is his definition of meaning as 'the translation of a sign into another system of signs' (4.127). (Jakobson 1985, 251)

Jakobson is pointing to the elaboration of the notion of Interpretant by Peirce that is closely related to Jakobson's view of the importance of translation:

> The problem of translation is indeed fundamental to Peirce's views and can and must be utilized systematically. Notwithstanding all the disagreements, misunderstandings, and confusions which have arisen from Peirces's concept of "interpretants," I would like to state that the set of interpretants is one of the ingenious findings and effective devices received from Peirce by semiotics in general and by the linguistic analysis of grammatical and lexical meanings in particular. The only difficulty in the use of these tools lies in the obvious need to follow Peirce's careful delimitation of their different types and "to distinguish, in the first place, the Immediate Interpretant, which is the Interpretant as it is revealed in the right understanding of the sign itself, and is ordinarily called the meaning of the sign" (4.536): such an Interpretant of a sign "is all that is explicit in the sign itself apart from its context and circumstances of utterance" (5.474). I do not know of a better definition. This "selective" Interpretant, as distinguished from the 'environmental' one, is an indispensable but all too frequently overlooked key for the solution of the vital question of general meanings in the various aspects of verbal and sign languages. (Jakobson 1985, 251)

Distinguishing between an immediate, "selective" and a "contextual," "environmental" interpretant in a way reflects an old problem in the conceptualization of learning, that is, the notion that learning has to be understood also as learning *about* learning. I shall not go into the details of discussing whether Peirce actually was discriminating between an "immediate" and a "contextual" interpretant. As we shall see below, he was thinking about the role of "collateral experience" in semiotic processes.

In what follows, I would like to present a diagram that tries to capture the idea of a contextual interpretant as a variation. This variation attempts to picture an aspect of learning situations that is of such fundamental importance that it can be called a necessary and critical feature of human learning. I am talking here about metacognition, metaknowledge, decentering – something that makes learning possible in the first place and thus can be called a fundament for education in schools. With his vision of "deutero learning," Gregory Bateson (1972) drew attention to a characteristic feature of learning processes, namely, that progress in learning always aims at "learning to learn." Or, in other words, progress in learning is made possible because the learner can act from a higher or meta-level on to a previous level of learning. If one tries to imagine how climbing to a "higher" level might be achieved, it seems plausible to assume that this works according to a "metaphoric" principle as "something is seen as something else." This, in a certain sense, can be understood as a link to Peirce's idea of abduction.

What conceptual development has to achieve here is, basically, to show that the sign perspective or the semiotic view or, in Merlin Donald's terms (1991), the view of culture as representational make it possible to describe the centered, basic, elementary process of meaning making as well as the secondary process operating on the elementary process and thus express "the meaning of meaning."

If one understands, as is suggested above, the abductive process as a metaphor that includes viewing abduction as managing to translate something, as managing to "see something as something else," then it seems possible to view context as a decisive moment enabling metaphorical reflection. Here, too, Bateson (1972) has prepared the ground for understanding "higher" forms of learning or the development of the new in terms of the "development of contexts."

In Peirce's thought, we find context in relation to the Interpretant primarily connected to what he called "collateral experience," as he writes in a letter to William James:

> Now let us pass to the Interpretant. I am far from having fully explained what the Object of a Sign is; but I have reached the point where further explanation must suppose some understanding of what the Interpretant is. The Sign creates something in the Mind of the Interpreter, which something, in that it has been so created by the sign, has been, in a mediate and relative way, also created by the Object of the Sign, although the Object is essentially other than the Sign. And this creature of the sign is called the Interpretant. It is created by the Sign; but not by the Sign quâ member of whichever of the Universes it belongs to; but it has been created by the Sign in its capacity of bearing the determination by the Object. It is created in a Mind (how far this mind must be real we shall see). All that part of the understanding of the Sign which the Interpreting Mind has needed collateral observation for is outside the Interpretant. I do not mean by "collateral observation" acquaintance with the system of signs. What is so gathered is not COLLATERAL. It is on the contrary the prerequisite for getting any idea signified by the sign. But by collateral observation, I mean previous acquaintance with what the sign denotes. Thus if the Sign be the sentence "Hamlet was mad," to understand what this means one must know that men are sometimes in that strange state; one must have seen madmen or read about them; and it will be all the better if one specifically knows (and need not be driven to presume) what Shakespeare's notion of insanity was. All that is collateral observation and is no part of the Interpretant. But to put together the different subjects as the sign represents them as related – that is the main of the Interpretant-forming. Take as an example of a Sign a genre painting. There is usually a lot in such a picture which can only be understood by virtue of acquaintance with customs. The style of the dresses for example, is no part of the significance, i.e. the deliverance, of the painting. It only tells what the subject of it is. Subject and Object are the same thing except for trifling distinctions. [---] But that which the writer aimed to point out to you, presuming you to have all the requisite collateral information, that is to say just the quality of the sympathetic element of the situation, generally a very familiar one – a something you probably

> never did so clearly realize before – that is the Interpretant of the Sign, – its "significance" (The Essential Peirce, Vol. 2, 493-494)

What would the well-known triad of sign-object-interpretant (see *Figure 1*) look like if we were to try to incorporate context into the diagram for that triad?

If we were to try to incorporate "context" into that diagram, we would be wanting to express the specific quality of the interpretant to become the object of another triad. This seems to be exactly the point characterizing a learning meta-perspective because learning is made an object of learning.

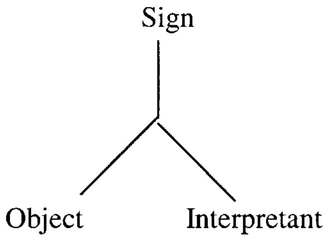


Figure 1. The semiotic triad according to Peirce

A diagram capturing this specific situation might look like *Figure 2*.

The diagram illustrates that a meta-perspective can primarily be taken because the interpretant is changing its position: Now, it has become the object of another contextual triad with yet another interpretant. However, if we look closely at the following excerpt from Peirce, we can see that this quality of the "meshing" of triads seems to be an effect of the fact that, ultimately, the "meaning of a representation can be nothing but a representation":

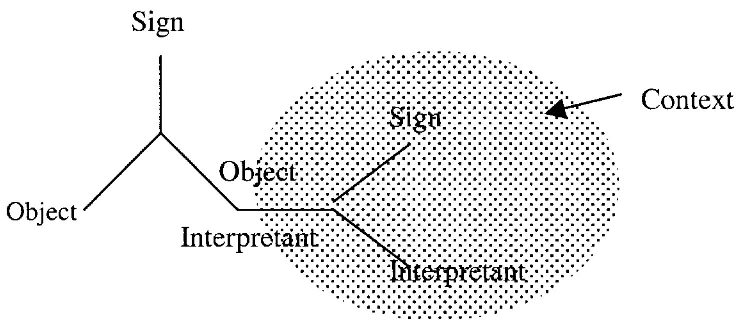


Figure 2. A "metaphorical" triad

> A sign stands for something to the idea which it produces, or modifies. Or, it is a vehicle conveying into the mind something from without. That for which it stands is called its object; that which it conveys, its meaning; and the idea to which it gives rise, its interpretant. The object of representation can be nothing but a representation of which the first representation is the interpretant. But an endless series of representations, each representing the one behind it, may be conceived to have an absolute object at its limit. The meaning of a representation can be nothing but a representation. In fact, it is nothing but the representation itself conceived as stripped of irrelevant clothing. But this clothing never can be completely stripped off; it is only changed for something more diaphanous. So there is an infinite regression here. Finally, the interpretant is nothing but another representation to which the torch of truth is handed along; and as representation, it has its interpretant again. Lo, another infinite series. (A Fragment, CP 1.339, Not dated)

Something important becomes apparent when the original Peircian triad is extended as in *Figure 2*. It shows only the first step of an endless spreading of these meshed triads: The triads are woven together in infinite processes of semiosis. An end of the actual and potential weaving cannot be determined.

This web-like structure, this weaving, surprisingly echoes much of the idea of a *rhizome* put forward by Gilles Deleuze.[4] For the present discussion, only the following features of rhizomatic structures should be briefly mentioned: Each point of a rhizome can be connected to any other point; a rhizome starts to grow from the middle, but it has no center; if a rhizome is cut or interrupted, it continues to grow at any given place. The rhizomatic form of representation *par excellence* is the map: representations are no longer layered and hierarchically organized, but spread over on the flat surface of the map.

> In contrast to centered (even polycentric) systems with hierarchical modes of communication and pre-established paths, the rhizome is an acentered, non-hierarchical, non-signifying system without a General and without an organizing memory or central automation, defined solely by a circulation of states. (Deleuze & Guattari 1987, 21)

In conclusion, I would like to come back to the *problématique* mentioned at the outset. It seems that the dichotomy of the reception and production of signs closely connected to the antagonistic approaches of representationalism and instrumentalism ultimately does not make sense. Obviously, it is not so exciting to continue the age-old discussion between a Platonic approach, emphasizing the instrumental character of signs and the Aristotelian approach, emphasizing the representational nature of signs (see, e. g., Keller 1995). It would be more interesting to find and define perspectives, that build on the complementarity of these two paradigms. Here, these paradigms have been closely connected with the work of Vygotskij and Peirce; Vygotskij building his approach of semiotic mediation on the idea of thinking and acting as fundamentally mediated by signs, and Peirce putting forward the idea of an endless semiotic web based on the incessant activity of the interpretant.

If we manage to stop being trapped in the familiarity of the juxtaposition of representation and instrument, of the reception and production of signs, we can begin to ask new questions and do research in new fields. It is quite obvious that abduction, to take an example, is not only a passive process of correctly "reading" the signs. Already here we can see that "producing" new elements, new approaches, also plays a role – and it is certain that abduction cannot fully be understood until the interplay between the productive and the receptive, the representational and the instrumental character of signs is better understood.

I would like to illustrate this aspect with another example. This example will also touch upon another aspect that seems to be of key importance for a semiotically inspired theory of teaching and learning. Briefly, this aspect comes into focus if one takes into account that the semiotic relations, the complementarity of mediating and weaving and of producing and reading signs, have to be brought alive by real persons. It will not be enough to point out that everything is connected to everything else; it is the quality of the relation that will be decisive.

This point can be illustrated by taking an example from developmental psychology and psychopathology. With the development of his theory of attachment, John Bowlby (1969, 1973, 1980) has laid the ground for a very fruitful exploration of the reasons why socialization sometimes fails. It seems that being closely and securely attached to a family is the decisive factor for growing up positively. More recent studies are increasingly asking how this attachment relation is perceived by children and adults. The "secure base" that children can attach to and the attachment relations themselves have to be represented by a corresponding "internal working model" (see, e. g., Bretherton & Mulholland 1999). It is obvious that neither is independent from the other: Good attachment relations co-occur with a good internal working model and vice versa.

In his work, Michael Otte has often pointed out that the mathematics teacher has to be, as Gramsci coined it, an *exemplary intellectual* (see, e. g., Otte 1994). This imperative is another expression of the necessity for a "secure base": for the relations of students to their teachers as well. I do not want to say that there has to be an attachment to teachers like the attachment to parents and family – although in the first grades, attachment-like relations to teacher tend to be the rule rather than the exception. The relation of the teacher to the students should be one in which the teacher takes over responsibility for providing orientation in the endless weaving by importing authentic "collateral experience" as Peirce would say. It seems a far cry from a "secure base" in learning, if students are taught to construct their own meaning all the time. The assumption that this might be possible at all, underestimates the creativity of the sign processes operating even in the reception of signs.

At the same time, mediating and re-mediating the relations to the "secure base" by using signs as instruments, seems essential. I feel that it would be most rewarding to put more efforts into research designed to reveal the semiotic processes, the complementarity of mediating and weaving.

Institut für Didaktik der Mathematik, Universität of Bielefeld

NOTES

[1] Barthes 1981, 76

[2] Although Vygotskij uses the two classical categories of behaviorism here, he is nothing less than a behaviorist. For historical reasons and for reasons lying beyond the scope of this article, Vygotskij is using

here the categories of Pavlovian reflexology while simultaneously criticizing the approach (for more detail cf. van der Veer & Valsiner 1991; Veresov 1999)

[3] It is interesting to note that from this vantage point, Vygotskij's educational-psychological approach has a very strong "constructivist," productive orientation. This is surprising, because the dominant tone of the constructivist critique on Vygotskij had been that his conception of appropriation was too much based on the notion of a "reception" of knowledge instead of its construction.

[4] But also other approaches could be mentioned here, such as the concept of dissipative structures put forward in the context of a theory of self-organization by Prigogine (see Nicolis & Prigogine 1977; Prigogine & Stengers 1981) or similar ideas on the manufacturing of social order without a central steering ordering power formulated by Bourdieu (1979).

REFERENCES

Barthes, R. (1981). *Das Reich der Zeichen.* Frankfurt/Main: Suhrkamp.

Bateson, G. (1972). *Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology.* San Francisco: Chandler.

Bourdieu, P. (1990). *Outline of a theory of practice.* Cambridge: Cambridge University Press.

Bowlby, J. (1969 – 1980). *Attachment and loss.* Vol. 1 *Attachment* (1969), Vol. 2 *Separation* (1973), Vol. 3 *Loss* (1980). London: Hogarth Press.

Bretherton, I., & Mulholland, K. A. (1999). Internal working models in attachment relationships: A construct revisited. In J. Cassidy & Ph. R. Shaver (Eds.), *Handbook of Attachment: Theory, research, and clinical applications.* New York: Guilford Press, 89-111.

Deleuze, G., & Guattari, F. (1987). *A Thousand Plateaus. Capitalism and Schizophrenia* (Brian Massumi, Trans). Minneapolis: University of Minnesota Press.

Donald, M. (1991). *Origins of the Modern Mind. Three Stages in the Evolution of Culture and Cognition.* Cambridge, MA: Harvard University Press.

Hoffmann, M. H. G. (1999). Problems with Peirce's concept of abduction. *Foundations of Science* 4(3), 271-305.

Jakobson. R. (1985). A few remarks on Peirce, pathfinder in the science of language. In R. Jakobson, *Selected Writings, Vol. VII: Contributions to Comparative Mythology. Studies in Linguistics and Philosophy,* 1972 – 1982. Berlin: Mouton Publishers, 248-253.

Keller, R. (1995). *Zeichentheorie – Zu einer Theorie semiotischen Wissens.* Tübingen: Francke.

Nicolis, G., & Prigogine, I. (1977). *Self-organization in nonequilibrium systems: From dissipative structures to order through fluctuations.* New York: Wiley.

Otte, M. (1984). Komplementarität. *Dialektik* 8, 60-75.

Otte, M. (1990). Arithmetics and geometry – Some remarks on the concept of complementarity. *Studies in Philosophy and Education* 10, 37-62.

Otte, M. (1994). *Das Formale, das Soziale und das Subjektive – Eine Einführung in die Philosophie und Didaktik der Mathematik.* Frankfurt/Main: Suhrkamp.

Peirce Edition Project (Ed.) (1998). *The Essential Peirce.* Volume 2 (1893 – 1913). Bloomington, IN: Indiana University Press. *(EP)*

Prigogine, I., & Stengers, I. (1981). *Dialog mit der Natur: Neue Wege naturwissenschaftlichen Denkens. München*: Piper

Seeger, F. (2002). Research on discourse in the mathematics classroom: A commentary. *Educational Studies in Mathematics* 46, 287-297.

Seeger, F. (2003). Entwicklung und Vernetzung als Grundbegriffe einer semiotisch inspirierten Theorie des Lernens. In M. H. G. Hoffmann (Ed.), *Mathematik verstehen – Semiotische Perspektiven.* Hildesheim: Franzbecker, 119-143.

Shank, G. (1998). The extraordinary powers of abductive reasoning. *Theory and Psychology* 8(6), 841-860.

Trabant, J. (1996). *Elemente der Semiotik.* Tübingen: Francke.

van der Veer, R., & Valsiner, J. (1991). *Understanding Vygotsky.* Oxford: Blackwell.

Veresov, N. (1999). *Undiscovered Vygotsky. Etudes on the pre-history of cultural-historical psychology.* Frankfurt/Main: Peter Lang.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1987). Thinking and speech (Norris Minick, Trans.). In R. W. Rieber and A. S. Carton (Eds.), *The collected works of L. S. Vygotsky: Vol. 1. Problems of general psychology.* New York: Plenum Press, 39-285. (Original work published 1934)

MARIA G. BARTOLINI BUSSI – MARIA ALESSANDRA MARIOTTI – FRANCA FERRI

SEMIOTIC MEDIATION IN THE PRIMARY SCHOOL:

DŰRER'S GLASS

Abstract. "During the seventeenth century geometrical perception became separated, so to say, into two relatively distinct forms of geometry, into two different geometrical styles. One of these is represented by the work of Descartes: the geometry of mechanical-metric activity. The straight line in Cartesian geometry corresponds to an axis of rotation or to the stiffness of a measuring rod. The other geometrical style is represented by the work of Desargues. The straight line of Desarguesian geometry is the ray of light or the line of sight." These sentences were written by Michael Otte in 1997. Aim of this paper is to present the rationale, design and early findings of a teaching experiment, where the Desarguesian form of geometry was approached at by 5th graders, through the use of a cultural artifact and guidance of the teacher.
Key Words: cultural artefacts, Dürer's glass, embodiment, geometry, polyphony, polysemy, primary school, semiotic mediation, teaching experiment, visual pyramid.

1. INTRODUCTION

The experiment presented in this paper concerns the Desarguesian form of geometry.

As found in a previous experiment (Bartolini Bussi 1996) replicated several times, this field of experience (Boero et al. 1995) allows even younger pupils to construct a germ-theory within which they may produce examples of theorems (Mariotti et al. 1997): in this case, the theory is based on a single axiom, i. e. the conservation of straight lines in the projections from one plane to another. Aim of this paper is to present the rationale, design and early findings of a new teaching experiment in a 5th grade classroom: the activity developed in the previous experiment is enriched with the introduction of a big-size model of Dűrer's glass, i. e. an instrument for perspective drawing, reconstructed by Marcello Pergola, in the Laboratory of Mathematical Machines of the Department of Mathematics of the University of Modena (Bartolini Bussi et al. 1999 b).

The theoretical framework draws on works by Vygotsky and Bachtin with additional elements coming from activity theorists (such as Engestroem and Wartofsky). The above elements are, however, filtered so as to meet the needs of designing and analysing effective teaching experiments in the mathematics classrooms.

M. H. G. Hoffmann, J. Lenhard, F. Seeger (ed.). Activity and Sign – Grounding Mathematics Education. Festschrift for Michael Otte, 77 – 90.

2. THEORETICAL FRAMEWORK

In this section we shall sketch very briefly (a longer version is in preparation) the main elements of a theoretical framework of classroom activities centred on artifacts. The crucial issue is given by the transposition of the theoretical construct of *semiotic mediation* (Vygotsky 1974; 1987; 1992) into educational design and classroom implementation.

Three different poles are to be considered in a didactic application of the Vygotskian construct of semiotic mediation:

- the cultural-historical pole, to describe the features of technical and psychological tools which have the potentiality of creating "new forms of a culturally-based psychological process" (Vygotskij 1987, 64).
- the didactic pole, to describe the way of designing, implementing and analysing processes of semiotic mediation;
- the cognitive pole, to describe the process of internalisation of interpsychological activity, that creates the plane of individual consciousness.

2.1. First pole: the cultural historical perspective

Vygotskij distinguishes between the function of mediation of *technical tools* and that of *psychological tools* (or *signs* or *tools of semiotic mediation*), discusses on their relation and offers a list of examples:

> language, various systems for counting, mnemonic techniques, algebraic symbol systems, works of art, writing, schemes, diagrams, maps, and mechanical drawings, all sorts of conventional signs and so on (Vygotskij 1974, 227).

However, as Engestroem (1987) writes, "the exciting relations between technical and psychological tools were not elaborated concretely by Vygotsky." To deepen the discussion, Engestroem explicitly refers to Wartofsky's (1979) discussion about cultural artifacts (i. e. primary, secondary and tertiary artifacts). with an explicit identification of technical tools with primary artifacts and of psychological tools with secondary artifacts (Engestroem 1987, 62).

> What constitutes a distinctively human form of action is the creation and use of artifacts, as tools, in the production of the means of existence and in the reproduction of the species. *Primary artifacts* are those directly used in this production; *secondary artifacts* are those used in the preservation and transmission of the acquired skills or modes of action or praxis by which this production is carried out. Secondary artifacts are therefore representations of such modes of actions (Wartofsky 1979, 200 ff.).

There is also another class of artifacts (*tertiary artifacts*)

> … which can come to constitute a relatively autonomous 'world,' in which the rules, conventions and outcomes no longer appear directly practical, or which, indeed, seem to constitite an arena of non-practical, or 'free' play or game activity. This is particularly true … when the relation to direct productive or communicative praxis is so weakened, that the formal structures of the representation are taken in their own right as primary, and are abstracted from their use in productive praxis (Wartofsky 1979, 208 ff.).

Mathematical theories are examples of tertiary artifacts, organizing the models constructed as secondary artifacts. Mathematical theories have the potential of being expanded to create something anew, that maintains links with practical and representative activities. Cartesian Geometry and Desarguesian Geometry are examples of tertiary artifacts which, as Otte clearly discussed (1997) because of their origins, constituted two mutually exclusive theoretical systems for nearly two centuries.

However, the links with the origins can be neglected and mathematical theories correspond to a monological form of knowledge, as defined by Bachtin, who distinguishes between the case of exact and human sciences, as monological and dialogical forms of knowledge.

> Exact science is a monological form of knowledge: the mind contemplates a *thing*, on which pronunciation is performed: here there is only a subject: he who knows (contemplates) and speaks (pronounces). In front of him there is only the *dumb thing* (Bachtin 1988, 377).

This statement can certainly be applied to most of the 20th-century mathematics treatises (in Bourbaki's style). A different experience is made when the origins of some pieces of mathematics can be reconstructed, together with the primary and secondary artifacts that have characterised their historical development. In this way mathematical activity loses its monological feature and acquires

> [....] the specific task of re-establishing, transmitting and interpreting other people's discourse (Bachtin 1979, 160).

In this perspective, the practical, representative and theoretical aims are supposed to be (at least potentially) embodied in the activity with the same artifact which, in this way, acquires polysemy or multivoicedness (Engestroem 1990). This has two consequences for education. On the one hand, the presence of an artifact does not mechanically determine the way in which it is actually used and conceived of by the students; on the other hand, the presence of an artifact may call to life different aims of the activity, through specific tasks and guidance of the teacher.

2.2 Second pole: the didactical perspective

The polysemy or multivoicedness of cultural artifacts makes them good candidates to rouse and sustain mathematical discussions in the classroom. The term *mathematical discussion* has been introduced by Bartolini Bussi (1996):

> A mathematical discussion is a polyphony of articulated voices on a mathematical object (e. g. a concept, a problem, a procedure, a structure, an idea or a belief about mathematics) that is one of the motives of the teaching-learning activity. The term voice is used following Bachtin, to mean a form of speaking and thinking, which represents the perspective of an individual, i. e. his/her conceptual horizon, his/her intention and his/her view of the world.

Different activities can be designed, which are centred on an artifact, whose introduction defines the teacher's role as follows: *exploiting the (potential) polysemy of the artifact, by constructing occasions and supporting the articulation of different voices.*

The teacher's role in exploiting the artifact polysemy may be rephrased also in terms of *semiotic mediation.* This process is started when the teacher intentionally articulates a primary artifact (e. g., a concrete instrument to be handled in the solution of a problem, like abacus, compass, drawing instrument, gear, perspectograph) and a secondary artifact (e. g., a text or a system of signs, describing how and why it should be constructed and used). Whilst the primary artifact is initially a technical tool (oriented outwards). the secondary artifact may become a psychological tool (oriented inwards):

> The use of signs leads humans to a specific structure of behaviour that breaks away from biological development and creates new forms of a culturally-based psychological process. (Vygotsky 1987).

In this way, conditions are created for the appropriation of those tertiary artifacts, historically rooted in the praxis with those primary and secondary artifacts.

2.3 Third pole: the cognitive perspective

From the individual perspective, the appropriation of tertiary artifacts may be described as the construction of mathematical meanings and ways of thinking. The most powerful Vygotskian instrument for analysing individual cognitive processes is *internalisation* (Vygotskij 1974, 200 ff.) Elsewhere, the same author emphasizes the need to combine three different contributions:

> Children solve practical problems helping themselves with language, eyes and hands. This unit of perception, language and action, which in the end produces interiorization of the visual field, constitutes the central theme for all types of analyses regarding the origin of exclusively human forms of behaviour (Vygotskij 1987, 45).

This internal visual field is a part of the student's internal context where to carry on mental experiments, also supporting the production of mathematical reasoning. This emphasis on the body (eyes, hands and action) is consistent with a recent position in cognitive science, according to which mathematical ideas are, to a large extent, grounded in sensory-motor experience (Lakoff & Nunez 2000).

3. TOWARDS CLASSROOM EXPERIMENTS ABOUT DESARGUESIAN STYLE: A PRIORI ANALYSIS

The pivot of our analysis is a primary artifact, i. e. Dűrer's glass, that is considered as the germ from which secondary and tertiary artifacts have been historically produced.

3.1 Dürer's glass as a primary artifact

Since the 15th century, the production of illusionistic plane (2-D) representation of objects (3-D) has often been realised by means of physical instruments that help the painter to use one eye only, keeping it fixed when painting. The simplest perspectograph was composed by an eyehole and a transparent screen (Figure 1) where the painter traced directly the apparent contour of the object. This simple device is the ancestor of a rich family of instruments for perspective, which

Figure 1. Dürer's glass.

introduced additional equipment to expand its potentialities or to solve some practical problems (Bartolini Bussi, & Mariotti 1999; Field 1997).

3.2 From Dürer's glass to secondary artifacts

This early specimen of perspectograph was described in Alberti's and Dűrer's treatises. The descriptions are mixtures of rules for construction, rules for use and justifications of its functioning. From the ancient treatises (Piero della Francesca, *De Prospectiva Pingendi*, 1460; A. Dűrer, *Underweysung der Messung mit Zirkel und Richtscheit,* 1525; L. B. Alberti, *De Pictura*, 1540). texts may be easily extracted, speaking different voices, representing

– The need from which the production of primary artifacts is generated.

> It is necessary to be able to align on the plane in its own form everything man intends to do (Piero della Francesca).
>
> I don't believe infinite labour is required on the part of the painter, what one expects is a type of painting which looks raised and faithful (L. B. Alberti).

– The explanation of the ways in which the instrument can be built and used:

> A hole where you can place an eye, and this will allow you to see better (A. Dűrer).
>
> What you see through the glass inside the frame will be represented on the glass with the aid of a brush. This is advisable to all those wishing to portray somebody without being sure about their competences (A. Dűrer).

– Justification of functioning, by means of a mathematical model:

> Thus painting will be nothing more than intersection of the visual pyramid [...] in a surface[...] (L. B. Alberti).

Very often, the texts contain multivoiced utterances, as in the following example. The realization of the instrument is explained referring to a mathematical model which is however conceived as a physical object.

I place a veil between the eye and the object, so that the visual pyramid can penetrate owing to the subtle veil (L. B. Alberti).

The orthogonal projection of the painter's eye on the plane of the 'veil' (glass). defines a point (the centric point, also called 'eye' in ancient treatises) that is going to acquire a special status in the code for perspective drawing (Rotman 1987).

However, Dűrer's glass leads to the genesis of at least two mathematical models, related to each other: the plane section of a pyramid (or cone) for a plane representation and the centric point, as representation of the meeting point (infinitely far) of a pencil of parallel lines, orthogonal to the picture plane.

3.3 From Dürer's glass to tertiary artifacts.

Modern projective geometry is rooted in the tradition of perspective practice. Actually, the definition of conics as projective invariants (a manifestation of which are circles) draws on the projection from one plane to another, allowing conics to be considered as anamorphoses of circles (Bartolini Bussi & Mariotti 1999; Bartolini Bussi et al. 1999b).

4. A CLASSROOM EXPERIMENT IN A 5TH GRADE CLASSROOM

4.1 The design

The new experiment has been designed for a 4th – 5th grade classroom. The original design (not given here owing to space constraints) is structured in three phases, related to primary, secondary and tertiary artifacts. The three phases are not rigidly separated; in fact, the presence of the teacher, as orchestrator of classroom activities, introduces a teleological element that makes each phase a base for the following one. Moreover, from the very beginning, the teacher is aware of the polysemy of each artifact: the intentional choice of the problems and management of the collective interaction aim at introducing and developing polysemy for all the pupils.

In particular, a hypothesis can be stated, concerning the two kinds of artifacts (primary and secondary), intentionally introduced into classroom activities.

1st Hypothesis (Polysemy): The intrinsic polysemy of the artifact supports the production of the polyphony of voices, in classroom activities (Polysemy hypothesis).

This hypothesis is expected to shape the design and analysis from two perspectives: the teacher's perspective (how the teacher introduces and supports polysemy) and the pupils' perspective (how the pupils internalise polysemy).

There is, however, another intention in the time devoted to the exploration and use of the concrete artifact. The experience of the concrete simple structure of a pair of objects (a point, materialised by the eye-hole, and a plane, materialised by the glass) should produce, when the artifact is no longer available, gestures like closing one eye, or tracing a plane in the air with a flat hand. The latter, which substitutes the presence of a concrete plane in a fixed position, has a generalising feature, as the gesture may allude to a plane in whatever position. The central function of body experience is witnessed also by the rich metaphors in the historical sources. We thus have the following hypothesis:

2nd Hypothesis (Embodiment): The concreteness of the artifact supports the production of gestures and metaphors, that are maintained also in the step of secondary artifacts and beyond.

The data collected in the classroom are expected not only to test the above hypotheses, but also to allow an exploratory analysis to describe the role of both polysemy and embodiment in the individual construction of mathematical meanings and ways of thinking.

4.2 Implementation of a classroom experiment

Table 1. Description of the first (A) and second (B) phases of the teaching experiment with the start of the third (C) phase

Session	*Tasks*	*Short description*
1 A	Discussion	Interpretation of a primary artifact: a Dűrer's glass that shows the skeleton of a cube and its perspective drawing (Figure 2).
2 A	Individual drawings	Free drawing of the primary artifact (Figure 3)
3 B	Discussion	Recall with the observation of individual drawing of the previous year. Interpretation of secondary artifacts (the voices given in § 3.2 above)
4 B	Individual drawing and Discussion	Real life drawing of a table with some objects. Construction of the table of invariants in the shift from reality to representation (see Bartolini Bussi, 1996).
5 B and C	Discussion	Focus on the transformation of rectangular shapes. Construction of the definitions of several types of quadrilaterals.
6 B	Small group work and spokesman's presentation	Design of a tool for effective real life drawing
7 B	Individual text	Written report on the previous task

The experiment was started at the end of the 4th grade (May 2002) and continued in the 5th grade (from October 2002). The teaching experiment was articulated into three phases (see the table 1).

Figure 2. Looking at Dürer glass *Figure 3. Drawing Dürer glass*

The third phase was planned following the design of a previous experiment (Bartolini Bussi 1996). The first two phases are centred on primary and secondary artifacts (Dűrer's glass and texts about its use). In the third phase, the main objective concerns pupils' appropriation of styles of mathematical reasoning: in particular, producing definitions and 'theorems' within a 'germ theory' of Desarguesian Geometry, based on transformation and invariants. A first approach to Desarguesian Geometry, emerged during a mathematical discussion (Session 5): it was concerned with the 'general' definition of quadrilaterals as anamorphoses of rectangles.

4. 3 Some data from the classroom

Several kinds of data have been collected:
- individual protocols (texts, drawings);
- audio-recordings (and sometimes video-recordings) of classroom activities;
- photos of the pupils at work;
- teacher's and observer's notes.

All the collective verbal interchanges have been transcribed and analysis is in progress. In the following, owing to space constraints, only some short excerpts of a discussion (session 3) will be discussed, to illustrate the *Polysemy* and the *Embodiment Hypotheses*.

The discussion is centred on the interpretation of a collection of short sentences, drawn from different manuals on painting (see § 3.2 above). The text has a central role and is used by the teacher as a tool of semiotic mediation. It is given as a stimulus, a sign to be interpreted. The task ('read and interpret') is introduced by the teacher and repeatedly recalled over time. The model of Dűrer's glass is no more available in the classroom. The sentence that introduces the first mathematical model is the following:

> Thus painting will be nothing more than intersection of the visual pyramid [...] in a surface [...] (L. B. Alberti).

Surprisingly, the word pyramid seems to be easy to interpret, with the help of the reference to Egyptian monuments. There is a shift towards the description of the shape of the pyramid by means of the number and shape of its faces. Then a very interesting exchange takes place.

> Alessandro B.: If the base is triangular it has 4 [faces], if the base is square it necessarily has 5. It depends on the base. The one we are talking about has either a square or a rectangular base, because we imagine a painting or a piece of glass and the point of the triangles reaches the eye.
> Federica: Yes, but Leon Battista Alberti's is not a real solid, it's an imaginary solid which takes shape while you're looking at it. We can't see it, we can see it only when we think of it, if we want to see it. For example we can see it now because we have just read it.
> Assia: Of course it's imaginary, otherwise it would harm you and then it wouldn't even allow you to see.
> Voci: Can you imagine a solid getting into your eye!
> *[Many gestures, funny ones as well! A moment of confusion and jokes about the visual pyramid with participation of the entire class].*

What Alessandro B. says gives an example of a multivoiced discourse, where the pivot element is the pyramid. Although the word 'pyramid' does not appear explicitly, there is cross reference to the school pyramid, to the vision pyramid and to its origin in Dűrer's glass. Other pupils take part in the discussion. Communication is possible because, although different interpretations rise from a particular sign, they find a shared context in the common experience of the use of Dűrer's glass. Moreover, meanings can evolve because each context contributes to enriching the interpretation given in the other.

The emergence of the joke confirms that the multivoiced discourse is mastered by many pupils. The joke witnesses the 'pleasure of absurdity' given by wavering between two different contexts and, at the same time, as Freud (1972) tells us, it may relieve the tension caused by 'critical reasoning' which, under this circumstance, reveals a conflictual interpretation of the term pyramid.

A new example is offered by the interpretation of an everyday word. Pupils guess that the word 'intersecazione' (the ancient spelling for 'intersezione,' i. e. intersection) is related to the word 'segare' (i. e., to saw) and interpret the excerpt as follows:

> Alessandro B.: [...] If you saw the visual pyramid you obtain the painting.
> Luca: How can you possibly saw the visual pyramid which is a solid that does not exist?
> Alessandro B.: Exactly how you imagine it. If you see it because you imagine it, you can saw it as well. You have to work with the mind.
> Elisabetta: It's like working with imagination: you have to imagine things and then they seem true.
> Marcello: It's not as it is with imagination, but with the mind because you have to saw well where you want, in order to draw what you want to do.

Federica: Yes, all right, but in any case you have to imagine it. I understood this, if you saw it near the object you obtain a large image, if you saw it near the eye you get a smaller image.
[*With gestures, many children cut, saw the visual image. They trace many imaginary planes which are parallel to an imaginary painting.*]
Alessandro B.: If you go down straight, because with our hands we form a kind of plane parallel to the one of the objects [*With his hands he traces two parallel planes in space*]. In this way you certainly obtain a figure which is exactly the same as the base of the pyramid, but smaller.
Luca: Franca [*the teacher*], try to draw it on the blackboard so that we can understand better.
Teacher: I'll try, but I can't guarantee that you will understand better. [*the teacher draws, see the Figure 4*]
Marco: Now we can understand better why it comes out with the same figure …
Alessandro B.: That's because you're sawing it in parallel. If you saw it obliquely, you obtain another figure, but I don't know what it is. Leon Battista Alberti tells us painting is only what there is if you saw the visual pyramid where you want. And then Federica is right when she says that if you saw the visual pyramid near the eye your drawing becomes very small.
Federica: Franca if you read what is written afterwards, written by Dürer, it teaches you what you have to do to saw the visual pyramid. When he says that you have to paint on glass, it's as if we went to the glass of the window and on a little piece of glass, sort of on a triangle, it's as if we were drawing what we see outside, right inside that little painting. [*She gets up, goes towards the glass, where she traces a rectangle with her finger and then pretends to draw inside what she sees*]. It's as if the glass had sawn/intersected the visual pyramid.

As in the previous excerpt, a dialogue takes place between the concrete referent and the ideal model. They support each other: Dürer's glass may be used to interpret the mathematical model and viceversa. A crucial role is played by certain words of the given texts, which, by themselves, may evoke different contexts; however, words are not the only signs involved in this complex semiotic activity.

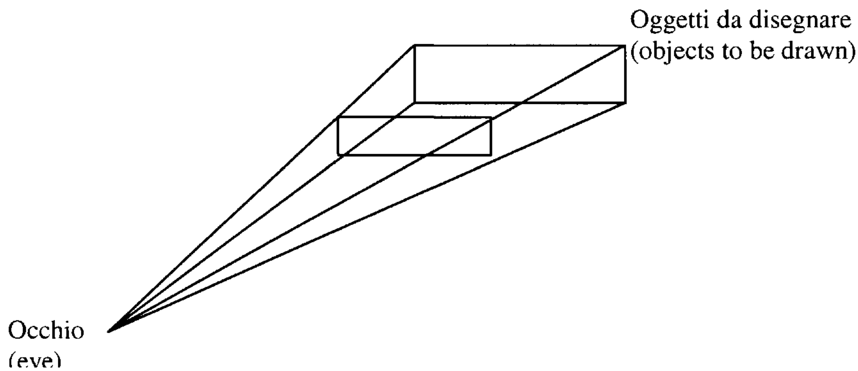


Figure 4. What the teacher draws on the blackboard

Firstly, there are gestures. Widely used by the pupils, gestures miming planes and lines in whatever position constitute a fundamental support to image a pyramid. Like the word 'pyramid,' gestures may represent both the concrete and the ideal pyramid, but better than the words, they may afford to represent general geometrical (spatial) properties.

Secondly, there are drawings. The teacher's drawing, requested by a pupil, seems to support pupils' mental imaging. Both hypotheses are confirmed in these short excerpts.

The sense of the text and the sense of the instrument are reconstructed with cross reference to each other. The discourse produced by the pupils is multivoiced, not only in the collective context (alternation of voices as alternation of turns) but also at the individual level (a pupil may control both voices by him/herself). Hence the polysemy of the artifacts is not only introduced by the teacher but also internalised by many pupils.

The Embodiment Hypothesis is also confirmed: we observe a rich production of gestures; pupils' gestures support their mental imagery attempt to adapt the pyramid model to the painting situation to be explained. In particular, it seems that, after the intervention of Elisabetta, many pupils look for and find a confirmation of their interpretation in miming the cutting of the pyramid.

The space constraints of this paper do not permit to go any further in our analysis. However, we think that the previous examples give sufficient evidence of how the articulation of voices is related to the polysemy of the artifact and how gestures and metaphors may provide a powerful base to interlace different meanings.

5. CONCLUDING REMARKS AND OPEN PROBLEMS

As the previous example shows, the theoretical framework presented in this paper, which originated the two hypotheses, seems to provide a powerful tool both to design and interpret educational intervention. Deeply rooted in the Vygotskian theory, this frame is an attempt to explain the complex functioning of artifacts of different types in the construction of mathematical meanings. The analysis of the interrelation among different artifacts, as modelled by Wartofsky's classification, is the main point, characterizing the evolution of this theoretical framework, in respect to those presented in previous works.

The polysemy of a primary artifact, Dűrer's glass, and the following emergence of a mathematical theory were reconstructed through the historical analysis, based on a number of texts, describing how and why to use this tool in the drawing. The teacher was in charge of introducing polysemy in the mathematics classroom. The new teaching experiment aimed at refining the previous hypotheses (Bartolini Bussi 1996) concerning the key role of texts (secondary artefacts) as instruments of semiotic mediation, and focused on the importance of a direct link with the experience on a primary artefact. A text may play a crucial role, and we have shown how the teacher may use it: intepretation of a text shows its power in triggering a semiotic game within which meanings may be evoked and evolve. However, the emergence of mathematical modelling seems to be strictly related to the fact that it can be grounded in the concrete experience with a primary artefact, experience which provides all the pupils with a source of meaning. In particular, direct manipulation of the artifact, constitues the basic element for interpreting a text, but also provides the concrete reference within which one interprets the mathematical model evoked in the text.

The findings of the first part of the teaching experiment confirm our hypotheses:

- the function of polysemy of an artifact, as means to fuel multivoiced discourse, which may be directed according to the intention of the teacher and the possibility that polysemy is internalised by most pupils;
- the function of body experience, that is of concrete manipulation and its mimic, both in the social interaction, when the class look for a shared meaning, and in the individual action, when the tension of interpretation asks for an immediate reference.

These findings are consistent with those of other experiments, which all show the potentialities of signs (secondary artifacts) with respect to both the primary and the tertiary artefacts (gears, Bartolini Bussi et al. 1999 a; compass, Bartolini Bussi 2002; drawing instruments and pantographs, Bartolini Bussi 1998, 2001; Bartolini Bussi & Pergola 1996; abacus, Bartolini Bussi, 2003; big size models of conic sections, Bartolini Bussi, to appear). In the quoted experiments the primary artifacts are concrete, physical instruments, taken from the history of mathematics and technology, and represented by historical sources (secondary artifacts).

These findings are also consistent with those coming from other experiments carried out within a Vygotskian perspective and centred on the semiotic mediation processes related to the use of microworld (Mariotti 2001, 2002, Mariotti & Cerulli 2001, Laborde & Mariotti 2002). where specific elements of the microworld (dragging facility, commands available, macro ...) may be used as instruments of semiotic mediation. In the case of a microworld, primary and secondary artifacts are intimately linked so that sometimes it is difficult to separate them, to the extent that the polysemy of signs is even more evident. On the contrary, the specificity of 'manipulating virtual objects' surely demands a reformulation of the Embodiment Hypothesis. However, we think that similarities and differences between concrete instruments, coming from historical tradition, and new technologies, generate a number of crucial questions and provide a stimulating open field of investigation. In this new field we intend to design one of our future research projects.

ACKNOWLEDGEMENT

Research study funded by MIUR: Project 'Problems about the teaching and learning of mathematics: meanings, models, theories' (n. 2003011072). We wish to thank Marcello Pergola, who has built models for the classroom; Simona Vangelisti who has been carefully observing the experiment for months; all the colleagues of the research groups who have carried out the experiments and taken part in the protocol analysis.

Dipartimento di Matematica, Università di Modena e Reggio Emilia (Italia)
Dipartimento di Matematica, Università di Pisa (Italia)
Scuola Elementare Palestrina, Modena (Italia)

REFERENCES

Bachtin, M. (1979). *Estetica e romanzo*. Torino: Einaudi.

Bachtin, M. (1988). *L'autore e l'eroe. Teoria letteraria e Scienze umane*. Torino: Einaudi.

Bartolini Bussi, M. G. (1996). Mathematical Discussion and Perspective Drawing in Primary School. *Educational Studies in Mathematics* 31.1 – 2, 11-41.

Bartolini Bussi, M. G. (1998). Drawing Instruments: Theories and Practices from History to Didactics. *Documenta Mathematica* – Extra Volume ICM 3, 735-746.

Bartolini Bussi, M. G. (2001). The Geometry of Drawing Instruments: Arguments for a didactical Use of Real and Virtual Copies. *Cubo Matematica Educacional* 3.2, 27-54

Bartolini Bussi, M. G. (2002). The Theoretical Dimension of Mathematics: a Challenge for Didacticians, *Proc. 2000 (24th) Annual Meeting of the Canadian Mathematics Education Study Group* (Montreal): 21-31.

Bartolini Bussi, M. G., & Boni, M. (2003). Instruments for semiotic mediation in primary school classrooms. *For the Learning of Mathematics* 23.2, 12-19.

Bartolini Bussi, M. G., (to appear). The Meaning of Conics: historical and didactical dimension. In C. Hoyles, J. Kilpatrick, & O. Skovsmose (Eds.), *Meaning in Mathematics Education*. Kluwer Academic Publishers.

Bartolini Bussi, M. G., & Mariotti, M. A. (1999). Instruments for Perspective Drawing: Historic, Epistemological and Didactic Issues. In G. Goldschmidt, W. Porter, & M. Ozkar (Eds.), *Proc. Of the 4th International Design Thinking Research Symposium on Design Representation* III, Massachusetts Institute of Technology & Technion – Israel Institute of Technology, 175-185.

Bartolini Bussi, M. G., & Pergola, M. (1996). History in the Mathematics Classroom: Linkages and Kinematic Geometry. In H. N. Jahnke, N. Knoche, & M. Otte. (Eds.), Geschichte der Mathematik in der Lehre. Goettingen: Vandenhoeck & Ruprecht, 39-67.

Bartolini Bussi, M. G., Boni, M., Ferri, F., & Garuti, R. (1999a). Early Approach To Theoretical Thinking: Gears in Primary School. *Educational Studies in Mathematics* 39.1 – 3, 67-87.

Bartolini Bussi, M. G., Nasi, D., Martinez, A., Pergola, M., Zanoli, C., Turrini, M. et al. (1999b). *Laboratorio di Matematica : Theatrum Machinarum*. 1° CD rom del Museo, Modena: Museo Universitario di Storia Naturale e della Strumentazione Scientifica <http://www.museo.unimo.it/theatrum>.

Boero, P., Dapueto, C. Ferrari, P., Ferrero, E., Garuti, R., Lemut, E., Parenti, L., Scali, E. (1995). Aspects of the Mathematics-Culture Relationship in Mathematics Teaching-Learning in Compulsory School. In *Proc. 19th PME (Recife)* 1, 151-166

Cerulli M., Mariotti, M.A. (2001). L'algebrista: a microworld for symbolic manipulation. In H. Chick, K. Stacey, J.Vincent, & J. Vincent (Eds.), *The Future of the Teaching and Learning of Algebra, Proceedings of the 12th ICMI Study Conference*. The University of Melbourne, Australia.

Engestroem, Y. (1987). *Learning by Expanding: an Activity Theoretical Approach to Developmental Research*. Helsinki: Orienta – Konsultit Oy.

Engestroem, Y. (1990). When is a tool? Multiple meanings of artefacts in human activity. In *Learning, Working and Imagining: Twelve Studies in Activity Theory*. Helsinki: Orienta-Konsultit Oy, 171-195.

Field, J. V. (1997). *The Invention of Infinity: Mathematics and Art in the Renaissance*. Oxford University Press.

Freud, S. (1972). Il motto di spirito e la sua relazione con l'inconscio. In *Opere* 5.Torino: Boringhieri.

Laborde, C., & Mariotti, M. A. (2002). Grounding the notion of function and graph in DGS. *Actes de CabriWorld 2001*. Montreal.

Lakoff, G., & Nunez, R. (2000). *Where Mathematics Comes From. How the embodied mind brings mathematics into being*. New York: Basic Books.

Mariotti, M. A., Bartolini Bussi, M. G., Boero, P., Ferri, F., & Garuti, R. (1997). Approaching Geometry Theorems in Contexts: From History and Epistemology to Cognition. In *Proc. 21st PME Int. Conf.* 1, Lahti, Finland, 180-195.

Mariotti, M.A. (2001). Introduction to proof: the mediation of a dynamic software environment (Special issue). *Educational Studies in Mathematics* 44,1 & 2, 25-53.

Mariotti, M. A. (2002). The Influence of Technological Advances on Students' Mathematics Learning. In L. D. English et al. (Eds). *Handbook of International Research in Mathematics Education*. Lawrence Erlbaum Associates Publishers, 695-724.

Otte, M. (1997). Mathematics, Semiotics and the Growth of Social Knowledge. *For the Learning of Mathematics* 17.1, 47-54.

Rotman, B. (1993). *Signifying Nothing: The Semiotics of Zero*. Stanford University Press.

Vygotskij, L. S. (1974). *Storia dello sviluppo delle funzioni psichiche superiori e altri scritti. Firenze:* Giunti.

Vygotskij, L. S. (1987). *Il processo cognitivo*. Torino: Boringhieri.

Vygotskij, L. S. (1992). *Pensiero e Linguaggio*. Bari: Laterza Editore.

Wartofsky, M. (1979). Perception, Representation, and the Forms of Action: Towards an Historical Epistemology. In *Models. Representation and the Scientific Understanding*. D. Reidel Publishing Company: 188-209.

HEINZ STEINBRING

DO MATHEMATICAL SYMBOLS SERVE TO DESCRIBE OR CONSTRUCT "REALITY"?

Epistemological Problems in Teaching Mathematics in the Field of Elementary Algebra

Abstract. By means of an epistemological analysis of a teaching episode from a mathematics classroom the paper tries to exemplify and to concretize some fundamental ideas developed in a theoretical perspective by Michael Otte towards the basic role of visualizations and metaphors for mathematics teaching. The metaphor "The equation is a balance" is taken as a paradigmatic case. "*The equation is a balance* is a sentence not to be taken verbally, and the seemingly abstract (the algebraic equation) cannot be limited to the seemingly concrete and empirical (the balance) in a process of reduction and of visualization, but on the contrary, the balance represents the highly general meaning of the interaction or the reciprocity or it stands for the dynamic and compensation. ... The real balance has meaning for the equation. The algebraic concept of equation could not have been constituted without the experience of the balance" (Otte 1984).

Key words: epistemology, epistemological triangle, mathematical knowledge, sign, sign language, symbol.

> Visualizations in mathematics are no pictures, no illustrations nor visual exemplifications. They act as metaphors. A metaphor must be spontaneously acceptable and intuitively evident; ...the probably oldest and most widely spread metaphor in mathematics [is]: "The equation is a balance" (Otte 1984).

1. INTRODUCTION: ELEMENTARY ALGEBRA AS A MATHEMATICAL SIGN LANGUAGE

In general mathematics teaching, elementary algebra is of special interest. Following the first years of school, in which arithmetic mainly stands in the centre of mathematics instruction, a new difficulty, yes, often a break in the development of the students' mathematical thinking takes place with the content of algebra. The most striking features of algebra are the new signs, the letters, the variables, the operation signs, the chains of sign-like combinations. Among other things, algebra represents something like a mathematical language, a sign language.

> Mathematics is often regarded as a difficult and mysterious science because of the numerous signs it uses. Obviously, there is nothing more inconceivable than a sign lan-

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 91 – 104.

> guage we do not understand. Likewise, a sign language, which we understand only partly and whose use we are not familiar with, is hard to follow. (Whitehead 1948, 35)

This aspect of algebra as a mathematical sign language will be at the centre of this analysis. Which are the particularities of this algebraic language? How does one "speak" in this language? Which meaning do the "words" and the "phrases" in this language have? How are sense and meaning communicated in this language? Which special social communication and mathematical culture is constituted with and by means of this language?

In a first understanding, the algebraic sign language mainly has an economic function; the algebraic signs serve to abbreviate, simplify, clarify, and so forth complex circumstances. With the signs, other objects are named directly; with the algebraic operations; the concrete relations and the concrete treatment of objects are represented directly in a symbolic way. Thus, the algebraic sign language is understood as a "bijective" translation of objects out of reality into mathematics.

Before analyzing the role of the *algebraic sign language* in mathematics instruction (or in *mathematical technical language*), it is necessary to obtain clarity on certain tacit, implicit or also open, explicit presumptions on the role of language; the perspective taken on the relation between language and reality should be explained for the analyses of mathematical communication.

Mathematical signs do not represent empirical things, but embody relations. Raymond Duval formulates this fact as the "paradoxical nature of mathematical knowledge:"

> there is an important gap between mathematical knowledge and knowledge in other sciences such as astronomy, physics, biology, or botany. We do not have any perceptive or instrumental access to mathematical objects, even the most elementary, ... We cannot see them, study them through a microscope or take a picture of them. The only way of gaining access to them is using signs, words or symbols, expressions or drawings. But, at the same time, mathematical objects must not be confused with the used semiotic representations. This conflicting requirement makes the specific core of mathematical knowledge. (Duval 2000, 61)

Mathematical knowledge must be represented by signs or symbols within a semiotic system that is of fundamental importance for mathematical activity. This is where the described paradox develops: In order to work on a not directly accessible mathematical concept and to understand it, a suitable symbolic representation system is required. However, in order not to confuse this sign system with the mathematical concept, and to operate meaningfully within this system, knowledge of the particular mathematical concept is necessary (cf. Duval 1993, 37f; Steinbring 1997; 1998a).

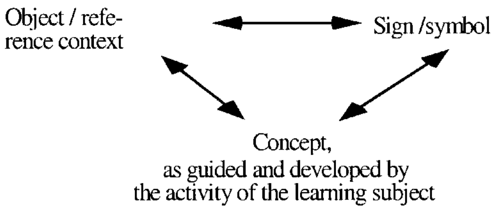


Figure 1

The epistemological triangle represents a theoretical instrument for tackling this problem that one requires signs and symbols for mathematical knowledge, but that these signs and symbols themselves are not the knowledge.

Mathematical knowledge cannot be reduced to signs and symbols. The connection between the signs to code the knowledge and the reference contexts to establish the meaning of this knowledge can be represented in the *epistemological triangle* (cf. Steinbring 1989; 1991a; 1998b). The relations between the corner points of this triangle are not defined explicitly; they form a balanced system, that reciprocally supports itself. In the ongoing development of the knowledge, the interpretations of the sign systems and the chosen according reference contexts will be modified or generalized by the epistemological subject or the learner.

Mathematic-didactic problems in particular, as they become visible in everyday instruction practice under an epistemological perspective, have been an essential reason to develop the epistemological triangle.

> The practice of mathematics, especially in school, is ... usually seduced to an identification of sign and signified by the automatization, the algorithmic, as the formula expresses it as a calculating procedure, or, if one makes the threefold distinction of concept, sign and object, which would actually be necessary, it is seduced to an identification of sign and object, neglecting an independent conceptual. (Otte 1984, 19)

Similar triangular schemes for the analysis of the semiotic problem, how relations between symbols and referents are realized have been developed in philosophy of mathematics, in linguistics and in the philosophy of language (e. g., Frege 1969; Ogden & Richards 1923). The perspective of this triangular diagram leads to the question whether the algebraic sign (the verbal instrument) is essentially a fixed name for a certain thing, for a specific object, or what the sign or symbol could mean otherwise.

In didactics of mathematics and also in fields of philosophy of mathematics, a very popular conception is that mathematical signs are merely names for things, even though mathematical signs in particular are very exact and unequivocal, precise terms for certain objects – in contrast to everyday language.

> In mathematical technical language, as in any other scientific technical language, one tries to avoid the ambiguities existing in everyday language and this is done by means of assigning each used sign one and only one well-determined meaning in the frame of a theoretical connection. Reversibly, the same technical sign is supposed always to be assigned to each used concept or meaning content. (Maier & Bauer 1978, 142-143)

This conception of the role of language and thus also of the role of mathematical language – is fundamentally criticized from a philosophical perspective. The illusion that the words and sentence constructions in the language correspond with things in reality in an unequivocal way is questioned. Lakoff and Johnson criticized this attitude as the myth of objectivism.

> According to the myth of objectivism, the world is made up of objects; they have well defined properties, independent of any being who experiences them, and there are fixed relations holding among them at any given point in time. These aspects of the myth of objectivism give rise to a building-block theory of meaning. If the world is made up of

> well-defined objects, we can give them names in a language. If the objects have well-defined inherent properties, we can have a language with one-place predicates corresponding to each of those properties. And if the objects stand in fixed relations to one another (at least at any given instant), we can have a language with many-place predicates corresponding to each relation.
> Assuming that the world is this way and that we have such a language, we can, using the syntax of this language, construct sentences that correspond directly to any situation in the world. The meaning of the whole sentence will be its truth conditions, that is, the conditions under which the sentence can be fitted to some situation. The meaning of the whole sentence will depend entirely on the meanings of its parts and how they fit together. The meanings of its parts will specify what names can pick out what objects and what predicates can pick out what properties and relations. (Lakoff & Johnson 1984, 202)

An important aspect that reveals the independence of language – and also of mathematical language is metaphors, metonymies, or also language games. An excellent illustration of the implications of language as an independent, developing means of formation of reality and not merely a kind of duplication of reality is sign language for the deaf and dumb (see Sacks 1990; see, also, Pinker 1998). With gesture and sign languages one has often proceeded from the assumption that gesture signs equivalent to the spoken "signs," the words of those who can speak, ought to be developed for the deaf-mute. Hence, in the background, we once again find a kind of "objectivism," that the signs of sign language only serve for the translation into another language, or to code things of reality. A true sign language, however, is not simply a sign-like translation.

> True sign languages, however, are complete in themselves: Their syntax, grammar and semantic do not require a supplement, but they differ in their nature from the ones of all other articulated or written phonetic languages. Thus it is not possible to translate a spoken language word by word, sentence by sentence into sign language – their structures are fundamentally different. (Sacks 1990, 53)

And with sign language – just like spoken language – one can behave actively, construct one's own world, enter the symbolic world, differentiate and create the world in a new way by means of the metaphorical use of language signs. "Language creates experience in a new way. ... By means of language ... one can introduce the child into the purely symbolic sphere of past and future, of distant areas, ideal relations, hypothetical events, of utopian literature, beings, imaginary entities – from the werewolf to π-mesones ..." (Church 1971, 96, as cited in Sacks 1990, 68).

The deaf-mute are also able to "play with pictures, with hypotheses, with possibilities or to enter the empire of imagination or of metaphors" (Sacks 1990, 65) with their sign language. This, however, makes it necessary not to understand sign language as "pantomime or a gesture-code or maybe as a kind of broken English in hand signs." (Sacks 1990, 108) In contrast, one should not conceive gestures as "pictures ..., but complex abstract symbols with a complex inner structure." (Sacks 1990, 108) Sign language differs in aspects from spoken language; but it is important to realize that it is a "complete," not a deficit language; indeed, in its difference to spoken language, it also has its own features and advantages.

Certainly, algebraic sign language is neither a sign language nor a common spoken language. But it is also a "living" language, not merely a sign code, or an unequivocal picture of real or mathematical objects. Which particular characteristics does the autonomous language of elementary algebra possess? With the help of algebraic sign language, is it also imaginable "to play with pictures, with hypotheses, with possibilities or to enter the empire of imagination or of metaphors" (Sacks 1990, 65)? What does the metaphorical character of algebraic sign language consist in? How can it be brought forth and maintained?

2. THE INTERACTIVE DEVELOPMENT OF MATHEMATICAL MEANING: ANALYSIS OF A CLASSROOM EPISODE IN ALGEBRA TEACHING

So far, two contrary positions to the role of language have been emphasized:

- Language serves for the *description of reality*; it pictures aspects of reality; it gives – as exact and differentiated as possible – reproductions of reality.
- Language is a *means of constructing reality*; language is an autonomous, self-referential system and instrument for creating reality by means of identifying and emphasizing things in reality; with abstractions and metaphors, structures and relations in reality are created.

In a similar way, this difference – these two contrary points of view – can also be worked out for algebraic sign language:

- Algebraic sign language serves for *describing reality*; it pictures aspects out of technical contexts, out of other mathematical structures or reference domains (out of geometry, arithmetic, diagrams, algebraic structures).
- Algebraic sign language is a means of *constructing reality*; it is an autonomous, self-referential system and instrument with its own rules for changing and creating algebraic language elements; the algebraic equations do not just describe a findable reality; the algebraic expressions and equations formulate construction conditions for the creation of (ideal) objects within the reality to be created.

For algebra – and already for elementary algebra as well – these features of autonomy and the construction of its own reality are justified in the following: Algebraic connections or equations, such as the Pythagorean theorem $a^2 + b^2 = c^2$ or the law of force in physics: $F = m \cdot a$ (force equals mass multiplied by acceleration; see, for a discussion of theoretical terms Jahnke 1978) are no mere descriptions of findable reality in the sense that all aspects of reality are described down to the last detail with them; they are postulates or axioms to reality, conditions to which (ideal) objects, references, and so forth are produced.

The following examines an instruction episode on elementary algebra from this conceptual perspective on the role of algebraic sign language as a means of constructing reality.

2.1 *Analysis of the Teaching Episode: "System of two linear equations and balance"*

During the episode considered here, the students are asked to interpret operations on two linear equations with the help of a balance. The balance is a familiar metaphor for an algebraic equation as well as for the admissible operations and rearrangements with the equation (MacGregor 1998). The main statements by both students and their teacher will be summarized in the following.

The episode can be structured into the following phases and subphases:

Phases	*Contribution*	*Theme*
1	*1 – 7*	*Posing a problem: Two balances with weights and objects*
2	*8 – 16*	*Equations: An algebraic description of the balance situation*
3	*16 – 52*	*Developing a "practical" solution of the balance problem*
3.1	*16 – 23*	*Conditions for the practical solution*
3.2	*24 – 33*	*Proposals for solution: Estimations and comparisons*
3.3	*34 – 52*	*The "practica" solution*

During the first phase (1-7) of this episode, the teacher shows the following transparency to the students:

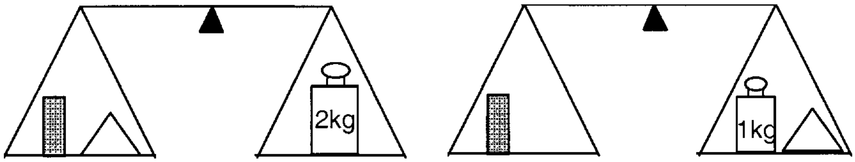


Figure 2

First, the teacher explains the situation:

4 T.: ... Two balances are shown ... so ... Yes, there are weights on it. Here we have such a 2 kilo piece made of iron, and there a one kilo piece; and now there are two objects, yes, a red one, which could be a box, or something, and a kind of cone, here, and there too.

The box is colored red, and the cone is colored blue.

The teacher makes some limiting conditions for possible solutions of the problem:

6 T.: ... you are only allowed to use these two weights. You don't get any other, additional, that is not, again 500 g and 100 or whatever other exists. ...

During the second phase (8-16), one student quickly furnishes a mathematical description of both balances with the help of algebraic signs and equations: "This cone there we call x and the red y, then we have x plus y equals two kilos. ... And the other down there is then, one kilo plus x equals y." (11, 13). He has correctly described the two equations:

$$x + y \quad = 2 \text{ kg}$$

$$1 \text{ kg} + x = y$$

But the teacher wants a "practical" solution. In the course of the third phase (16-52), such a practical solution is exemplarily discussed and elaborated. In subphase 3.1

(16-23), the teacher specifies some conditions for the acceptability of a practical solution.

16 T.: ... I first of all would solve it with you in the practical way,
... I could say, one could take off something here and put it into some other place, that's allowed.
...You have to move some objects at least, put them somewhere.

In their following statements, the students seem to have in mind some kind of weighing procedure, as it is generally known and also applied: All objects to be weighed have to be put on one side of the balance, and then they are weighed out with suitable weights:

17 S.: ... all of them onto the other side, or so ...
18 S.: ... all on one side ... take half a kilo ...
19 T.: No, well, sawing through is not possible, no, the weights ...

But the teacher refuses this proposal; it is not allowed to saw through the weights. In other words, one does not have enough many weights for an exact measuring procedure, only the 1 kg and 2 kg weight.

Thereupon one student objects that indeed one object could weigh 0.5 kg and the other object could weigh 1.5 kg; accordingly the condition is specified.

20 S.: ... but if this, ehm, the blue one would be now 0.5 and the red one 1.5, then it would work.
21 T.: This would be ..., that must not be right, it could be that, for my sake, that this then is, eh, 1.3 and the other 0.7, that would also be OK ... this could also fit, yes. But this we cannot say.
22 S.: ... but it doesn't work!
23 T.: Well, this one, ... practically, how could one now, ... if one had to work with it now, could one find out something now by rearranging? ... Stephanie?

The result of this discussion seems to be that there are not enough concrete different weights for a concrete procedure of measurement; but the weight of an object could differe from 1 kg or 2 kg. Is a "pure practical" solution at all possible on this basis?

In subphase 3.2 the students propose an approximation towards the solution, that tries to follow the teacher's conditions as strictly as possible: There is one 1 kg weight, one 2 kg weight, a box, and a cone. One student says:

24 S.: When one takes off the blue, ehm, there above, and the red then goes up, then one knows, that it is lighter than two kilos.

He means that the following will happen to the first balance when one takes off the blue cone: The left scale will go up, the right scale will go down; therefore the red box weight less than 2 kg.

The teacher confirms the correctness of his argument, but he wants a more exact determination of the value of measurement. Is this possible on the basis of the given conditions?

During the third subphase 3.3 (34-52) the "practical" solution as expected by the teacher is elaborated. A student proposes:

34 S.: Now, the red one here is as heavy as the blue one and one kilo.

With this remark he points to the lower balance: "red" = "blue" + 1 kg

The student continues:

34 S.: ... Then one can take off the red one and therefore add another blue one and a one kilo piece.

Now he is referring to the upper balance, in which one can substitute the red box through the blue cone plus the 1 kg weight.

This argumentation is resumed again after the teacher's requirement.

38 S.: Well, the red one there above ...
40 S.: ... yes, ehm, this is as heavy as a blue part and one kilo, ...
42 S.: Yes, and that, ehm, there above, take the red off.
44 S.: ... and the others, ehm, the blue parts and the one kilo piece therefore ...
45 T.: ... substitute them.

Now it becomes obvious that other conditions of the concrete situation are no longer fulfilled. At least, in the beginning, the teacher has not indicated that there are several similar exemplars of red and blue objects at one's disposal. The situation given only allowed two objects and two weights, at one time, and the other balance situation could not be presented at the same time but only consecutively.

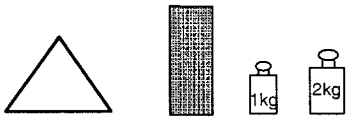


Figure 3

Because the weights did not exist in all possible combinations, as the teacher required in his assumptions before, – a condition that is normally fulfilled in every practical weighing procedure – one has to conclude that the concrete objects – the red box and the blue cone – do exist in several exemplars. (Are there perhaps several exemplars of the two weights existing at the same time? And in a way, later, one is allowed to saw through the numbers – representing the weights.)

The student introduces the variability of the objects with the following remarks: " ... and therefore add *another blue one* ..." (34); and later: " ... the *blue parts* and the one kilo piece therefore ..." (44). The teacher legitimizes it by saying: "... you then would have here above *two blue* and a one kilo piece..." (47). This variability of the given objects to be weighed makes the concrete balance a "mathematical balance" – in contrast to the restricting conditions of the teacher that the concrete objects cannot be measured by matching weights. The student's description of the solution approach with the characteristics of the objects becomes a direct translation of the symbolic-algebraic representation as given at the beginning of the episode by another student:

$$\begin{aligned} x + y &= 2\text{ kg} \\ 1\text{ kg} + x &= y \\ \text{"red"} + \text{"blue"} &= 2\text{ kg} \\ 1\text{ kg} + \text{"blue"} &= \text{"red"} \end{aligned}$$

This makes the solution possible, not in a "practical" way, but as a kind of theoretical deduction:

49 T.: What could you then deduce from it, then, Ismail?
50 S.: That, ehm, the two, eh, blue parts are one kilo heavy. One then would weight 500 g...

and, in the end, the procedure used here in a hidden manner is identified as the substitution procedure.

2.3 Aspects of an Epistemological Analysis of the Teaching Episode

In the course of this episode, the variables x and y, the measuring numbers (1 kg, 2 kg) and the objects with their properties (red box and blue cone) are used to describe a measurement situation with two different states. The symbolic variables x and y and the "concrete" variables "red" and "blue" are, in a way, exchangeable. According to the concrete weighing procedure, one would expect these variables to serve as names or marks for objects again; this also seems to correspond to the teacher's intentions.

Table 1

	Practical Weighing	Algebraic Weighing
Conditions	• The objects to be weighed exist only in one exemplar • There exist a number of matching different weights	• the objects to be weighed (variables $x, y, z, \ldots$) exist as often as necessary • the weights/numbers can be divided, multiplied, etc. (with other numbers) as often as necessary
Action	Each single object has to be weighed out with matching weights as exactly as possible; in this way determining the solution	Drawing conclusions from one (or more) situations of equilibrium; operating with the given "relations;" substitute objects/variables according to "rule of equilibrium" and determine the solution in this way
Concept	Weighing out by a **direct** (or approximate) **comparison** between **object and weight**	Weighing out (solution) by an **indirect comparison**; given states of equilibrium/**conditions of equivalence** are compared; the solution is deduced by a logical transformation

The process of negotiation between teacher and students regarding the admissible conditions and procedures of the measurement situation for solving the problem reveals the difference between a concrete, practical and a theoretical, mathematical way of measurement (Table 1).

Bearing the structure of the epistemological triangle in mind (Figure 1), one can observe that the teacher intends to use the picture of the two balances as an explanatory reference context for the elementary algebraic formula, for instance, in the following way: The concrete situation of a balance with objects and with weights should here be, in a way, a clear, familiar reference context serving for a possible interpretation of the algebraic symbol system of "equation" and to provide it with meaning (Figure 4).

At the beginning of the episode, the aim of this reference context of the two balances was to provide the students with a direct and concrete interpretation of the two equations with two unknowns. In the course of the discussion between the teacher and some of the students about the conditions that are admitted and that are forbidden in the procedure of weighting, the "inversion" can be observed.

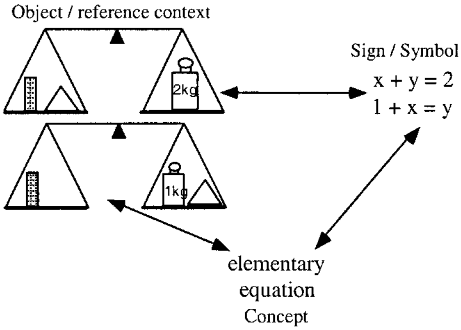


Figure 4

It becomes increasingly clear that perhaps in the beginning, the concrete balance could give some support for the understanding of an equation, but then in the course of interactively specifying the conditions of the process of elaborating and solving the problem, it turns out that the algebraic equation now describes and determines how the situation of the balance has to be understood and interpreted.

Here, we can observe a general issue: Elementary algebra – here the simple expression of "two equations with two unknowns" – is first understood as a name for a concrete object, that is, for two states of a balance with weights and objects. But during the course of interactive negotiation between teacher and students, the proper mathematical object is constructed, the mathematical balance. This mathematical object has a different status than the concrete balance: Whereas the concrete balance is a material object with concrete properties, the mathematical balance is not given directly; it is described by a number of defining conditions. The mathematical discourse in this episode has started with the teacher's intention to find a "practical" solution to the problem by using the balance in a concrete measurement process. But with the constraint that only two weights are allowed (1 kg and 2 kg), the further interaction increasingly specifies the conditions under which a mathematical balance can be defined. A mathematical object – such as the mathematical balance in our example – is described by defining conditions. A concrete object is given materially and is described by modes of use. The discourse in this episode has transformed the concrete properties of a real balance into the defining conditions of a mathematical balance.

The process of changing the function of a mathematical "sign/symbol" of being a mere name for a concrete object into a symbolic condition for constructing a mathematical object needs both sides: the balance and the algebraic sign system of an equation. The signs alone, the algebraically written equation, are not sufficient for defining the concept of an equation. In this way, the metaphor: "The equation is a balance." has turned into its inversion: "The balance is an equation."

The mathematical equation is created from being a name for the balance by adopting some structural similarities from concrete balances through a constructive act in which new relationships (conditions) changing the concrete balance into a mathematical balance are introduced. In this way, the algebraic equation constructs the mathematical balance: "The balance is an equation."

3. CLOSING REMARKS

The considerations on the role of language as a means of constructing reality and the analysis of the example episode from mathematics instruction in elementary algebra have revealed differentiated aspects to the meaning of elementary algebra as a *mathematical language*. One theoretical means of analyzing mathematical concepts is the epistemological triangle (Figure 1). In this triangle, aspects of "reality" are represented by the "object/reference domain" and aspects of (algebraic, mathematical) language are represented by the "sign/symbol." Concepts are constituted in a co-operation between language and reality.

The analysis has drawn attention to a possible development in the relation between "object/reference context" and "sign/symbol" that can be stated in elementary algebra, and it has also emphasized the way in which these reciprocal actions are created interactively within mathematics instruction.

In the following, I would like to characterize three levels of the relations between "object/reference context" and "sign/symbol:"

(1) Algebraic signs and symbols serve as *names* for objects within the reference context: description of a reality.
(2) Algebraic signs and symbols describe *relations* and *structures* within the reference context.
(3) There is a *reciprocal action* between algebraic signs and symbols and structures and relations: construction of a reality.

(1) The function of algebraic signs as names for things predominates in everyday mathematics instruction. Such an empirical interpretation of the "mathematical reality," for which the algebraic signs are to be introduced and defined simply as names

for given things, so that their use can be controlled, also seems to be very appropriate for the demands of teaching in a methodical way.

However, already for this interpretation of "signs as names for things," a purely empiristic use of mathematical signs must be relativized: An algebraic sign is not always simply a name for one single, fixed object, such as a single number. Algebraic signs can also refer to several numbers, even to "all numbers." Signs such as π or e describe one certain number in each case, letters such as a, b, c, ... can describe all numbers; x can potentially describe all numbers, but with x in a given problem context, one is usually looking for only one or a few possible numbers. We can see that generalizations are already contained potentially in algebraic numbers. They do not merely relate to one certain object, but to a whole class of objects (which have a certain feature).

(2) This potential general name function of the signs reveals a connection to the next level: The algebraic, symbolic terms and equations do not describe the constellation of (general) objects of a reality; equations formulate relations: The area formula for the trapezium is no mere description rule for the appropriate geometric elements of the trapezium and of the correct calculation:

$$F_T = h \cdot \frac{(g_1 + g_2)}{2}$$

With the formula, relations and structures between the elements are described systematically. The representation of a possible formula given here, expresses the relation between the height h and the midline $m = \frac{(g_1 + g_2)}{2}$. By the way, one should also pay attention to the way in which the midline in this algebraic formula is symbolically formulated as an "average value" of the two parallel sides of a trapezium.

On this level, the algebraic signs and symbols obtain a new interpretation: They do not remain mere names for objects; they describe structures and relations. Thus, a change from an empiristic sign use to a general characterization of relations takes place.

(3) On this third level, it now comes to an interpretation, in which the structures/relations on the side of the reference context and on the side of the sign/symbol start a reciprocal action. The roles of sign/symbol and of object/reference context become "exchangeable;" which side plays the role of sign/symbol and which one takes the role of "object/reference context." The (geometric, algebraic, arithmetical, ...) structures and relations can adopt the role of a reference context for signs/symbols as well as be the signs/symbols (an iconic language) for another reference context themselves. "...Mathematical signs play a *creative* rather than a merely descriptive function in mathematical practice. Those things that are 'described' – thoughts, signifieds, notions – and the means by which they are described – scribbles – are mutually constitutive: each causes the presence of the other." (Rotman 2000, 34-35)

This change becomes particularly visible in the example of the episode on the interpretation of the equation as a balance. In the interaction between teacher and students, the concrete balance was given a new interpretation as a "mathematical balance." The algebraic equation no longer remains a "name" for a given balance situation. In the two domains – the balance domain and the domain of the algebraic equation – the respective relations and structures come to the foreground and are related to each other. This reciprocal action makes it possible for the equation – the algebraic verbal expression – to construct the mathematical balance – and, in a certain way, the "mathematical balance" becomes a description for equations.

For the epistemological triangle (Figure 1), we thus obtain the following interpretation: The relation between the three corners is open in the sense that there is no fixed reference point (e. g., a given object), proceeding from which one could work out the other components of the triangle in a systematic and unequivocal way. In all cases, it is about reciprocal actions, which must be produced by the epistemological subject.

Ultimately, no fixed objects are given a priori in the epistemological triangle; in social and cognitive processes of mathematical concept creation, signs/symbols with objects/reference contexts are put into relations and in contrast to each other, interpreted with reference to each other, and developed. In the cognitive process, these relations and interpretations are produced and changed; they gain a new epistemological status. Ultimately, the mathematical objects are not merely found to already exist and be described in a logically clear way by the mathematical signs and symbols; the objects of mathematical reality are constructed with mathematical, symbolic language. Through this construction, mathematical concepts in processes of generalization develop; and, conceptual ideas and conditions influence and guide the interpretation of symbols in reference contexts and the conception of structures and relations within reference contexts.

The didactic-epistemological perspective on the role of mathematical language together with the qualitative analysis of an exemplary instruction episode has made it clear that in mathematical instruction interactions as well, attention must be paid to what Brian Rotman formulates for the practice of the researching mathematician as follows: "(W)hat present-day mathematicians think they are doing – using mathematical language as a transparent medium for describing a world of pre-semiotic reality – is semiotically alienated from what they are ... doing – namely, creating that reality through the very language which claims to 'describe' it" (Rotman 2000, 36-37).

Universität Dortmund

REFERENCES

Church, J. (1971). *Sprache und die Entdeckung der Wirklichkeit.* Frankfurt am Main: Suhrkamp.

Duval, R. (2000). Basic Issues for Research in Mathematics Education. In T. Nakahara, & M. Koyama (Eds.), *Proceedings of the 24th International Conference for the Psychology of Mathematics Education.* Hiroshima, Japan: Nishiki Print Co., Ltd. I., 55-69.

Duval, R. (1993). Registres de représentation sémiotique et fonctionnement cognitif de la pensée. In *Annales de Didactique et des Sciences cognitives* 5, 37-65.

Frege, G. (1969). *Funktion, Begriff, Bedeutung. Fünf logische Studien.* Göttingen, Vandenhoeck & Ruprecht.

Jahnke, H. N. (1978). *Zum Verhältnis von Entwicklung und Wissensbegründung in der Mathematik – Beweisen als didaktisches Problem.* Materialien und Studien des IDM, Band 10, Bielefeld: Universität Bielefeld.

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By.* Chicago: The University of Chicago Press.

MacGregor, M. (1998). How Students Interpret Equations: Intuition Versus Taught Procedures. In H. Steinbring, M. Bartolini-Bussi, & A. Sierpinska (Eds.), *Langugage and Communication in the Mathematics Classroom.* Reston, VA: National Council of Teachers of Mathematics, 262-270.

Maier, H., & Bauer, L. (1978). *Zum Problem der Fachsprache im Mathematikunterricht.* In Schriftenreihe des IDM Bielefeld 18. Bielefeld: Universität Bielefeld, 137-153.

Ogden, C. K., & F. A. Richards (1923). *The Meaning of Meaning.* London: Routledge and Kegan.

Otte, M. (1984). *Was ist Mathematik?* Occasional Paper 43 des IDM, Bielefeld: Universität Bielefeld.

Pinker, S. (1998). *Der Sprachinstinkt.* München: Knaur.

Sacks, O. (1990). *Stumme Stimmen.* Hamburg: Rowohlt.

Rotman, B. (2000). *Mathematics as Sign. Writing, Imagining, Counting.* Stanford, California: Stanford University Press.

Steinbring, H. (1989). Routine and Meaning in the Mathematics Classroom. *For the Learning of Mathematics* 9.1: 24-33.

Steinbring, H. (1991). The Theoretical Nature of Probability in the Classroom. R. Kapadia and M. Borovcnik (Eds.), *Chance Encounters: Probabilitiy in Education.* Dordrecht: Kluwer Academic Publishers, 135-167.

Steinbring, H. (1997). Epistemological investigation of classroom interaction in elementary mathematics teaching. *Educational Studies in Mathematics* 32, 49-92.

Steinbring, H. (1998a). Elements of Epistemological Knowledge for Mathematics Teachers. *Journal of Mathematics Teacher Education* 1.2, 157-189.

Steinbring, H. (1998b). Mathematical Understanding in Classroom Interaction: The Interrelation of Social and Epistemological Constraints. In F. Seeger, J. Voigt and U. Waschescio (Eds.), *The Culture of the Mathematics Classroom.* Cambridge, UK: Cambridge University Press, 344-372.

Whitehead, A. N. (1948). *Eine Einführung in die Mathematik.* Bern: Francke Verlag.

NORMA PRESMEG

METAPHOR AND METONYMY IN PROCESSES OF SEMIOSIS IN MATHEMATICS EDUCATION

Abstract. Building on Michael Otte's insights regarding the roles of icon, index, and symbol in mathematical signification, definitions of these categories of representation are explored in terms of metaphors and metonymies. A nested model of signs, based on Peirce's triadic formulation, is described, along with his trichotomic distinction among interpretants that are intentional, effectual, and communicational (leading to the *commens*). The theoretical argument and its utility is illustrated in terms of an episode of creating a proof in a college geometry class. The significance of the theoretical notions for creativity in mathematics is seen to reside in metaphorical and metonymical processes.

Key words: commens, icon, iIndex, metaphor, metonymy, representamen, symbol, universals

> The reasoning of mathematicians will be found to turn chiefly upon the use of likenesses, which are the very hinges of the gates of their science. The utility of likenesses to mathematicians consists in their suggesting, in a very precise way, new aspects of supposed states of things. (Peirce 1998, 6)

1. HOW AND WHY IS SEMIOTICS USEFUL IN MATHEMATICS EDUCATION?

After all, some of the originators of theories of semiotics were linguists. Ferdinand de Saussure's (1959) book, *Course in General Linguistics*, is a seminal work in this area. And Charles Sanders Peirce, himself fluent in Latin, Greek, and several other languages, makes it abundantly apparent in his writings (e. g., 1998, Vol. 2) that semiotics under girds and illuminates the study of languages and their structure. Why, then, is semiotics, defined as the study of semiosis (activity with signs), useful to mathematics educators? A hint of an answer to this question is given in the initial quotation from Peirce, and in this chapter I analyze semiotic aspects of metaphor and metonymy in particular, showing the relevance of "the use of likenesses" for deductive thinking and problem solving in the learning of mathematics. In a triadic model of nested signs based on the formulation of Peirce, the categorization of signs as iconic, indexical, or symbolic relates to the uses of metaphor and metonymy in semiosis. Informed by the insights of Michael Otte, I have found these constructs to be powerful lenses in my research, both on ways of connecting home activities of students with formal mathematical concepts in school and college (Presmeg 2002), and in understanding the ways that signs support learning of mathematics at all levels. After an initial description of this triadic nested model of signs and their uses,

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 105 – 115.

I analyze in more depth how metaphor and metonymy are implicated in the model, concluding with an analysis of an episode in a college geometry course.

2. A TRIADIC NESTED MODEL OF SEMIOSIS

Implicit in Peirce's triadic model of semiosis is a nesting effect. However, his writing is dense with ideas, some of which are explicated in careful detail, while others are barely sketched and further elaboration is left to the reader (e. g., his ten trichotomies, 1998, 481-491). Thus it is useful to consider in detail how this nesting of *signs* occurs. In this chapter I shall refer mainly to the trichotomy that by his own account he used most often, namely, a trichotomy designating three kinds of *signs*, which he called icons, indices, and symbols. But first it is necessary to say a few words about terminology. In the previous two sentences I have used the word *signs* in two different ways, and this is what Peirce does too, in different parts of his writings. To avoid confusion I shall use the word *sign* to refer to the totality of object, representamen, and interpretant (the former usage above – these signs are nested), and not to the representamen specifically (the latter usage). Thus when Peirce designates icons, indices, and symbols as three kinds of "signs", he is referring to three ways in which the representamen may stand in relation to its object; and the interpretant is then the result of reflection on this relationship – and thus is indirectly implicated. If we take an example suggested by Peirce (1998) and elaborated by Whitson (1994, 1997), the falling barometer (representamen) suggests that it will rain (object), but an act of interpretation is involved, and the observer may decide to take an umbrella (interpretant). That the interpretant is in terms of an action – the taking of an umbrella – reminds us that Peirce was one of the founders of pragmatism; but this aspect of his writings will not be pursued in this chapter. "To define a sign we therefore need an object as well as an interpreter" (Otte 2001, 5).

Michael Otte, in his writings on semiotics (e. g., 2001), usually invokes the representamen as "sign" following this usage in most of Peirce's work. As Otte (2001) remarked, Peirce (1998) defined a sign as anything that stands for something (called its object) in such a way as to generate another sign (its interpretant or meaning). This definition involves a double use of the term. The import of the definition seems to be that a sign is anything that *stands for* something else. In this case the interpretant (meaning) also stands for the relationship of the first two components. Although Otte's usage is consistent with this definition of Peirce, I shall not follow it here, because I am particularly interested in clarifying and describing the *totality* of object, representamen, and interpretant as a sign that becomes reified (Sfard 1992) as a new object in a nesting process that could continue indefinitely (as is also implied by Peirce's definition). A diagram casts light on the structure of the relationships.

Each of the rectangles in Figure 1 represents a sign consisting of the triad of object, representamen, and interpretant, corresponding roughly to signified, signifier, and a third interpreted component, respectively. This interpretant involves meaning making: it is the result of trying to make sense of the relationship of the other two components, the object and the representamen. It is important to note that the entire first sign with its three components constitutes the second object, and the entire second sign constitutes the third object, which thus includes both the first and the second signs. Each object may thus be thought of as the reification of the processes in the previous sign. Once this reification occurs, this new object may be represented and interpreted. Resonating with the cyclic nature of the processes involved, the construction of the representamen in the form of icon, index, or symbol, and its interpretation, also inform the creation of this new object. This formulation also illustrates Otte's (2001) statement that "the immediate object of a symbol is a sign itself" (15).

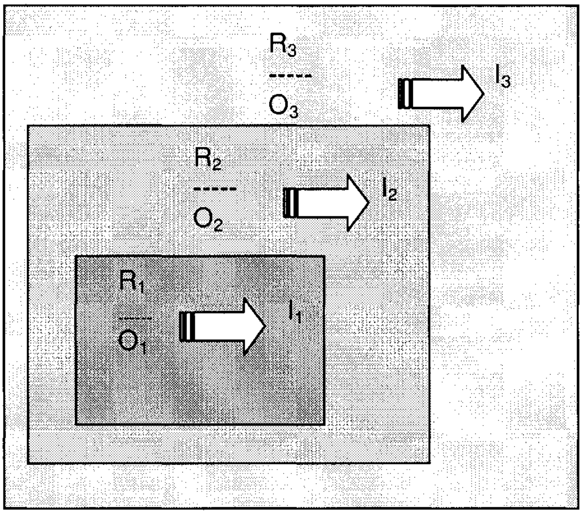


Figure 1: A representation of a nested chaining of three signs
O = Object (signified)
R = Representamen (signifier)
I = Interpretant
These three components together constitute the Sign, thus the three nested rectangles represent Signs 1, 2, and 3 respectively

This is a model of thinking and feeling, imagination and reason: semiotics eliminates Descartes' dualisms in this regard. The model has the potential to constitute a web of signs. One nested component of such a web may also be related metaphorically or metonymically with another such component. Icons, indices, and symbols and how they embrace metaphors and metonymies, are the subject of the next section.

3. METAPHOR AND METONYMY IN RELATION TO ICON, INDEX, AND SYMBOL

In Peirce's "trichotomy of signs," as he stated, there are three kinds of signs (referring to the representamen):

> Firstly, there are *likenesses*, or icons, which serve to convey ideas of the things they represent simply by imitating them. Secondly, there are *indications*, or indices, which show something about things, on account of their being physically connected with them. Such is a guidepost, which points down the road to be taken, or a relative pronoun, which is placed just after the name of the thing intended to be denoted, ... Thirdly, there are *symbols*, or general signs, which have become associated with their meanings by usage. Such are most words, and phrases, and speeches, and books, and libraries (Peirce 1998, 5; his emphasis).

In these definitions, Peirce is speaking broadly. With regard to icons, it may or may not be an actual physical likeness that connects the representamen with its object. It is the qualities of the object that are imitated, or its structure. As Otte (2001, 16) wrote, "The resemblance may be the extreme likeness of a photograph (CP2.281) or it may be subtler." But even with regard to photographs, Peirce (1998) qualifies this classification:

> Photographs, especially instantaneous photographs, are very instructive, because we know that they are in certain respects exactly like the objects they represent. But this resemblance is due to the photographs having been produced under such circumstances that they were physically forced to correspond point by point to nature. In that respect, then, they belong to the second class of signs [indices], those by physical connection (5-6).

Peirce here gives a hint of the interconnectedness of icons, indices, and symbols. An algebraic formula is also an icon (Otte 2001; Peirce 1998). But in what sense can a formula be considered a "likeness" of a mathematical idea? Surely there are elements of habit, of being "associated with [its] meanings by usage" (Peirce, quoted earlier) in the employment of an algebraic formula? Would not such a formula then partake more of the nature of a *symbol*, a "general sign", than it would of an *icon*? It could be argued that both of these categorizations are accurate. It is in the depiction of the structure of the relationships involved that a formula is also, indeed, iconic. And while it is not possible for new knowledge to evolve from a symbolic representamen because it is established by habit (Otte 2001; Peirce 1998), it is this property of depicting relationships that gives the icon its power in creative mathematical thinking. And this is where the role of metaphor comes in, because the use of an icon involves comparing two domains and noting their structural similarities. Because an icon strips some of the superfluous baggage from a mathematical relationship, it has the power to highlight the essential elements of the structure. In the language of metaphor, the source (object) and the target (representamen) domains are united in the icon. These two domains are also called the "vehicle" and the "tenor" respectively, in English language literature (Presmeg 1998), but the terms "source" and "target" better portray that it is the mathematical object that is the source of the structural relationships, and the icon is the representamen that captures them. Otte (2001, 16) pointed out that "icons are of key importance in mathematics" by virtue of their highlighting the analogy and structural similarity that play a fundamental role.

Having claimed that symbols cannot be the fountain of new knowledge, I do, however, want to qualify this statement. The symbols chosen to represent mathematical knowledge, even if they are conventional and become established by habit, can have a profound effect on the direction that new knowledge takes.

Moreover, the symbols chosen may further the construction o f new knowledge, or they may impede it, because the nature of a symbol is that it is a rule that will determine its interpretant (Otte 2001, Peirce 1998). One has only to consider the history of mathematics through the ages to be convinced of this claim. Time and again, as in the example of Leibniz's notation for calculus chosen over Newton's infinitesimals, or the problems occasioned by the ambiguous use of symbols for negative numbers, the symbols were a crucial factor in the furtherance – or hindrance – of mathematical thinking. In section 5, I shall give a more mundane example from a college geometry course, which nevertheless exhibits that it is important for mathematics educators to pay careful attention to all three kinds of representamen, the iconic, the indexical, *and* the symbolic, in teaching mathematical concepts.

Having established that icons depict structural relationships of mathematical objects by means of metaphor, it is the taking into account of *context* that suggests the importance of indices as metonymies. *Metonymy* is defined in Webster's dictionary as "a figure by which one word is put for another on account of some actual relation between the things signified" (Presmeg 1998, 29). This *actual relation* resonates with Peirce's aspect of *physical connection* in his definition of "indications, or indices". There is a sense in which icons, indices, and symbols all partake of metonymy, because in each of these cases, the representamen stands for the object by means of some actual relation. This relation is implied in the way the representamen is defined, as "a First which stand in such a genuine triadic relation to a Second, called its *Object*, as to be capable of determining a Third, called its *Interpretant*, to assume the same triadic relation to its Object in which it stands itself to the same Object" (Peirce 1998, 272-273; his italics). But the way in which the representamen stands in relation to its object in the case of an index is the closest to the literary use of the term metonymy, in which the context is needed for interpretation (as in the example, "Washington is talking with Moscow" – meaning that the governments centered in those cities are communicating). There is another forms of metonymy, called synechdoche, in which the part stands for its whole, or the whole for its parts (Presmeg 1998). One use of synechdoche in mathematics is when a particular drawing of a triangle, for instance, is taken to stand for the class of all triangles. It might be debated whether the relationship in this case is iconic or indexical. But if even as direct a link as a photograph is both iconic and indexical, as discussed earlier in this chapter, then I am inclined to say that a drawing of a triangle, taken to represent all possible triangles, is both iconic and indexical. Not only is structure preserved across the domains of particular and general triangles (icon/metaphor), but there is also a virtual "physical" link (index) – notwithstanding that it is impossible to draw a general triangle! Some of the problems associated with generality, and the ancient philosophical question of the status of universals, are relevant in a semiotic analysis, and these will be the topic of the next section.

4. COMMENS AND GENERALITY

Bertrand Russell wrote in 1912, "Relations ... must be placed in a world which is neither mental nor physical" (Russell 1959, 90). His argument for this claim hinged on two sentences he gave as examples. If I say, "I am in my room," does the preposition "in", which defines the relation, exist in the same sense that "I" and "my room" exist? No, the preposition does not share this physical existence. Is the relation the work of the mind? The answer is no, because it is not thought that produces the truth of the proposition. To clarify even further, he suggested, "An earwig is in my room." This statement may be true even if no human is aware of it, or cognizant of what an earwig is. Such relations are examples of *universals*. Plato, in attempting to address the existence of universals, called them "forms" or "ideas" – pure essences, like justice or whiteness, which cannot exist in the world of sense because they are not particular. Plato's solution put them in the supra-sensible, unchangeable world of ideas. But Russell pointed out that there are four kinds of universals, namely, substantives other than proper names, adjectives, prepositions, and verbs. Plato's solution, like those of Berkeley and Hume much later, concentrated on the first two types, ignoring relations as universals and concentrating on qualities such as "triangularity". In Russell's judgment, this omission led to error. If one thinks of whiteness, it is not the whiteness that is in the mind, but the act of thinking of whiteness. Common nouns and adjectives refer to particular things, even if the qualities are universal. It is the latter two types of universals, prepositions and verbs, which refer to relations. Universals exist then, but they are not "merely mental." By this is meant that "Whatever being belongs to them is independent of their being thought of or in any way apprehended by minds" (Russell 1959, 97).

The relevance of these issues for the semiotic discussion in this chapter is that the world of universals is "delightful to the mathematician" (Russell 1959, 100). More than that, mathematics, like language, could not function without universals. Semiotics casts light on why this is so. Peirce wrote as follows, in 1906:

> I use the word "*Sign*" in the widest sense for any medium for the communication or extension of a Form (or feature). Being medium, it is determined by something, called its Object, and determines something, called its Interpretant. ...In order that a Form may be extended or communicated, it is necessary that it should have been embodied in a Subject independently of the communication; and it is necessary that there should be another Subject in which the same Form is embodied only in consequence of the communication. The Form (and the Form is the Object of the Sign), as it really determines the former Subject, is quite independent of the sign; yet we may and indeed must say that the object of a sign can be nothing but what that sign represents it to be. Therefore to reconcile these apparently conflicting truths, it is indispensable to distinguish the *immediate* object from the *dynamical* object. (Peirce 1998, 477)

Relations, as represented by prepositions and verbs, refer to the latter class of dynamical objects. They are what Peirce called "dicisigns" because they signify the "form this sign represents itself to represent" (478). In the example he gave, "John is in love with Helen," John and Helen are the immediate object, but the dynamical object, the dicisign, resides in the relation "is in love with." Thus Peirce avoids the error of considering only qualities as universals, and embraces all four of Russell's categories in universals that a representamen might stand for. This opens the door to communication of mathematical structure by means of representamen and interpretant. It is likely that the whole debate about Platonism in mathematics conflates this distinction between immediate objects and dynamical objects.

In communication of mathematical ideas, there is a further Peircean construct that is relevant, and that is his trichotomy of three forms of interpretant:

> There is the *Intentional* Interpretant, which is a determination of the mind of the utterer; the *Effectual* Interpretant, which is a determination of the mind of the interpreter; and the *Communicational* Interpretant, ... which is a determination of that mind into which the minds of utterer and interpreter have to be fused in order that any communication should take place (Peirce 1998, 478, his emphasis).

It is this fused mind that he called the *commens*. He characterized the commens as all that is well understood between utterer and interpreter at the outset, in order that the representamen should fulfill its function. The commens clearly partakes of the context, and thus serves a metonymical purpose. Whether the communication is spoken or written on a chalkboard or transparency for an overhead projector, the commens is an important construct for mathematics education.

In this section I have barely skimmed the surface of the deep issue of how semiotics can inform the communication of mathematical ideas – which is the purpose of mathematics education. In the next section I return to how types of representamen, and the interpretants they determine by means of the commens, are implicated in the learning of mathematics.

5. USE OF METAPHOR AND METONYMY IN MATHEMATICAL SEMIOSIS

I shall open this section with an anecdote from an episode in a college level geometry course, *Euclidean and non-Euclidean geometry*, which I taught in spring of 2003. The topic of discussion was the Nagel point of a triangle, that is, the point of intersection of the line segments from the vertices of the triangle to the points of tangency of the opposite escribed circles. In figure 2, N is the Nagel point of triangle ABC. To establish the existence of the Nagel point by Ceva's theorem, it is necessary to prove:

$$\frac{AF}{FB}\cdot\frac{BD}{DC}\cdot\frac{CE}{EA}=1 \qquad (1)$$

In preparing for class several days earlier, I had proved the relationship at home without difficulty, and because this particular day turned out to be a very full one, I did not check the proof before class, confident that I knew how to do it. After I had drawn the diagram on the chalkboard, without the lower-case notation for the line segments shown in figure 2, and we had established together what was required to prove, the students were given time to think, and later to talk with others in the established groups of four students that had become part of classroom practice. Not one of the 29 students could even approach a solution. What was worse, neither could I! I remembered that the proof had used the following facts:

$$AB + BD = AC + CD;\ BC + CE = BA + AE;\ \text{and}\ CA + AF = CB + BF \qquad (2)$$

So I wrote this on the board, hoping to give the students and myself a clue about how to proceed (having told the students – to their delight and amazement – that that was genuinely all I could remember of the solution). After the 50 minutes of class, the problem was still unsolved, and I assigned it for homework.

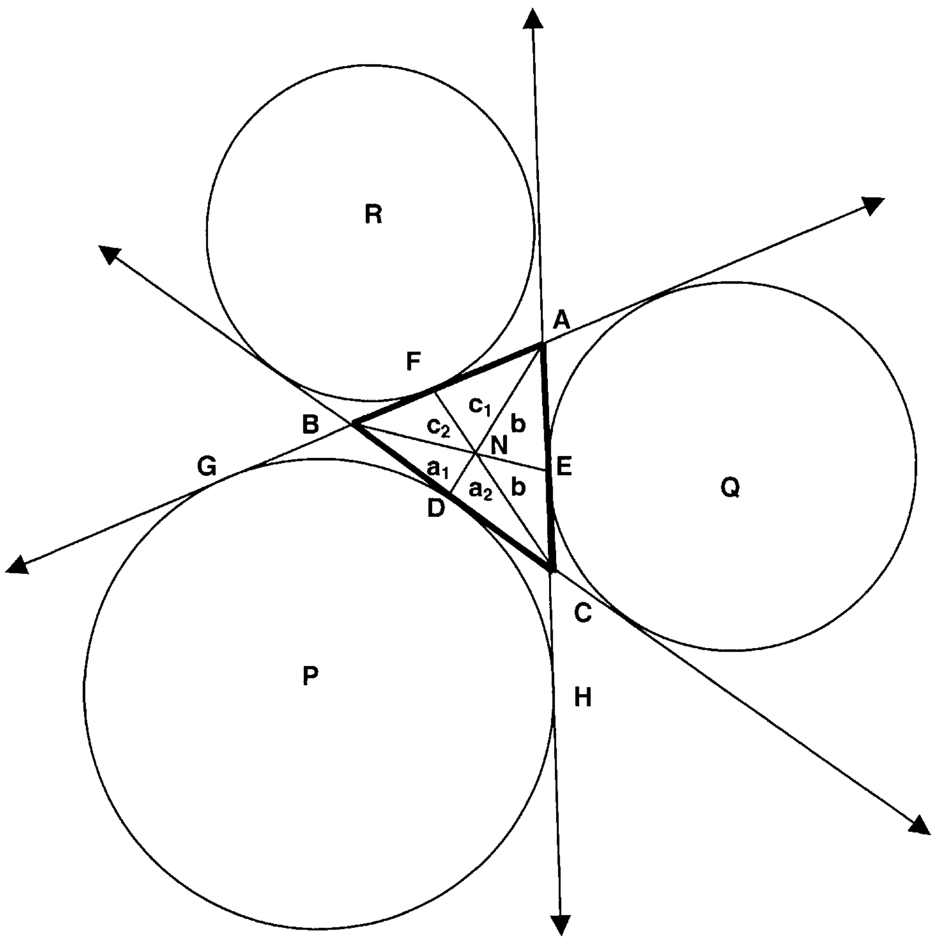


Figure 2. Illustration of the Nagel point, N, of triangle ABC.

Back in my office, it took only a few moments, glancing at my previous work, to see that the proof was not difficult if one used a different notation, the lower-case letters in figure 2, to stand for the line segments involved. In fact, the proof came out in seven lines.

Proof:

$AG = AH$ (tangent segments from same point to circle P).
Similarly $BG = BD = a_1$ and $CD = CH = a_2$
Therefore $c_1 + c_2 + a_1 = b_2 + b_1 + a_2$ = semi-perimeter of triangle ABC.
Similarly $a_1 + a_2 + b_1 = c_2 + c_1 + b_2$ = semi-perimeter of triangle ABC,
and $b_1 + b_2 + c_1 = a_2 + a_1 + c_2$ = semi-perimeter of triangle ABC.
Therefore $a_1 = b_2$, $b_1 = c_2$, and $c_1 = a_2$.
Then

$$\frac{a_1}{a_2} \cdot \frac{b_1}{b_2} \cdot \frac{c_1}{c_2} = 1 \qquad (3)$$

and the concurrency follows (by Ceva's theorem).

The following day, the students (who had not been able to solve it for homework), had little difficulty working out the proof with the support of group discussions, once I introduced the new lower-case symbols in the diagram.

This episode was dramatic, and left me wondering what it was about the commens or the symbolism that led to such different results in the classroom practices on the two days. The first day had been frustrating for the students and for me; the second resulted in some students expressing enough confidence in their reasoning for them to be willing to come and demonstrate it on the board. It was, after all, the same structure that was represented in the notation I had used on the first day as on the second. But in semiotic terms, the representamen on day one,

$$AB + BD = AC + CD, \qquad (4)$$

was not a sufficiently iconic likeness of the semi-perimeter of the triangle to make it clear, firstly, that the object *was* a semi-perimeter, and secondly, that each of the six given expressions represented a semi-perimeter, and that they were therefore all equal to one another. Somehow the representamen

$$c_1 + c_2 + a_1 = b_2 + b_1 + a_2, \qquad (5)$$

together with the similar equations, had helped to make the relationships apparent. Both of these representamens were symbolic: both were conventional ways of expressing relationships among the lengths of line segments. "A symbol is a representamen – a rule that will determine its interpretant" (Otte 2001, 14). In both cases, the interpretants were all-important in determining the outcomes of the actions taken, successful or otherwise. The difference between the two symbols appears to reside in their metaphoric and metonymic features, that is, how they behaved as icons and indices in representing the abstract structure of the configuration of circles and lines. By breaking each side of the triangle into two constituent line segments, the lower-case symbols served to highlight basic building-blocks of the structure, thus subtly pointing to a commonality that was less easy to deduce from the more holistic upper-case symbols. As a metaphor, the icon of the lower-case representamen was more efficient in suggesting an interpretant that would unpack the relationships.

I have not yet mentioned the indexical features of the two notations. In each case, there was a physical connection between the representamen and its object, the mathematical relationships among the line segments. Thus both were clearly indices. As metonymies, they pointed to subtly different contexts – contexts that were sufficiently different to result in different interpretants, one unsuccessful and the other successful, in attempting the proof. I am not claiming that these interpretants would be identical for everybody. It is possible that some readers may have completed the proof using only the upper-case notation at the outset.

In spite of the symbol being a rule, governed by the commens, that will determine its interpretant, there is sufficient individual variation in the construction of an interpretant to justify Peirce's insertion of the phrase "upon a person" – his "sop to Cerberus" – in the definition of a sign (as representamen) in one of his letters to Lady Welby:

> I define a Sign as anything which is so determined by something else, called its Object, and so determines an effect upon a person, which effect I call its Interpretant, that the latter is thereby mediately determined by the former (Peirce 1998, 478).

Otte (2001) claimed that "the interplay of iconic and indexical representations [is] most important to understand mathematical cognition" (22). Then we cannot avoid the use of both metaphor and metonymy in mathematics, and semiotics can provide a tool for helping us to understand the relationships involved. Regarding mathematics as "a constructive and visual art" (ibid.), would imply that metaphorical thinking plays a fundamental role, in "seeing an *A* as a *B*" (as Otte expressed it). Then the question that he asked is crucial: What leads us in creating good metaphors? After all, "everything seems similar to everything in at least some aspects" (22). This statement may be too broad; but the point is that the comparison of domains needs to be useful for a mathematical purpose. The ability to see structure across domains – discarding irrelevant details – is one that Krutetskii (1976), based on his research, considered to be a vital component of effective mathematical thinking. These issues will continue to be of significance for mathematics educators.

Illinois State University, Normal, IL

REFERENCES

Krutetskii, V. A. (1976). *The structure of mathematical abilities in schoolchildren.* Chicago: University of Chicago Press.

Otte, M. (2001). *Mathematical epistemology from a semiotic point of view.* Paper presented in the Discussion Group, Semiotics in Mathematics Education, at the 25th Annual Meeting of the

International Group for the Psychology of Mathematics Education, Utrecht, The Netherlands, 12-17 July 2001.

Peirce, C. S. (1998). *The essential Peirce. Volume 2*, edited by the Peirce Edition Project. Bloomington: Indiana University Press.

Presmeg, N. C. (1998). Metaphoric and metonymic signification in mathematics. *Journal of Mathematical Behavior 17.*1, 25-32.

Presmeg, N. C. (2002). A triadic nested lens for viewing teachers' representations of semiotic chaining. In F. Hitt (Ed.), *Representations and mathematical visualization.* Mexico City: Cinvestav University, 263-276.

Russell, B. (1959). *The problems of philosophy.* London: Oxford University Press.

Saussure, F. de (1959). *Course in general linguistics.* New York: McGraw-Hill.

Sfard, A. (1991). On the dual nature of mathematical conceptions. *Educational Studies in Mathematics 22,* 1-36.

Whitson, J. A. (1994). Elements of a semiotic framework for understanding situated and conceptual learning. In D. Kirshner (Ed.), *Proceedings of the 16th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Baton Rouge, Nov. 5 – 8, 1994, Vol. 1, 35-50.

Whitson, J. A. (1997). Cognition as a semiosic process: From situated mediation to critical reflective transcendance. In D. Kirshner & J. A. Whitson (Eds.), *Situated cognition: Social, semiotic, and psychological perspectives.* Mahwah, New Jersey: Lawrence Erlbaum Associates, 97-149.

ANNA SIERPINSKA

ON PRACTICAL AND THEORETICAL THINKING AND OTHER FALSE DICHOTOMIES IN MATHEMATICS EDUCATION

Abstract. I owe much of my understanding of the difference between synthetic and analytic thinking in mathematics to my reading of Michael Otte's papers and the conversations we had with him within the BACOMET group. One of the first sources of inspiration for me has been his work on arithmetic and geometric thinking. In the paper I shall outline the consequences of the distinction for analyzing processes of mathematics teaching and learning in my own research. I shall further use this distinction to look critically upon the recent trend in mathematics education of considering mathematics as a kind of "discursive practice."

Key words: epistemological obstacles, manipulatives, mathematics, practical thinking, Pythagoras theorem, teaching, theoretical thinking.

1. INTRODUCTION

This paper, dedicated to Michael Otte, is about practical and theoretical thinking as complementary epistemological categories and the use of this distinction in mathematics education. The distinction is presented as one among many "false dichotomies" that are common in the domain. The dichotomies are first discussed in the light of Michael Otte's papers on complementarity. An alternative view is then proposed in terms of couples of epistemological obstacles. A possible use of the practical/theoretical distinction in mathematics education is illustrated by means of a thought experiment about a teacher educator planning to discuss the use of manipulatives with his student teachers. The thought experiment points to the complexity of the system of objects of thought in mathematics education and the extreme fragility, in practice, of the distinction between theoretical and practical thinking. It also highlights the crucial role that epistemological analyses, such as those offered in Michael Otte's papers, play for research in mathematics education.

In his comments on one of my papers about epistemological obstacles (Sierpinska 1996), Michael Otte was saying that, where I saw a couple of obstacles, he could see only one, namely

> … the problem that in order to understand mathematics one has to take into account [the fact] that mathematics is simultaneously meta-mathematics … [T]he problem lies in an empiricist or concrete epistemology [that] does not think of mathematical objects in relational or structural terms. … [M]athematics is difficult for the learner not because of the technical complications of its method, but because of the specificity of its objects. (Letter dated 24.1.1994)

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 117 – 135.

He went on to say that he has been busy his entire didactical career with this one problem [the problem of sources of difficulties in mathematics learning] and with the question of the nature of mathematical objects and concepts. In my own research, I have tried to engage more directly with the practice of teaching, with designing and experimenting didactic sequences. Somehow, I always ended up discussing these same problems. They are very powerful attractors indeed in the dynamics of research in mathematics education.

2. EPISTEMOLOGICAL OPPOSITES

Theory of mathematics education is replete with pairs of opposite categories of knowing and thinking such as the empiricist-structural distinction mentioned above, intuition versus formal knowledge, instrumental versus relational, or operational versus structural understanding. In my research I first resisted using such global categories, explaining both the meaning of particular mathematical concepts and students' difficulties by the existence of "epistemological obstacles" specific to concrete mathematical concepts. But, as I went on in my research, I realized (and thus agreed with Michael) that many obstacles were related not to specific concepts but to mathematics in general. And thus I ended up with, first, three categories of thinking in linear algebra: synthetic-geometric, analytic-arithmetic and analytic-structural, and then attributing students' difficulties in linear algebra to their tendency to practical as opposed to theoretical thinking (Sierpinska et al. 1997; Sierpinska 2000; Sierpinska & Nnadozie 2001).

It is tempting to think that these categories refer to some ontological reality; that there exists an identifiable brain activity such as, for example, theoretical thinking, with no trace whatsoever of its opposite, namely practical thinking. But, as Michael Otte has argued in Otte (1990b), these distinctions should be regarded as epistemological, not ontological distinctions. They are our simplified ways of knowing human cognitive activity in mathematics; they are not kinds of human cognitive activity.

> I have argued that, whenever we see mathematical proof as involving only a mechanical aspect, we are driven to see that it involves, as well, an intuitive one. And whenever we are tempted to see mathematical proof as involving only a solitary aspect, we are driven to seeing that it is also a social matter. And whenever we are tempted to see a mathematical argument of the kind found in proof, namely a chain of tautologies or of equalities, as merely, or perhaps the ideal of, literal expression, we are forced to see that it is, in fact, essentially metaphorical. (Otte 1990b)

This is why Otte preferred to speak of "complementarity" (*p* and not *p*) rather than of dichotomy (*p* or not *p*).

3. COMPLEMENTARITY

In his philosophical considerations on mathematics and its teaching, Otte has explored in depth the idea of complementarity of object and method in science (Otte 1990a), or, broadly speaking, the idea that "every scientific explanation simultaneously contains a meta-communication, i. e. it represents, in an exemplary way, an answer to the question what it means to explain an object or a fact at a certain historical point in time." This notion of complementarity comprises issues such as relationships between intuition (which focuses on discovering the object of study) and logic (whose problem is to systematize methods of validation of the findings), immediate perception (synthetic thinking) and discursive procedures (analytic thinking), or between theoretical representation and technology of measurement or computational technique. These issues constitute the philosophical underpinnings of debates on the teaching of mathematics focusing on the problems of striking a balance between "theory" and "practice," knowing why and knowing how, letting the students engage in free explorations and express themselves as they like and teaching them the "right" mathematical discourse and standards of methodological rigor.

Complementarity of these categories could be expressed also in terms of epistemological obstacles. An epistemological obstacle is a way of thinking that stands in the way of another way of thinking, but it would not exist (as an obstacle) without this other way of thinking. Thus it does not make sense to speak of single epistemological obstacles, but only of their pairs. The epistemological categories mentioned above can be seen as pairs of epistemological obstacles in the philosophy of knowledge. Intuition and formal knowledge is such a pair of obstacles, for, without intuition, formalism would have nothing to doubt; there would be no need to formalize in order to confirm or remove the doubt; on the other hand, without formalism, intuition would remain in a state of either permanent self-satisfaction or permanent doubt.

Similarly, theoretical and practical thinking can be viewed as a pair of epistemological obstacles. Thinking is not either theoretical or practical but arises in a tension between the two. The "practical sense" decisions are acts of discarding all but one possible course of action; but this decision would not be necessary if these possible courses of action were not available to the mind. They are available as a result of hypothetical, theoretical thoughts, however primitive, swift and unconscious. On the other hand, the mind would not engage in thinking about the possible courses of action and their outcomes if no action were envisaged at all. As Otte was saying, in his polemic with Piaget's concept of empirical abstraction, which he considered "too primitive" by being completely separated from reflective abstraction (Otte 1990a):

> One has to emphasize that theoretical consciousness demands to conceive the objects and phenomena of reality not just in the form of knowledge and contemplation but as parts of activity also ... [T]he relationship between the conceptual-reflective and the algorithmic-logical elements of mental activity is only conceivable as an interaction of two poles of a relationship the basis of which is the activity. (Otte 1990a)

"This is all very well" – a mathematics teacher might say at this point – "but what difference does it make for my teaching practice, whether I see these pairs of categories as dichotomies or as complementary couples?"

4. THE QUESTION OF RELEVANCE OF EPISTEMOLOGICAL DISTINCTIONS FOR MATHEMATICS EDUCATION

Saying that mathematical thinking is, at the same time, intuitive, formal, practical, and theoretical is anything but an astounding discovery. Of course, the realization of the epistemological complementarity discussed above might save one the inevitable failure of organizing one's teaching on the basis of the assumption that, say, "real" mathematical thinking is formal and theoretical, and that the intuitive and practical aspects of knowledge construction are only the necessary contingency of some shameful "didactic transposition" from scholarly research knowledge to the social and cultural institution of teaching mathematics to masses of students. It might save one, as well, from trying to "derive" theoretical concepts from concrete, "hands-on" experience, based on the belief that the meaning of these concepts is somehow already there in the empirical relations. Steinbring has convincingly demonstrated the ineffectiveness of such approaches, using the theory of epistemological triangle and detailed analyses of classroom interactions (e. g. Steinbring 1991, 1993).

But the mere realization of complementarity cannot help mathematics educators in understanding what exactly is difficult in learning this or that mathematical idea in a particular teaching/learning situation, never mind helping them in planning and organizing such situations. In each case, the mathematics educator must "roll up his sleeves" and do the epistemological and didactic analysis almost from scratch. This is no trivial task, as can be seen from the above-mentioned papers by Steinbring.

The mathematics educator must also be more specific in describing the epistemological categories if he[1] intends to use them in analyzing particular teaching situations; he needs to "operationalize" them. With respect to the theoretical/ practical distinction, for example, saying that learning mathematics is difficult because it requires theoretical thinking is almost a tautology. In the next section I present a characterization of the theoretical/practical distinction, which we developed for the purposes of our research on linear algebra teaching and learning (Sierpinska et al. 2002).

5. A CHARACTERIZATION OF THEORETICAL/PRACTICAL THINKING

Michael Otte once told me that the difference between synthetic and analytic thinking is that the former holds a direct relationship with its object while in the latter this relationship is mediated by one or more sign systems. The same can be said of the difference between practical and theoretical thinking, since it is normally assumed that theoretical thinking is analytical.

Thus what is theoretical or practical is not thinking as such but *the relationship between thinking and its object.* It makes sense to conceive of this object as some kind of action, actual or imagined, present or past, performed or planned, since both theory and practice are normally related to purposeful action. "Action," here, could mean proving a mathematical statement as well as carving squares from a plank of wood. Practical thinking could be viewed as *thinking-in-action*, whereby changes in thought directly influence changes in action. Thus, the relationship between thinking and this very action of thinking is practical. If a philosopher ponders a theoretical question, the relation of his thinking to this very activity of thinking is necessarily practical; he is not thinking about his thinking. He is just thinking; he is engaged in the practice of philosophizing.

For theoretical thinking to even begin, the thought and its object must belong to different planes of action (Figure 1). Thinking-*in*-action must become thinking-*about*-action. The moment the philosopher reflects back on his thinking, verifying if it is well founded, eoretical.

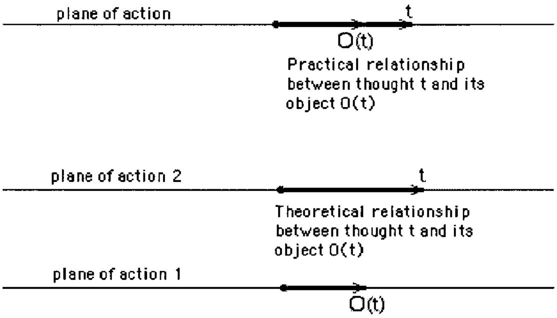


Figure 1. Relations between thought and its object in theoretical and practical thinking.

Let me illustrate this idea with one more example. Imagine a student who solves an equation and then substitutes the obtained result into the original equation. In the phase of solving the equation, the student is engaged in the practice of solving equations: his thinking and his activity of processing the algebraic expression belong to the same plane of activity. In the phase of substitution, the student may be taking a step back from his previous activity, which would now become the object of his thinking. He may be verifying if the result he obtained indeed satisfies the equation. In this case, we could say that the student is "engaged in theoretical thinking" or, more precisely, that the relationship between his thinking and its object is theoretical. But the student may also do the substitution as part of what he understands as the school task of "solving an equation," without viewing it as a means of control of the result obtained in the first phase. It is well known that many students indeed hold this conception and are not bothered if they obtain a contradiction through substitution. These students think practically in both phases of the task.

Obviously, one cannot assume that belonging to different planes of action is a sufficient condition for the relation between thought and its object to be theoretical. Musing about days gone by, day-dreaming, or rotating three-dimensional shapes in one's mind would then count as theoretical thinking and this is not what we intend to mean. More restrictions on the relationship between thought and its object are needed for a satisfactory characterization of theoretical thinking.

The most obvious characteristic of what we normally call *theoretical* thinking is that its ultimate purpose is the production of theories or conceptual systems.

One consequence of this assumption is that theoretical thinking is not about techniques or procedures for well-defined actions, although these might be derived from or explained by the theories. Theoretical thinking is *reflective* in that it does not take such techniques or procedures for granted but considers them always open to questioning and change. In this sense, therefore, theoretical thinking is opposed to mythical thinking, in which knowledge is considered as "natural" or "sacred" and therefore in no need for justification (Steinbring 1991).

Another consequence is that theoretical thinking is *systemic*, i. e. its objects are not particular actions but systems of relations between actions, and systems of relations between these relations. As Otte was saying,

> The history of science may be briefly sketched as a transition from thinking about objects to relational thinking. Theoretical thinking, accordingly, is not concerned with concrete objects, nor with intrinsic properties of such objects, and theoretical terms, in particular, are not just names of objects. Rather, science is concerned with the relationships between objects or phenomena. As the historical transition took place, it became increasingly obvious that a theoretical term will receive its solid content, its clear form, only from its relationship to other concepts. (Otte 1990a)

The systemic character of theoretical thinking entails *sensitivity to contradictions*; otherwise, conceptual systems would collapse. Vygotsky has particularly stressed this characteristic of scientific, as opposed to everyday concepts (Vygotsky 1987, 234). Actually, the very concept of contradiction makes no sense outside a system of concepts. Contradiction is a type of logical relationship between propositions; there can be no contradiction between events occurring in space and time; their meanings change with the context in which take place. Contradiction thus requires *stability of meanings* in the frame of reasoning. This can be achieved by definitions and other agreed upon characterizations.

The combination of reflective and systemic thinking implies that theories do not grow by simple addition of new concepts, but that new developments may cause a restructuring of the whole system. The system is always reflected upon as a whole. This feature of theoretical thinking is sometimes called "*reflexivity*" (Steinbring 1991).

Concern with non-contradiction implies that attention is being paid to problems of *validation*, both at the level of the systems themselves and at the meta-level, i. e. at the level of *methodology*. Theoretical thinking asks not only, *Is this statement true?* but also *What is the validity of our methods of verifying that it is true?* Thus theoretical thinking always takes a distance towards its own results.

Thinking within conceptual systems can only produce conditional truths; it is *hypothetical thinking*. Theoretical thinking is concerned with problems of the sufficient, necessary, essential, complete character of conditions of truth in each case.

As mentioned, the assumption of belonging to different planes of action already implies that theoretical relationship between thought and its object is *analytic*, i. e. mediated by systems of signs. But, if we assume that the results of theoretical thinking are conceptual systems or theories, which have to be formulated in some coherent terminology and symbolic notation, then we must also require that theoretical thinking have an analytic relationship with sign systems themselves. Theoretical thinking not only is mediated by systems of signs; it takes systems of signs as an object of reflection and invention.

In brief, theoretical thinking is thinking where thought and its object belong to distinct planes of action, and whose purpose is the production of internally coherent conceptual systems, based on specially created systems of signs. Theoretical thinking is, therefore, reflective, systemic and analytic.

I have argued elsewhere (Sierpinska et al. 2002) how highly relevant, a priori, are the above features of theoretical thinking in understanding linear algebra, and how irrelevant they can be for high achievement in linear algebra courses. In this paper, I will focus on the complementarity between theoretical and practical thinking in actions related to teaching and learning of mathematics.

6. A THOUGHT EXPERIMENT: THE INTERPLAY OF THEORETICAL AND PRACTICAL THINKING IN A TEACHER EDUCATOR'S PLANNING OF AN ACTIVITY ON THE USE OF MANIPULATIVES WITH STUDENT-TEACHERS

Thinking in and about mathematics education involves simultaneously several planes of action. In particular, the object of study for a mathematics education researcher may comprise several levels of recursion of the act of "theoretical reflection on practice." For example, when a researcher reflects theoretically on the practice of a teacher educator, he may use his practical experience of being a teacher educator, a schoolteacher, a learner and doer of mathematics, and a researcher knowledgeable of the theories and methodologies of his field. He may entertain, with each of these planes of action, a practical or a theoretical relationship.

In any concrete activity of reflection, these relationships are closely intertwined and dependent on each other. Their identification and categorization is possible in a methodological analysis, but not in actual fact. This is what I would like to show in the following thought experiment.

Suppose a researcher reflects on the work of a teacher educator preparing an activity for his student teachers aimed at a reflection on the use of manipulatives in mathematics teaching, on the example of the learning, by high school students, the meaning of the Pythagorean theorem. In the first section (6.1), the narrator is the hypothetical teacher educator. In the second (6.2), a researcher interprets and analyzes the actions of the educator, focusing on the interplay between his theoretical and practical thinking.

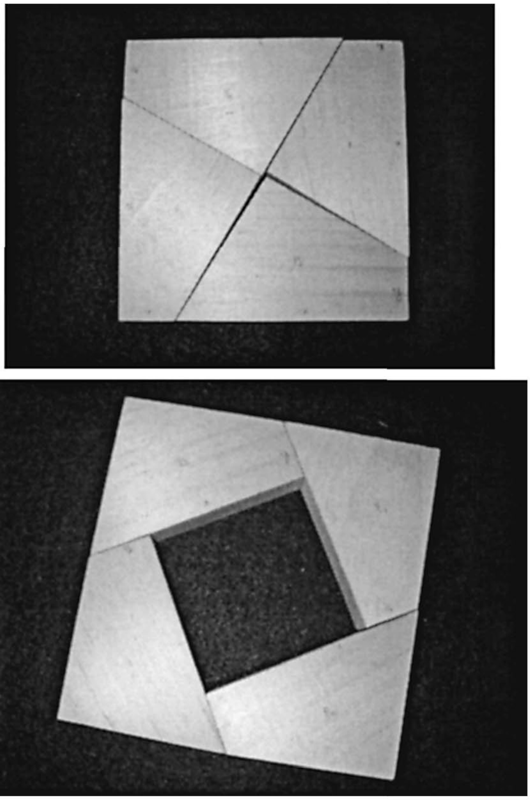
6.1 Teacher educator prepares a class on the use of manipulatives

[1] Suppose I am a teacher educator preparing a session with student teachers on the problem of using manipulatives in the teaching of mathematics. I want to convince them that mathematics is not there, in the manipulatives, but, at best, in the interplay between the practical and theoretical tasks based on actions with the manipulatives.

[2] Let me prepare for a worst-case scenario. Suppose the pre-service teachers in my class want a straightforward judgment such as, "manipulatives are good" (or bad). Also, suppose they expect that teaching with manipulatives is easy:

one just goes into the classroom with a bag of manipulatives, lets the students play with the them, and the students thus "naturally" discover the mathematical concept planned for this particular lesson.

[3] What situation could help my students realize that there is no simple recipe and that it all depends on the manipulatives, what you want to teach with them and how you set up the didactic situation? I know that just telling teachers "it depends" will not help them understand the complexity of the issue. I need to engage them in planning a concrete lesson with concrete manipulatives. Suppose I take the wooden puzzle that I got at the last NCTM[2] meeting and ask student teachers to imagine if and how they could use it to introduce the Pythagorean theorem.



Figures 2a & 2b. Two ways of arranging the pieces of the puzzle.

[4] This, I feel, is bound to show them that while manipulatives may embody mathematical ideas for those who already have them in their minds, they are not necessarily helpful in bringing these ideas to the minds of those who hadn't seen them before.

[5] The puzzle has 4 pieces, which can be arranged into shapes like those in Figure 2.

[6] This pair of shapes brings to mind the "puzzled" proof of the Pythagorean theorem, as in Figure 3.

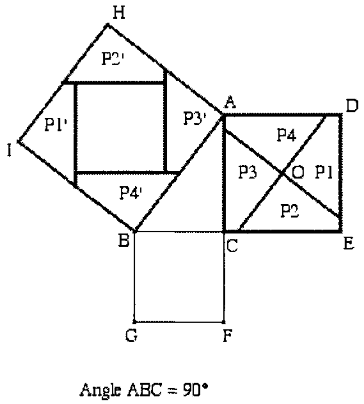


Figure 3: The idea of the popular "puzzled" proof of the Pythagorean theorem.

[7] The student teachers will probably recognize the Pythagorean theorem in the puzzle, and they will take it for granted that their students will "see" it as well, in spite of never having heard of the theorem before.

[8] I will show them that this need not necessarily be so. I will invite them to imagine, step by step, what may happen if they bring the puzzle to the classroom and ask the students to first play with it freely and then to construct squares. I will ask them to assume that students in the classroom are mostly practically minded. I don't know what scenarios they may come up with, but let me do this exercise myself, so I can be better prepared for arguing with their claims.

[9] Most students want to make nice looking material objects. They do not think of a shape first and then try to construct it, but just move the pieces around, trying in which ways they best "fit" with each other. Their decisions about when to stop and consider the shape done are based on visual and tactile clues and their spontaneous esthetic feelings. These may be explained by their previous encounters with cultural artifacts, but not by some explicit esthetic principles such as "symmetry," "compactness," or "balance" (Figure 4).

[10] If the students only want to play with the puzzle in this rather random fashion, they will never be brought anywhere close to the Pythagorean theorem. Let me now think of the next-to-worst scenario. The students start noticing some relations between the pieces. They might discover that the four pieces of the

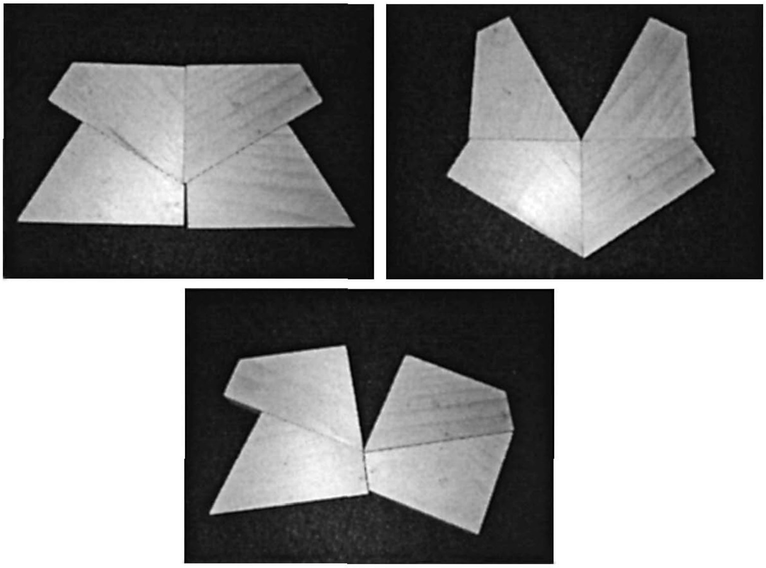


Figure 4. Shapes that could be obtained by students through free play with the puzzle.[3]

puzzle are not identical. The lengths of their sides differ a little. Especially one piece is quite off the shape of the other three. Also the angles that look like right angles are not exactly so, because, when the pieces are put side by side, they do not form a straight line exactly (Figure 5).

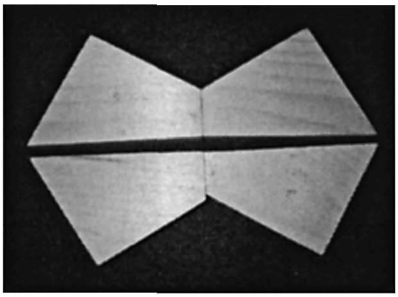
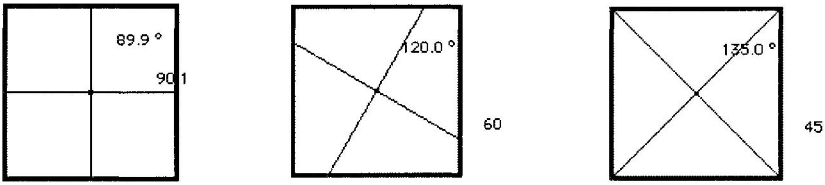


Figure 5. The pieces of the puzzle are not all identical.

[11] Students may decide to ignore the differences (as technical errors of the person who cut the pieces). Suppose now that some students are technically minded or have been inspired by their recent experiences in the woodwork class. Some of them may start thinking about the technology of producing the puzzle. This may lead them to viewing each puzzle as made from a single square piece of wood (like in Figure 2a) cut along two perpendicular lines passing through the center of the square, constructed as the intersection of its diagonals. Some stu-

dents may measure the angles at which the inner segments fall on the sides of the square, find that they are approximately 60° and 120°, and include these measures in their definition of the puzzle. Other students may see these angles as arbitrary and only constrained by the requirement of producing a non-trivial puzzle, i. e. one made of four quadrilaterals with unequal sides and not four squares or four right-angled triangles (Figure 6).



Figures 6a, 6b, 6c. Two trivial and one non-trivial puzzle.

[12] These technological concerns of the students could perhaps be considered as the most "natural" outcome of playing with the puzzle in a high school mathematics class. This situation could give the teacher an opportunity to generalize the puzzle as a set of four identical quadrilaterals with two opposite right angles and the sides of one of the right angles[4] being equal. The other two angles add up to 180°, because the sum of angles in a convex quadrilateral is equal to 360°. Thus, if one of the angles measures a, the other measures 180° – a. If a = 90°, the piece is a square[5] (Figure 6a); if a = 135° the piece is a triangle (Figure 6c).

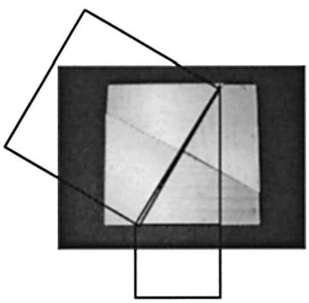


Figure 7. The three squares seen as built on the sides of a right angled triangle.

[13] The question is: Is it at all possible to bring students to think about the Pythagorean relation from playing with the puzzle? Is there a best-case scenario? Suppose the students construct the squares in Figures 2, either by themselves or in response to the teacher's explicitly formulated task. Suppose they even notice that there are three squares in these two figures and that the area of the external square in Figure 2b is equal to the sum of the square in Figure 2a plus

the area of the square built on a segment which can be seen as a certain part of the side of the square in 2a. Suppose, even more optimistically, that they notice that all three squares can then been seen as built on the sides of a right-angled triangle (Figure 7).

[14] Making these observations requires that the students be highly theoretically oriented. It requires seeing the shapes obtained with the puzzle as structures composed of segments of different lengths and mutual positions. It also requires reflecting about the relations between the different shapes obtained with the puzzle (possibly in a situation where only one shape is available to the senses at a time). These observations are not a result of direct visual and tactile perception: they are a result of a construction of a geometrical model of the puzzle (Figure 8).

[15] Communication of these observations among students would require coding the different segments of the pieces of the puzzle. Students would not know where their observations would be leading them, so they would be likely to use ad hoc representations, such as color. However, color is not functional if algebra is to be used later on in the representation of the Pythagorean relation and its proof. If the teacher imposes a notation, this will immediately destroy the "naturalness" of the situation. The students will know that their initiative does not count and this is not real exploration but the well known ritual of fake "discovery teaching," where students are left in the dark till they are eventually explicitly told what they were expected to have discovered. But suppose that somehow students are brought to using letters to denote lengths of segments, as in Figure 8.

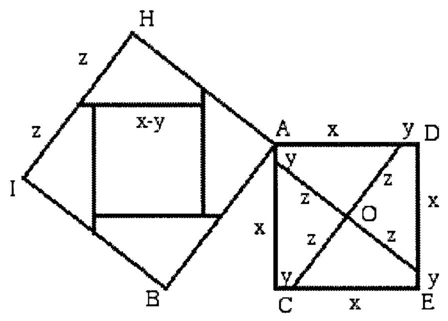


Figure 8. Using a diagram to compare the sides of the three squares.

[16] The students would be probably quick to notice, but also take it for granted that, in the left-hand side square in Figure 8, the side of the external square is 2z and the side of the internal square is $x - y$. It could also be obvious for them from the figure that the side of the right-hand side square is $x + y$. Using the known formula for the area of a square, the students might write the relation: $(2z)^2 = (x + y)^2 + (x - y)^2$. The students would now see in front of them a famil-

iar mathematical object: an algebraic expression. Their aim might become to simplify the expression (to $2z^2 = x^2 + y^2$). This is what they have always done in such situations.

[17] The teacher would not be satisfied with this result. Not because it is not true. It is, but it is also irrelevant from the point of view of his didactic goal. The relation could be obtained directly from looking at the square ACED, by noticing that $4(xy/2 + z^2/2) = (x + y)^2$. Taking y = 0, it could lead to the formula for the diagonal of a square ($2z = \sqrt{2}\,x$), which, in the curriculum, is only derived as a consequence of the Pythagorean theorem.

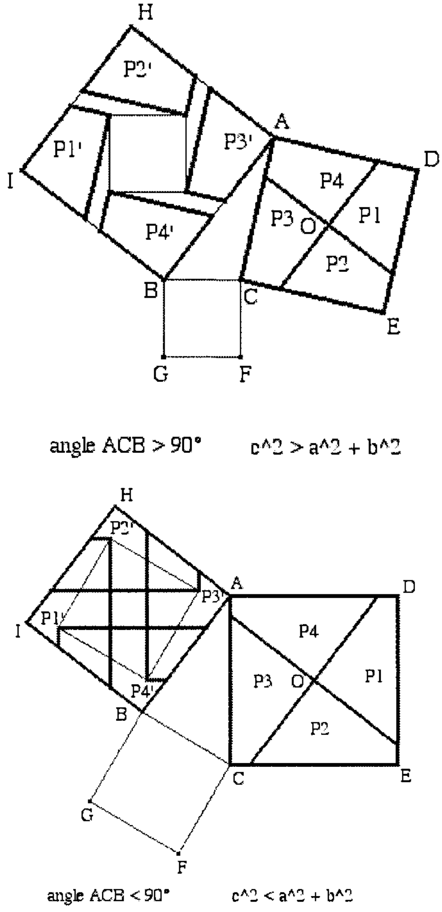


Figure 9. The failure of the Pythagorean identity $c^2 = a^2 + b^2$.

[18] At this point, the student teachers should be convinced that, in order to even start discussing relations among the three squares that can be obtained with the puzzle, students have to forget about the puzzle as a puzzle altogether. They should also realize that students would have to be heavily directed to focus

their attention on the relation between the areas of the squares without simplifying it any further. But even if the hypothetical teacher achieves all that, his students will still be extremely far from the "discovery" of the Pythagorean theorem. This is because, for the puzzle, the relation between the areas is always true. In fact, it is quite obvious and trivial. But the Pythagorean theorem speaks about one very exceptional situation. The rigid wooden puzzle illustrates this one single exceptional situation without hinting at the class of situations of which it is an exception. It is, indeed, quite exceptional that the areas of squares built on two sides of a triangle add up to the area of the square built on the third. It only happens when one of the angles of the triangle is a right angle. The puzzle, in itself, is unable to provoke students to think about the conditions of the Pythagorean relationship between the areas of squares built on the sides of a triangle. At best it illustrates a possible way of proving the theorem once it is realized as a conjecture.

[19] But the student teachers should not be left with the impression that the only way to introduce the Pythagorean theorem is to state it on the board and have the students learn it by heart. They have to understand that their students will not appreciate the significance of the theorem this way, either. Suppose I suggest that student teachers try to imagine starting a lesson by directly asking the theoretical question: What is the relation between the areas of squares built on the sides of a triangle? and allowing their students to work within a dynamic computer environment (Figure 9).

The discussion would then be organized on their views of the potential of this type of more sophisticated "manipulatives" in the teaching of the Pythagorean theorem.

6.2 Analysis of the thought experiment

This section presents a possible theoretical reflection of a researcher on the role of theoretical thinking in the work of a teacher educator planning a teaching activity with student teachers. The analysis will make references to the narrative of the hypothetical educator in the form of paragraph numbers in square brackets. It will also make explicit the evaluation, as theoretical (t) or practical (p), of the narrator's thinking about the practices of research (R), teacher education (E), teaching (T), learning (L), doing mathematics (M). The analysis highlights in italics words that are related to particular features of theoretical thinking.

In [1] the researcher engages in hypothetical thinking ("suppose") about the action of a teacher educator, so his relationship with E is theoretical (tE). The choice of the topic, however, is based on his experience with E; he knows that manipulatives is a "hot issue" and is likely to attract student teachers' attention (pE). He also knows that this is a controversial issue in mathematics education (pR) and has a theory about the epistemological relationship between manipulatives and mathematics (tM). This theory re-surfaces now and again in his reflection ([4], [7], [8], [10], [13], [14], [18]).

In [2] it is the hypothetical teacher educator who speaks. His consciously adopted methodology of preparing a class (tT) is to first "prepare for the worst" and then gradually consider more optimistic scenarios. His worst-case scenario is based on the assumption, founded on his experience with teaching (pT), that his actual student teachers as well as hypothetical teachers and pupils have a strongly practical attitude towards their tasks. He considers this to be the worst-case scenario, because he assumes that mathematics is theoretical knowledge par excellence (tM).

In [3] the educator decides against just telling the student teachers that the use of manipulatives can be more or less effective depending on circumstances. Following perhaps a "socio-constructivist approach" (tT), he plans to confront his student teachers with a specially designed situation, let them draw the conclusions for themselves, and then engage in an argument with them, negotiating alternative ways of thinking. This operationalization of the socio-constructivist epistemology in terms of didactic choices is based on his familiarity with its common interpretations within the community of teacher educators to which he belongs (pE). He chooses to use a wooden puzzle as an example of a manipulative, because it is there on his desk, reminding him of his recent activities with children of various ages playing with the puzzle (pT).

In [5] the educator reflects on (tM) his personal experience with the puzzle; he has played with the puzzle, trying to make mathematically meaningful shapes (pM). Based on this experience, he assumes, in [6], that knowing the Pythagorean theorem allows one to construct a material model of the idea of the proof of the theorem with the puzzle (pM, tM).

In [7] the educator reasons as follows: Since the student-teachers know the Pythagorean theorem, and, according to the worst-case scenario, they hold the naïve belief that mathematical patterns are there in nature and things (tL), waiting to be discovered, it is very likely that they will expect high school students to "discover" the theorem through playing with the puzzle (tT).

In [8] the educator reflects on the possible moves (tT) in this situation, based on his experience as a teacher (pT). The best thing would be to ask the student teachers to actually perform an experiment with a student who has never seen the Pythagoras theorem before. But the constraints of time as well as the practical difficulties of access to such students and of the control of the experiment by the teacher educator make him opt for a collective "thought experiment" instead.

In the sequel of the thought experiment ([9]-[18]), the teacher educator speculates about how his students could be led to the realization of the non-transparency of manipulatives by imagining what could happen in a classroom started by a free play with the puzzle.

The educator imagines the course of events based, again, on his methodology of going from the worst-case scenario to gradually more optimistic scenarios regarding the agents' theoretical thinking (tT, [9], [10]). His speculations are founded on his informal observations of students playing with the puzzle (pT, tL, [9], [10]), his theory of people's relationship to cultural artifacts (tL, [9], [10]), and his reflection on his experience of mathematizing the relationships between the elements of the puzzle (pM, tM, [10], [11], [12]).

In [12] he reflects on the outcome of these speculations (tT); he considers a technical approach to the puzzle as quite natural in students. On the other hand, thinking about the Pythagorean relation in the context of the puzzle does not appear as natural; based on his reflection on his activity of relating the puzzle configurations with the Pythagorean relation, he realizes that this would require noticing unobvious quantitative relations and highly theoretical thinking (pM, tM, [13], [14]). This would also require a graphical representation of two special configurations of the puzzle and a mathematically consistent coding of the elements of the puzzle (pM, tM, [15]).

Para	tR	pR	tE	pE	tT	pT	tL	pL	tM	pM
1		1	1	1					1	
2					1	1			1	
3				1	1	1				
4									1	
5									1	1
6									1	1
7					1		1		1	
8					1	1			1	
9					1	1	1			
10					1	1	1		1	1
11									1	1
12					1				1	1
13					0				2	1
14									2	1
15							1		1	1
16					1		1		1	
17					1		1		1	
18					1		1		2	
19					1				1	1
Sum:	0	1	1	2	11	5	7	0	20	9

Figure 10. A summary of the interplay of theoretical and practical thinking in the course of the educator's work of preparing his classroom activity. Theoretical thinking was invested mostly into the educator's relationship with the practice of doing mathematics and the practice of teaching. This thinking was strongly supported by the educator's experiences in these domains of practice. His thinking about learning was more speculative.

These reflections lead the educator to point to the shaky foundations of the so-called "discovery learning" (tL, [15]). Through [16]-[18] he demonstrates (tT, tL, tM) how unrealistic it is to expect that the puzzle will "naturally" lead students to thinking about the Pythagorean theorem, in all these scenarios, not only in the worst case scenario, based on a reflection on his own mathematization of the puzzle. He shows that even if students are highly theoretically minded, the puzzle cannot bring them to thinking about the Pythagorean theorem, if they hadn't seen it before, because the theorem points to the conditions of existence of a puzzle such as the given one. This existence is not put into question in the puzzle; the puzzle is a fact.

The educator eventually goes back to thinking about the possible reactions of his student teachers to the realization of the epistemological impossibility of obtaining the Pythagorean theorem through even theoretical modeling of the puzzle. Based, again, on his methodology of worst-case scenario, he prepares to counter the probable student teachers' conclusion that "manipulatives are bad" with a proposal of an alternative representation (tT, pM, tM, [19]).

Figure 10 contains a summary of the above analysis of the educator's engagement with the different domains of practice.

A striking overall characteristic of the teacher educator's thinking is the lack of one coherent theoretical framework or conceptual system, on which his planning would be based. The educator makes decisions based on bits of various "theories," while being strongly influenced by his own experience and practice of teaching, learning and doing mathematics. His relationship to the different objects of his reflection can be regarded as locally, but not globally theoretical. He makes conscious use of a methodology, but does not reflect on its validity. He does not verify for contradictions among his conclusions drawn at different points in his planning. His aim is to produce a rich learning experience for his student teachers, not to construct a theory of the use of manipulatives in the teaching of mathematics.

The next section contains a theoretical reflection of the researcher on the results of this thought experiment and, more generally, on research in mathematics education (tR).

7. CONCLUSION

It was not too difficult to write a characterization of theoretical as opposed to practical thinking. Innumerable philosophers did that, at least from the time of Aristotle. It was much harder to use this distinction in speaking about a concrete instance of thinking about teaching, learning and doing mathematics. One reason for this difficulty is the complementarity of the categories of practical and theoretical thinking. Both are related to action, one engaged with action from within, the other – from without. This is a subtle difference and it is easy for the researcher to mistake one for the other.

At any given moment, the thinking subject is involved in a practical relationship with an action, planning what to do next. But any decision that is being made in the course of this action depends on a consideration, however swift, of the hypothetical possibilities and the choice of one. The choice may be based on various degrees of theoretical analysis and construction. It is not possible to reliably judge such momentary choices as based or not on theoretical thinking – and this is another source of the difficulty. One can only speak of the presence of perhaps some features of this kind of thinking and one can never be sure if this short instance of thinking was done with some global and conscious intention of theory construction. "Intention" and especially "conscious intention" are categories that have caused enough problems in philosophy and psychology; it is very difficult to operationalize them in research.

Another reason of the difficulty is the complexity of what goes on in people's minds. Thinking takes place simultaneously at several planes of action, which can be considered separately only in theory, and even then, hypothesizing about the thinking at all of these planes in a subject at a given moment of an observation or interview may easily overwhelm even the most assiduous of researchers. This complexity cannot be ignored in mathematics education research, because its object is exactly the interplay of thinking at several levels of action at once. The construction of a coherent theoretical framework for the object of research in mathematics education is, therefore, an extremely challenging task (but not an impossible task; see, e. g. Brousseau 1996; Chevallard 1999).

It is not surprising, therefore, that so many researchers in mathematics education tend to reduce the complexity in their work, and either use eclectic approaches or focus on some chosen planes of action. Certainly cognitive and socio-cognitive issues, and philosophical questions related to the nature of mathematics have attracted much attention.

One is often tempted to deplore this state of affairs. However, as in the thought experiment described in this paper, the crucial argument in analyzing a teaching project is often found not by applying the most general and sophisticated theoretical framework, but by looking at the best-case scenario. Of course, if students are not interested or not intellectually mature for a topic, and the teacher makes pedagogical mistakes, the project will fail. But suppose students are capable and willing to think theoretically about mathematics, and the teacher is "good" according to the standards of some accepted instructional theory. If a teaching approach does not fulfill the expectations in this situation, the reason is not in the pedagogy but in the epistemology of the subject matter. Epistemological analyses of the mathematical ideas are, therefore, the foundation of any teaching project in mathematics education. This is why the work of philosophers such as Michael Otte is so important for our domain.

Concordia University, Montreal

NOTES

[1] The pronoun "he" is used throughout the paper as a generic pronoun, not as this author's political statement.
[2] Regional conference of the National Council of Teachers of Mathematics, Montreal, Canada, August 2002.
[3] These shapes were the first three produced by a 6:9 years old girl after she was given the puzzle and asked to "make some shapes with it." Asked why she made the first shape just so, she answered, "because it fitted with those triangles. And it also looks a bit like a flower, like those you get in a computer." The second shape was described as "it looks like a funny cat;" about the third she said, "it's a butterfly that's acting weird." Two grade seven students, asked to play with the puzzle, spontaneously constructed similar shapes. They were mainly interested in verifying how the parts of the pieces fitted with each other; e. g., if it was possible to make a straight line with two of them. These students were able to construct the

two squares in approximately 10 minutes. One of them believed that the inner square (in Figure 2b) is of equal size with the full square (in Figure 2a).

[4] Here, the meaning of the word "angle" may be based on an intuitive/visual idea of "corner."

[5] At this point, "square" means a quadrilateral with 4 right angles and equal sides, not a visually grasped regular shape.

REFERENCES

Brousseau, G. (1996). L'enseignant dans la théorie des situations didactiques. In R. Noirfalise, & M.-J. Perrin-Glorian (Eds.), *Actes de la Huitième Ecole d'été de Didactique des Mathématiques (Saint-Sauve d'Auvergne, 22 – 31 août 1995).* Clermont-Ferrand: IREM, 3-46.

Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques* 19.2, 221-266.

Otte, M. (1990a). Arithmetic and geometry: Some remarks on the concept of complementarity. *Studies in Philosophy and Education* 10, 37-62.

Otte, M. (1990b). Intuition and formalism in mathematical proof. *Interchange* 21.1, 59-64.

Sierpinska, A. (1996). The diachronic dimension in research on understanding in mathematics – usefulness and limitations of the concept of epistemological obstacle. In H. N. Jahnke, N. Knoche, M. Otte (Eds.), *History of Mathematics and Education: Ideas and Experiences.* Göttingen: Vandenhoek & Ruprecht, 289-318.

Sierpinska, A., Defence, A., Khatcherian, T., Saldanha, L. (1997). A propos de trois modes de raisonnement en algèbre linéaire. In J.-L. Dorier (Ed.), *L'Enseignement de l'Algèbre Linéaire en Question.* Grenoble: La Pensée Sauvage éditions, 249-268.

Sierpinska, A., Nnadozie, A. A. & Oktaç, A. (2002). *A Study of Relationships between Theoretical Thinking and High Achievement in Linear Algebra.* Concordia University: Manuscript.

Sierpinska, A., Nnadozie, A. A. (2001). Methodological problems in analyzing data from a small scale study on theoretical thinking in high achieving linear algebra students. *Proceedings of the 25th conference of the International Group for the Psychology of Mathematics Education* 4. Utrecht, 177-184.

Sierpinska, A. (2000). On some aspects of students' thinking in linear algebra. In J.-L. Dorier (Ed.), *On the Teaching of Linear Algebra.* Dortrecht: Kluwer Academic Publishers, 209-246.

Steinbring, H. (1991). The concept of chance in everyday teaching: Aspects of a social epistemology of mathematical knowledge. *Educational Studies in Mathematics* 22, 503-522.

Steinbring, H. (1993). Problem in the development of mathematical knowledge in the classroom: the case of a calculus lesson. *For the Learning of Mathematics* 13.3, 37-50.

Vygotsky, L. S. (1987). The Collected Works of L. S. Vygotsky 1. Problems of General Psychology, including the volume *Thinking and Speech.* New York and London: Plenum Press.

LUIS RADFORD

THE SEMIOTICS OF THE SCHEMA[1]

Kant, Piaget, and the Calculator

Abstract. What is the relationship between our mental activity and the empirical objects of the world? Kant raised this question in the *Critique of Pure Reason* and attempted to answer it by arguing that between the realm of concepts and that of sensuous phenomena lies the *schema*. Piaget re-elaborated the Kantian concept of schema and since then it has been extensively used in constructivist and psychological accounts of the mind. In this article, I discuss Kant's and Piaget's concept of schema from a semiotic-cultural perspective. Attention is paid to the epistemological premises on which the Kantian and Piagetian theoretical elaborations of the concept of schema were based and the role that signs played therein. I contend that the schema and its genesis can be better conceptualized if we take into account linguistic and non-linguistic mediated actions embedded in the social processes of meaning production and knowledge objectification. My discussion interweaves epistemological concerns with the semiotic analysis of a group of Grade 11 students dealing with the mathematical understanding and description of a natural phenomenon – the movement of a body along a ramp in a technological environment.

Key words: activity, cultural semiotics, gestures, Kantian and Piagetian epistemology, mediated action, phenomenology, schema.

INTRODUCTION

Kant believed – contrary to Hume, Locke and the empiricist tradition – that knowledge cannot be reduced to what impressions and senses give us. Ideas should certainly be more than the result of impressions that we receive from the contingent world. The guiding principles of experience should be more than customs if we are to avoid confining them to subjectivity. But Kant also believed – contrary to the rationalist tradition of Descartes, Leibniz and Wolff – that knowledge cannot be reduced to an inner mental activity governed by the a priori rules of Reason. Leibniz, for instance, had said that "our ideas, even those of sensible things, come from within our own soul" (Leibniz 1949, 15). If such were the case, Kant asked, how is it possible that the formal rules of Reason – removed from of all empirical content – can yield knowledge of the objects of the external world?

Kant constructed a sophisticated system that tried to accommodate both the empiricist and the rationalist traditions. In this system, the senses were no longer considered as superfluous or as with merely heuristic value, as in Leibniz[2]. Kant provided the senses with an epistemological import. In an important passage of the *Critique of Pure Reason*, he says that knowledge is constituted of both sensual perceptions and concepts (A50/ B74, 92)[3].

But knowledge is more than a cocktail of conceptual and sensual ingredients. The sensual perceptions, Kant claimed, have to be linked to their corresponding

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 137 – 152.

concepts. To distinguish between the pen on the table and the book beside it, we need to be able to differentiate among the perceptions. To accomplish this we need to judge. Otherwise, Kant said, we would be led to a "rhapsody of perceptions" (A 156/ B195, 193). Judgment is a "peculiar talent" that distinguishes whether something (a perception) goes under a certain concept or not (A133/ B172, 177). For Kant, the *schema* is precisely a function of the faculty of judgment. A schema is something mediating between the mind's logical machinery and the phenomenal world. Its task is to ensure the link between concepts and senses, that is to say, between Form and Content.

THE ENCOUNTER OF *FORM* AND *CONTENT*

The schema is a kind of *analogical procedure* – a "monogram", as Kant said – that *unveils* the link between the intellectual and the sensual in the course of its empirical execution.

Like the concepts, the schema for Kant is itself void of empirical content. Yet it must contain something which is represented in the object that is to be subsumed under the concept (A137/ B176, 180). While the schema, in one respect, must be *intellectual*, said Kant, in another, it must be *sensible* (A 138/ B 177, 181). But the schema does not have to be confounded with an image:

> If five points be set alongside one another, thus, ..., I have an image of the number five. But if, on the other hand, I think only a number in general, whether it be five or a hundred, this thought is rather the representation of a method whereby a multiplicity, for instance a thousand, may be represented in an image in conformity with a certain concept, than the image itself. For with such a number as a thousand the image can hardly be surveyed and compared with the concept. This representation of a universal procedure of imagination in providing an image for a concept, I entitle the schema of this concept. (Kant, A140/ B179, 182)

In saying that the schema is a method or universal procedure Kant meant that its execution can be repeated again and again. The schema entails, in fact, a principle of iteration linking thereby *knowledge* and *action*. Kant's epistemology supersedes here the passive receptivity of impressions of the empiricist school and the reduction of knowledge to inner mental activity effectuated by the rationalist tradition. As a result, "there is knowledge only in the schematized experience." (Chiurazzi 1990, 155). This is also what Piaget meant when he said that we know an object only when we act upon it (Piaget 1970a, 85).

Now, since the schema is not only intellectual but is also sensual, we can ask: What is the *material* of which the schema is made? In addition to the schema of number (quoted above), Kant mentioned other examples, among them the schema of a triangle and the schema of the concept of a dog. In the last two, the representation is made by drawing a figure that *during its execution reveals the method*; in the first one, the execution cannot reveal the method. There is no longer coincidence between execution and method. In the case of a number such as a thousand I can still draw point after point, except that, in this case, "the image can hardly be surveyed and compared with the concept." Judgments ("perceptual judgments", to use Peirce's term) do not work the same in geometry as in arithmetic. In the schema of arithmetic and algebraic objects highly cultural conventions underpin the very possibility of the execution of the method or universal procedure. It took Kant almost 10 years to disentangle the difference between these kinds of schemata. He came back to this difference in the third critique – *Critique of Judgment* – where, as Nichanian (1979) rightly observed, Kant met the symbol.

THE ROLE OF SYMBOLS IN KANT'S CONCEPT OF SCHEMA

It was, indeed, in the course of Kant's reflection on *Aesthetics* (by which he did not mean that which is related to art, but what in Greek is called "anaesthetic", i. e. "without sensation") that Kant encountered the symbol. How can we have or produce sensual presentations or re-representations of ideas (such as 'taste') "for which a commensurate intuition can never be given"? (Kant 1790, S57, 140)[4]. Kant wrote:

> All intuitions by which a priori concepts are given a foothold are ... either schemata or symbols. Schemata contain direct [presentations of the concept], symbols [contain] indirect presentations of the concept. Schemata effect this presentation demonstratively, symbols by the aid of an analogy (Kant 1790, S59, 148).

The schema for the geometric concepts is hence based on a certain resemblance – it shows *ostensively* a certain commonality between the concept and its sensual presentation. As in the case of ideas of 'taste' or 'beautiful', the schema of arithmetic and algebraic concepts is only symbolic. They

> express concepts without employing a direct intuition [i. e. sensual presentations – LR] for the purpose, but only drawing upon an analogy with one, i. e., transferring the reflection upon an object of intuition to quite a new concept, and one with which perhaps no intuition could ever directly correspond. (Kant 1790, S59, 148)

The analogical process that allows us to move from an object of intuition to a new concept opens a window for a new kind of reflection – a reflection that will go from analogy to analogy. In contrast to the ostensive schema that functions as an "emblem", here the symbolic schema needs to enter into a new realm, a realm of possible experience. "The symbol is the analogy of an analogy, an analogy in abeyance". (Chiurazzi 1990, 158).

With his *Critique of Judgment* Kant provided room for semiotic considerations and went beyond the borders of the *Critique of Pure Reason*. His epistemology reached a new point of development but the possibilities of development were limited by his own ontological stance[5]. To understand this point, we need to note that, in its execution or materialization, the symbolic schema produces symbols, but the symbols thus produced designate something whose mode of existence is prior to all experience. We may not know where the chain of analogies will lead us, but whatever the symbols are designating, their reference has always being *there*.

We have struck here one of the more fascinating and profound tensions in Kant's theory of knowledge. Although the symbol – as any intuition (presentation or representation) – has an *epistemological* import (as we saw in the previous section), the symbol cannot have an *ontological* constitutive role. Thus, it is unthinkable for Kant to conceive of a "pure symbolicity", i. e. a symbolicity without actual reference that, in its movement, could "participate" in the *constitution* of its own object. For Kant, the "symbol" can only be thought of in relation to a constituted reference: "the 'symbols' must always be 'symbolic' in the ... sense that [their] pure reference must be constituted in the exterior of them." (Nichanian 1979, 287) The problem is that Kant adopted the rationalists' view on concepts and that, consequently, he considered concepts as independent of, and prior to, all experience.[6] Although considering himself a good Kantian, Piaget parted from Kant exactly at this point, as we shall see in the next section.

PIAGET

In 1924, Piaget published a review of Léon Brunschvicg's *L'expérience humaine et la causalité physique* [Human experience and physical causality][7]. He was seduced by the way Brunschvicg dealt with these two concepts that were vital in Kant's theory of Knowledge. The 28-year-old Piaget rephrased Brunschvicg's position saying that experience is not, as Kant assumed, something invariable, something given once and for all. On the contrary, experience has a historical context. The object of Raison, Kant was right, is to inform experience, but, in turn, Reason is constituted *in* experience. This claim was no longer Kant's. "Experience and reason are not two terms that we can isolate: Reason regulates experience and experience adapts reason." (Piaget 1924, 587). For Piaget, an account of human reason has to give up Kantian apriorism.

To better understand Piaget's solution to the problem between experience and apriorism let us return to Kant's schema of a dog. We recognize a dog because the empirical data (intuitions) that we collect in our experience are identified and filtered by the schema. The schema is not an *abstraction* drawn from experience. Experience is possible, and the empirical data become thinkable, *because* of the schema, and not the other way around. This is why Kant's theory of knowledge does not include a theory of abstraction. What Kant needed was a theory of *subsumption*, i. e. a theory indicating how representations and perceptions are subsumed under an a priori concept. In giving up apriorism Piaget found himself in need of a theory of abstraction. Central to it was the concept of schema – a revised one. He said: "Whatever is repeatable and generalizable in an action is what I have called a schema" (Piaget 1970b, 42).

As in Kant's case, a schema for Piaget is based on iteration. But the emphasis is now on the *actions*. However, in terms of human cognition, what is important in Piaget's version of the schema is not that we can iterate actions of one kind and then actions of another kind. This would lead us to a wonderful 'panoply of schemata' (similar perhaps to Kant's "rhapsody of perceptions") that would remain in a chaotic situation in the absence of a higher organizing element. While Kant turned to the a priori concepts of the rationalist tradition, Piaget turned to structuralism:

> Any given scheme in itself does not have a logical component, but schemes can be coordinated with one another, thus implying the general coordination of actions. These coordinations form a logic of actions that are the point of departure for the logical mathematical structures. (Piaget 1970b, 42)

Piaget's anti-apriorism allowed him to conceive of symbols as playing a more decisive role in knowledge formation than they played in Kant's epistemology. Piaget's point of departure was the linking between action and representation. From the outset he insisted that it is a current mistake to reduce representation to language:

> Language is certainly not the exclusive means of representation. It is only one aspect of the very general function that Head has called the symbolic function. I prefer to use the linguists' term: the semiotic function. This function is the ability to represent something by a sign or a symbol or another object. (Piaget 1970b, 45)

In his book "*La formation du symbole chez l'enfant*" [The formation of symbol in children] – a particularly difficult book in its technical aspect because in it Piaget endeavored to show one of the central theses of his epistemology, namely that mental images are interiorized actions – Piaget argued that the symbol arises from non-symbolic schematism[8]. More specifically, Piaget was claiming that there is a *continuity* between the sensori-motor signifiers and the emergence of the first symbols in the children. In other words, that the sensori-motor intelligence prolongs itself into conceptual representation.[9]

The sensori-motor signifiers were seen by Piaget as 'indexes' or 'signals' but they still lack an independency vis-à-vis the signified object. The semiotic function begins precisely when there is a differentiation between signifiers and signifieds. This differentiation provides the signified with a spatial-temporal permanence and opens the possibility that a same signifier can be related to different signifieds.[10] For Piaget, the semiotic function includes differed imitations, symbolic play, mental images, gestures, and natural language. Following Saussure he distinguished between symbol and sign. A *symbol* is a « motivated » signifier, which means that the signifier bears a certain resemblance to the signified. A *sign*, in contrast, bears an arbitrary or non-motivated relationship to its signified. Thus, a letter that we use in an algebraic expression is a sign, while a figure standing for a triangle is a symbol.

If it is true that the constructive stance of his genetic epistemology led Piaget to pay careful attention to the way in which actions and gestures become conceptual representations, it is also true, however, that Piaget's attention to signs and symbols faded away in his analysis of older children's thinking. Reflective abstraction converts action into operations and signs come to symbolize the operations. Hence, in Piaget's epistemology, in opposition to Kant's, signs and symbol borne a constitutive ontological role, but because the primacy was given to the structure, signs and symbols were in the end merely the carriers and the expressions of a thinking measured by its structural features. Piaget wrote:

> reflective abstraction, which derives from the first concepts from the subject's actions, transforms the latter into operations, and these operations can sooner or later be carried out symbolically without any further attention being paid to the objects which were in any case "any whatever" from the start. (Beth and Piaget 1966, 237-238)

To sum up, Piaget elaborated a theoretical reformulation of the Kantian concept of schema. He emphasized the epistemological role of action and gesture. However, the emphasis on the operations' structure left little room for a thematization of the content of the operations and for a serious consideration of the semiotic systems and the cultural artifacts that the children use. Thus, for Piaget, the object that the hand holds in the schema is unimportant. It may be "any whatever" from the start, as he says in the last quotation. Verillon and Rabardel comment that

> the object submitted to the Piagetian subject is fundamentally non-historical and non-social: its main property is that it is structured by physical laws. ... The introduction of artifacts in classic Piagetian experiments is mainly due to their convenience for highlighting the invariant properties of reality. (Verillon and Rabardel 1995, 80)

Piaget's recourse to structuralism (even if it was a dynamic one) introduced irresoluble tensions in his epistemology – tensions that were proportional, we may say, to the ones Kant introduced in his by having recourse to apriorism.[11] While in Kant the tension appears between Form and Content, between concept and sensual representation, in Piaget it appears as the tension between structure and object. In both epistemologies, nevertheless, the common denominator is that mind's activity is, in the end, reduced to abstract mental labour.[12]

In the next section I will claim that, from an epistemological and a psychological viewpoint, the concept of schema needs to be broadened so as to include not only the instruments that the individual uses (which has been Rabardel's recent claim[13]) but its cultural context and other semiotic means such as speech and gestures that, more than mere ephemeral descriptors of reality, prove to be fundamental in knowledge formation.

SCHEMA AND ACTIVITY

Let us come back to Kant's concept of schema. As previously seen, for Kant, the distinctive epistemological trait of a schema is to present or exhibit, through the execution of a procedure, the "intuition" of an object (the object of knowledge). I will take this idea as my starting point. However, as Peirce contended (Peirce 1966, 43), the way in which the object thus becomes intuited has a *volitional character* that Kant did not take into account. The volitional character underpinning the schema and its genesis, should be studied in the context of the individuals' activity.[14] As such, it is related to the activity's *goal*. But complex activities are often comprised of chains of actions. A chain is directed towards the attainment of an *aim*. An aim (in contrast to the goal of the activity) is not necessarily something that is set from the beginning: it is a reference point that hypothetically can lead us closer to the goal. The formation of an aim is part of the heuristic process underlying the activity.[15] Bearing these remarks in mind, the schema, I would like to suggest, is an organization of actions or a chain of actions related to the attainment of the goal and aims of an activity.

In this perspective the schema has a double nature: (1) a *functional* and (2) a *phenomenological* one.

(1) The phenomenological aspect of the schema:

In its *phenomenological aspect*, the schema is a mode of presentation – a mode of "exhibition" of the object, as Kant used to say, – an effort to render something (e. g. a conceptual object or a process) available, noticeable – even if, ontologically speaking, the object or the process (in short, what Husserl called *objectity*[16]) does not have necessarily precedence over the action. In this case, the schema *produces* the object and functions as a form of disclosure (in Heidegger's sense[17]). The schema *objectifies* the object (Radford 2003a).

(2) The functional aspect of the schema:

The *functional aspect* of the schema means that the schema is governed neither by the Kantian rationalist apriorism nor by the Piagetian's normative character of logico-mathematical structures. Indeed, its adequacy is not examined against a grid of truth but against its practical results.[18]

The schema, as I am formulating it, is still both a sensual and an intellectual action or a complex of actions. In its intellectual dimension it is embedded in the theoretical categories of the culture. In its sensual dimension, it is executed or carried out in accordance to the technology of *semiotic activity* (Radford 2002b). We still save some of the characteristics of the Kantian formulation – figurative synthesis in the heuristic process, the difference between the execution of the schema and its result, its reiteration – but I place it in the broader context of the individual's subjective awareness that, in its constructive and creative endeavor, grows sustained and framed by the theoretical categories of the culture, its technology of semiotic activity and the historically constituted mode of knowing (Radford 2003b). In the next section, I turn to a classroom episode that will help clarify the previous ideas.

THE TECHNOLOGICALLY MEDIATED SCHEMA: FILLING THE HOLES

In an artifact-mediated classroom activity, Grade 11 students were asked to investigate the relationship between time and distance of a cylinder moving up and down an inclined plane[19]. In one of the parts of the activity the students performed two experiments using a TI 83+ calculator and a Calculator Based Ranger (CBR) motion detector. In the first one, the students propelled the cylinder upwards, from the bottom of the inclined plane and activated the CBR as soon as the cylinder was put in motion. In the second one, the cylinder was propelled one second after the CBR was activated[20]. In both experiments the CBR was placed at the top of the inclined plane (see Figure 1).

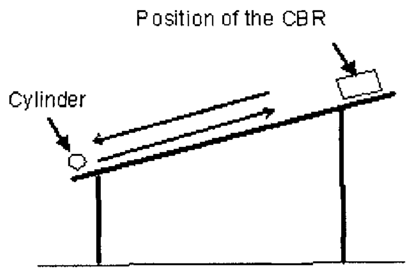


Figure 1. Inclined Plane or Table.

Figure 2 shows the calculator graph for the second experiment. In this part of the activity the students had to explain the shape of the calculator's graph. Another part of the activity consisted in two 'thought experiments.' Here the students were asked to sketch two graphs: one for a cylinder moving on an imagined ramp that had a greater slope than the one of the experiment, and one for a cylinder moving on an imagined ramp having a lesser slope than the one of the experiment.

I will discuss first the students schema that resulted from the delayed motion (t = 1; see Figure 2), and then I will comment on the use of this schema in the part concerning the 'thought experiments.'

In the first part, the schema, whose result is the graph shown on the calculator screen, consists of a sequence of actions, among them: (1) preparing the technological system calculator-CBR; (2) activating the CBR; (3) propelling the cylinder; (4) following the cylinder perceptually during its trajectory; (5) stopping the CBR when the cylinder comes back down and (6) making sense of the graph.

In order to better understand the schema we need to discuss the role of the technological system 'calculator-CBR' which was crucial in the experiment and in the generation of the graph. For sure, this technological system (TS) permitted a substantial economy in the carrying out of the experience. While Galileo went to great pains to figure out a way to measure the consumed time (a variable that, in contrast to distance, cannot be *seen*), the TS registered the measures of distance and time and, in the human-TS interaction, the calculator *produced* the graph.[21]

Now, the TS is more than a gadget to economize actions. It carries in itself, in a compressed way, socio-historical experiences of cognitive activity and scientific standards of investigation (Lektorsky 1984; Pea 1993). In addition to providing the students with economy and precision, the TS executes some of the human actions that it holds in a compressed way, and displays on its screen outputs of these actions. However, by taking over some of the human actions, certain aspects of the socio-historical experience that the TS holds remain "hidden" from the individuals using it. As a result, the schema loses an important aspect of the "sensuality" that it could have had for Galileo and the understanding that could have resulted from seeing, touching, and doing. The fact that the symbol-graph is not the result of the individuals' own actions but rather the result of the individuals' actions *and* those socio-historical ones that the TS executes, brings forward a very important element in the genesis of the schema: the resulting schema is a schema containing "gaps" or "holes". Indeed, while the *execution* of the symbol-figure of a triangle *reveals* the schema, in the technological experiment the *displaying* of the symbol-graph of the cylinder's motion on the calculator screen does not. There is no longer coincidence nor analogy between the execution of the procedure and the schema.

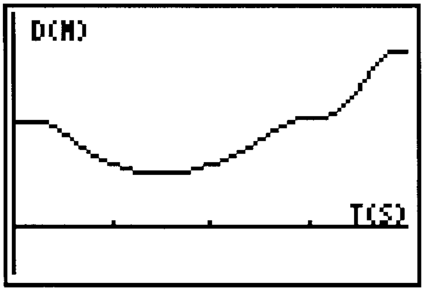


Figure 2. Calculator's graph.

To obtain the schema, the holes have to be filled. However, the problem is not to repair the holes induced by the division of labor with their original substance (which would be impossible anyway). The problem is to *make sense* of the symbol-graph thus produced. In contrast to the schema of the triangle or of the dog, and the schemata discussed by Piaget, such as the baby hitting an object with a stick, the semiotic activity does not end with the production of the "image" (Kant) of the schema. The semiotic activity goes beyond the image (here the symbol-graph). The question is not primarily to judge but to interpret and to produce meaning. The students have to *make sense* of the image and to do so they will have recourse to other semiotic systems such as gesture and speech, as we will now see.

The students noticed that the graph was not perfectly curved in the part after its minimum value and that, in the graph, the value of the variable D (distance) in the ending points is not the same (i. e. $D_f > D_0$; see Figure 2). While the first difference was explained by a slight turn of the cylinder when it was rolling upwards on the inclined plane, the second difference was more difficult to understand. After discussing different ideas Judith said:

Judith: ... (looking at the inclined plane) This thing there [the cylinder], does it go further? (the other two girls turn to see the inclined plane which was behind the students' desks) ... like this ... (she makes a gesture with her right arm; the gesture starts with her arm extended in front of her body and moves back, miming the cylinder motion in its coming back down trajectory) does it measure the ...? Oh! (she thinks she understood something)

Vanessa: What?

Judith: You started on the table [i. e. the table that served as the inclined plane for the experiment], right? (Vanessa : Yes) And when it was rolling it fell off the table (with a similar gesture her arm is bent again and goes beyond her desk, as the falling cylinder did during the final part of its motion when it fell off the inclined plane and was caught by the student)... I don't know...

Vanessa: It has nothing to do with that.

Judith: It does have something to do with that [...] That's the curve, right? Here (she points to the horizontal segment of the left part of the graph on the calculator screen) suppose this is when you started on the table and when you finished (she points now to the horizontal segment of the right part of the graph), you've finished further, that's further. [...] Let's say that your distance here would be 30, and 45, that's the error! [...]

In Lines 1 and 3 Judith makes an "iconic gesture", that is, a gesture that bears a resemblance with its referent. The iconic sign-gesture *enacts* the falling trajectory of the cylinder (see Figure 3). It allows Judith to call her group mates' attention to a specific part of the phenomenon. Like the Calculator-CBR system, the iconic gesture affords a segmentation of the phenomenon and operates a choice of what has to be taken into account. But in contrast to the Calculator-CBR system, the iconic gesture does not stress speed, time, accurate distance and other elements. What it stresses is the fact that the cylinder went off the table. The iconic gesture has made an important fact *evident* (i. e. capable of being seen). The fact that, in its way back down, the cylinder went off the table and, consequently it travelled more distance, allows Judith a new interpretation of the graph. The new interpretation is elaborated on Line 5. Indeed, in Line 5, Judith has recourse to an "indexical gesture": pointing with her finger, she indicates two parts of the calculator graph on the screen (see Figure 4). In this case, numbers (30 cm and 45 cm) come to play the role of the iconic gesture that has previously shown the cylinder falling off the table. The first number represents the students' estimated distance from the cylinder's maximum point to the bottom of the table. However, the cylinder never went 15 cm off the table (i. e. 45-30), for it was caught immediately as it fell off the table. By exaggerating the numbers, the particular element of the phenomenon is highlighted.

Figure 3. Judith makes an iconic gesture that mimes the cylinder coming back down. In the genetic constitution of the schema, the students have to interpret the results of the artefact-mediated actions. To do so, they have recourse to gestures and speech.

I have discussed in some detail the previous students' dialogue because this dialogue shows aspects of the students' efforts to fill the schema and, overall, because I take these efforts as an important part of the genesis of the schema.

Figure 4. Indexical Gesture.The students point to the right part of the calculator screen.

The students' dialogue suggests that to fill the holes in the schema the students produce a kind of simulation of the cylinder motion. The simulation was oriented towards understanding some 'remarkable points' on the graph. In the terminology of the previous section, these points are examples of *aims* and represent, as Arzarello and Robutti (2001, 37) indicate, strong connections between signs and experience. To attain the aims, the students had recourse to language. Through its rich arsenal of terms, in particular through some *objectifying deictics* (e. g. pronouns, locative words, time-related expressions), language allowed the students to "indexicate" and "iconize" essential features of their mathematical experience[22]. The students' dialogue also shows how language was coordinated with gestures in the production of meaning and understanding.

Once some understanding was reached and that the schema was apparently completed, the students could apply the schema to the proposed "thought experiments". To do so, the schema was significantly contracted. The technology of semiotic activity was not the same (now the students worked with pencil and paper). The key element that the students retained of the cylinder motion was the parabolic shape and the starting and ending points of the graph. They then produced the graphs shown in Figure 5.

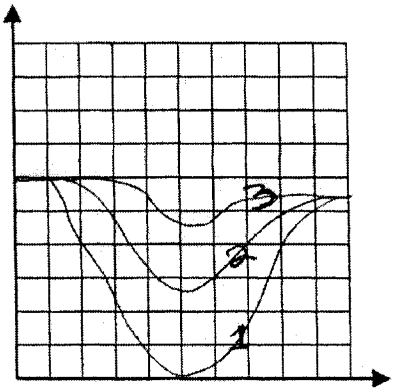


Figure 5. Graph for the cylinder motion on an inclined plane having a greater slope (graph 3) and a lesser slope (graph 1) than the original inclined plane (graph 2).

Graph 2 corresponds to the ramp of the original experiment. Graph 1 corresponds to the ramp having a greater slope and graph 3 to the ramp having a lesser slope. Of course, the results are not mathematically correct. The students focused on the kind of "effort" that it takes the cylinder to go up when the slope is greater and when the slope is lesser than the original one. All in all, the graphs show a partial understanding of the abstract mathematical spatio-temporal relationship of the cylinder motion.

SYNTHESIS AND CONCLUDING REMARKS

As we saw, Rationalists conceived of the mind as governed by a kind of abstract logical calculus ensuring deductions such as "$M \geq N$ and $N \geq P$, then $M \geq P$", regardless of the content of M, N, and P. Formal deduction removed from all empirical content, however, Kant argued, cannot yield knowledge. The question then was to explain how abstract concepts relate to their concrete content. In an important sense, the *Critique of Pure Reason* is an attempt to achieve this goal and the schema, in fact, was Kant's answer.

One of the distinctive theoretical features of Kant's concept of schema is that the individual is neither reduced to a passive receiver of impressions neither to a flesh box in whose interior logical calculations are effectuated. The schema entails the idea of an individual who, to acquire knowledge, has to become active. However, in Kant's theory of knowledge, the schema exhibits or unveils its concept – it does not produce it. Piaget retained the Kantian feature of an active individual, gave up apriorism and added a new ontological dimension: the schema was endowed with the power of producing concepts. Piaget's thesis, in fact, was stronger: concepts could not be produced in the absence of their correlated schemata[23]. The Piagetian elaboration of the schema opened a window for semiotic considerations. However, the concrete was rapidly evacuated and the relationship between Content and Form ended up being thematized against the rigid grid of logico-mathematical structures.

Since most of our actions are carried out using signs and artifacts, and since these actions are not performed in an arbitrary way but are framed by social goals and the cultural logic of meaning, I suggested that the schema can be reinterpreted as an organization of semiotical and artifactual actions or a chain of such actions related to the attainment of the goal and aims of an activity. The mediated nature of actions, nevertheless, leads to an important and difficult problem. Mediation means that, to accomplish something, we have recourse to an item of our environment (e. g. a word, an idea, a tool) that has already a social meaning. Carrying out a mediated action thus requires a lot of understanding. As a result of this intrinsic social nature of mediated action, the schema, generally speaking, cannot "exhibit" or show ostensively its object during its execution. The example of the graphic calculator and the motion sensor, I think, showed this point in a clear way. In the classroom episode, the students' schema was framed by a complex division of labor. The technological system calculator-CBR performed some key actions; as a result, even if the material product of the schema (i. e. the calculator graph) could be *seen*, the schema had "holes" that the students had to fill using creative imagination. The parabolic shape of the graph shown by the calculator underwent a process of interpretation. To do so, the sensual content of the cylinder motion had to be related to abstract aspects of the graph. Surely, language is a powerful means of objectification. However, in the genesis of knowledge, the relationship between conceptual descriptions and their referents cannot be reduced to linguistic terms (Otte 1998, 444). How then to account for the emerging schema and its encompassing description of the relationship between the concrete and the abstract? The interpretative process of the calculator graph (a crucial step in the formation of the schema), may shed some light on this problem.

In the course of this interpretative process, we saw the students displaying a range of semiotic forms of meaning production and knowledge objectification such as iconic and indexical reference (Figures 3 and 4) that were intermingled with language, intimating that the subsumption of a sensual content *A* into an abstract concept *B* by the schema may be much more complicated than perhaps Kant himself imagined. For one thing, both indexical and iconic reference involve types of "predication" different from those of the form "subject-copula-predicate", that is, of the form '*A is B*' that Kant emphasized following the classical logic's view on judgments. It may very well be that 'Reality' is much less homogeneous than what we usually think and that the distinction between the concrete and the abstract might be placed on an "infinite graduation of being, of perspective and of communication" (Otte 1998, 425) that language alone fails to capture.

If "the essential question of epistemology", as Otte suggests, is to understand that which "enables an *A* to stand for a *B*" (Otte 1998, 429) or that which makes an *A* to become subsumed into a *B*, a broader concept of predicative copula and relationship between *A* and *B* (between the concrete and the abstract or between Content and Form) would be required. In this line of thought, copular predication, I want to suggest, needs to be broadened so as to include other forms of semiotic reference capable of accounting for the dialectic ways of the constitution of subject and predicate, that is, of the semiotic processes through which the object of knowledge becomes noticed and socially thematized (in short, schematized) within a certain cultural discourse. It requires us paying attention to the technology of semiotic activity and its interaction with other semiotic systems in what Lotman (1990) calls the "semiotic space".

In the section titled Schema and Activity, I mentioned that I place the schema in the context of the individual's subjective awareness that grows sustained and framed by a historically constituted mode of knowing. I want to conclude these remarks by mentioning in what sense a schema relates to its cultural mode of knowing. I cannot find a better way to do so than to recall a phrase that Peirce wrote in a projected book that he never finished. Summarizing Kant's ideas, Peirce wrote: "Every cognition contains a sensual element."[24] In fact, every cognition (i. e. every phenomenon of our mental life) contains much more than a sensual element. It contains its cultural way of knowing. Thus, in the classroom activity, in addition to implicitly asserting, in a subtle way, the *existence* of a mathematical relationship between time and space that describes the cylinder motion, the design of the activity informs the students that such a relation becomes intelligible through *experimentation*. The question we asked the students and their conceptual procedures to answer it are framed and thus make sense within a particular, historically constituted mode of knowing. Had we asked a 17[th] century philosopher of nature this same question he would have certainly found it amusing – if not laughable. Vicenzo di Grazia (an Aristotelian philosopher and contemporary of Galileo), for instance, said:

> ... those who want to demonstrate natural accidents through mathematical methods are delirious... the natural philosopher [scientifico naturale] studies natural phenomena whose essence entails movement, while, instead, the subject matter of mathematics does not comprehend movement. (Quoted in Biagioli 1993, 205).

When I said, in the application of the schema to the two thought experiments, that the student *only* retained the parabolic shape and the starting and ending points of the graph, I was forgetting the most important thing: the students' schema embodies a way of inquiring and of knowing about nature that only habit makes us now take for granted and to see as "natural".

Université Laurentienne, Ontario.

NOTES

[1] This article is a result of a research program funded by the Social Sciences and Humanities Research Council of Canada.

[2] In *New Essays Concerning Human Understanding*, Leibniz says: "necessary truths ... must have principles whose proof does not depend upon examples, nor consequently upon the testimony of the senses, although without the senses it would never have occurred to us to think of them. This distinction must be carefully made, and was so well understood by Euclid, that he often proved by the reason, what is sufficiently seen through experience and by sensible images." (Leibniz 1949, 44)

[3] As usual in references to Kant's *Critique of Pure Reason*, A50 means page 50 of the 1781 edition; B74 means page 74 of the 1787 edition, etc. Page 92 refers here to the English translation of Norman Kemp Smith. I will use this format throughout this article.

[4] In Kant's vocabulary "intuition" means an effected immediate relation that objects have on us (see A19/B33, 65). Examples of "intuitions" are impressions, perceptions, representations, etc.
[5] The problem, of course, is not that Kant had an ontology. We all need a theory of Being (even if it is only an implicit theory) in order to make assumptions or hypotheses. As Adorno pointed out, "If you refuse to make any assumptions, if you attempt to understand a thing purely on its own terms, then you will understand nothing." (Adorno 2001, 13).
[6] Daval (1957) deals with this topic in detail.
[7] Brunschvicg 1922.
[8] In the beginning of the book he says: "We will attempt to show how the [emergence of the] symbol is prepared by the non-symbolic schematism" (schématisme pré-représentatif). (Piaget 1968, 8).
[9] Piaget 1968, 68-69. See also Piaget 1972.
[10] Piaget in Piattelli-Palmarini 1982, 58.
[11] One of the tensions in Piaget's epistemology is its problematic concept of necessity, related to the growth of knowledge. It has been discussed in Otte (1998, in press). Another one is related to the problem of objectivity. It has been discussed in Radford 2002a.
[12] For a detailed elaboration of this point see Adorno 2001 and Buck-Morss 1975.
[13] Rabardel 1995, 1997.
[14] I use the term activity here in Leontiev's sense (Leontiev 1984).
[15] Leontiev 1984, 117.
[16] Husserl 1961, 44.
[17] Heidegger 1971.
[18] In his interesting work, Vergnaud (1985) was also confronted with the problem of the adequacy of the schema. In dealing with this problem in terms of *invariants*, he certainly succeeded in avoiding the Piagetian normative problem of logical structures. Among the invariants, Vergnaud included propositions (i. e. something that is true or false) and "propositional functions" – abstract functions having propositions as "variables". However, since "truth" as a conceptual category is adopted without critical stance, it is not clear how, epistemologically speaking, invariants are dependent and sensitive to the concrete cultural contexts of learning.
[19] The episode is described in detail in Radford *et al.* (2003).
[20] Thus, in the first experiment, the cylinder motion started at t=0 and, in the second experiment, motion started at t = 1 sec.
[21] Commenting on the data collection in his experiment on an inclined plane, Galileo says: "As to the measure of time, we had a large pail filled with water and fastened from above, which had a slender tube affixed to its bottom through which a narrow thread of water ran; this was received in a little beaker during the entire time that the ball descended along the channel [carved on the inclined plane] or parts of it. The little amounts of water collected in this way were weighed from time to time on a delicate balance, the differences and ratios of the weights giving us the differences and ratios of the times, and which such precision that, as I have said, these operations repeated time and again never differed by any notable amount." (Galileo 1638, 170)
[22] A detail elaboration of the idea of "objectifying deictics" can be found in Radford 2002b.
[23] I will not dwell into this point here, limiting myself to mention that, to some extent, Radical Constructivism was elaborated as an effort to bring this point to its logical conclusions (for a critique see e. g. Lerman 1996; for a reply see Steffe and Thomson 2000).
[24] Peirce in Hooper (Ed.) 1991, 17.

REFERENCES

Adorno, T. W. (2001). *Kant's Critique of Pure Reason.* Stanford CA: Stanford University Press.

Arzarello, F., & Robutti, O. (2001). From Body Motion to Algebra through Graphing. In H. Chick, K. Stacey, J. Vincent, & J. Vincent (Eds.), *Proceedings of the 12th ICMI Study Conference*, The University of Melbourne, Australia, Vol. 1, 33-40.

Beth, E., W., & Piaget, J. (1966). *Mathematical Epistemology and Psychology.* The Netherlands: Reidel.

Biagioli, M. (1993). *Galileo, Courtier.* Chicago: Chicago University Press.

Brunschvicg, L. (1922). *L'expérience humaine et la causalité physique.* Paris: Alcan.

Buck-Morss, S. (1975). Socio-Economic Bias in Piaget's Theory and Its Implications for Cross-Culture Studies. *Human Development* 18, 35-49.

Chiurazzi, G. (1990). Schématisme et modalité: La doctrine kantienne du schématisme comme thématisation analogico-expérimentale de la connaissance. *Kant Studien* 91.2, 146-164.

Daval, R. (1957). *La métaphysique de Kant.* Paris: Presses Universitaires de France.

Galileo, G. (1638). *Two New Sciences* (translated by S. Drake, 1989). Toronto: Wall & Thomson.

Heidegger, M. (1971). *Poetry, Language, Thought* (translated by A. Hofstadter). New York: Harper & Row.

Hooper, J. (Ed.) (1991). *Peirce on signs.* Chapel Hill and London: The University of North Carolina Press.

Husserl, E. (1961). *Recherches Logiques* (Recherches I et II). Paris: Presses Universitaires de France.

Kant, I. (1781, 1787/ 1929). *Critique of Pure Reason* (translated by Norman Kemp Smith). New York: St. Marin's Press (second printing, 1965).

Kant, I. (1790). *The Critique of Judgement* (translated by James Creed Meredith), retrieved from: http://etext.library.adelaide.edu.au/aut/

Leibniz, G. W. (1949). *New Essays concerning Human Understanding* (translated from the 1705 text by A. G. Langley). La Salle, Ill: The open Court.

Lektorsky, V. A. (1984). *Subject, Object, Cognition.* Moscow: Progress Publisher.

Leontiev, A. N. (1984). *Activité, conscience, personnalité.* Moscou: Éditions du Progrès.

Lerman, S. (1996). Intersubjectivity in Mathematics Learning: A Challenge to the Radical Constructivist Paradigm? *Journal for Research in Mathematics Education* 27.2, 133-150.

Lotman, Y. (1990). *Universe of the Mind. A Semiotic Theory of Culture.* London/New York: I. B. Taurus.

Nichanian, M. (1979). *La question générale du fondement: Écriture et temporalité.* Thèse de Doctorat de troisième cycle de Philosophie. Strasbourg: Université des Lettres et sciences humaines de Strasbourg.

Otte, M. (1998). Limits of constructivism: Kant, Piaget and Peirce. *Science & Education* 7, 425-450.

Otte, M. (in press). Does mathematics have objects? In what sense? *Synthese.*

Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed Cognitions.* Cambrigde: Cambrigde University Press, 47-87.

Peirce, C. S. (1966). *Collected Papers,* (Ed. by A. Burks), Vol. 8. Cambridge: The Belknap Press of Harvard University Press.

Piaget, J. (1924). L'expérience humaine et la causalité physique. *Journal de psychologie normal et pathologique* 21, 586-607.

Piaget, J. (1968). *La formation du symbole chez l'enfant.* Neuchatel: Delachaux et Niestlé.

Piaget, J. (1970a). *Psychologie et épistémologie.* Paris: Éditions Gonthier.

Piaget, J. (1970b). *Genetic Epistemology.* New York: W. W. Norton.

Piaget, J. (1972). *Problèmes de psychologie génétique.* Paris: Denoël/Gonthier.

Piattelli-Palmarini, M. (Ed.) (1982). *Théories du langage, théories de l'apprentissage: le débat entre Jean Piaget et Noam Chomsky.* Paris: Seuil.

Rabardel, P. (1995). *Les hommes et les technologies.* Paris: Armand Colin.

Rabardel, P. (1997). Activités avec instruments et dynamique cognitive du sujet. In C. Moro, B. Schneuwly, & M. Brossard (Eds.), *Outils et signes. Perspectives actuelles de la théorie de Vygotski.* Bern: Peter Lang, 35-49.

Radford, L. (2002a). The Object of Representations: Between Wisdom and Certainty. In F. Hitt (Ed.), *Representations and Mathematics Visualization,* Departamento de matemática educativa Cinvestav-IPN, Mexico, 219-240.

Radford, L. (2002b). The seen, the spoken and the written. A semiotic approach to the problem of objectification of mathematical knowledge. *For the Learning of Mathematics* 22.2, 14-23.

Radford, L. (2003a). Gestures, speech and the sprouting of signs. *Mathematical Thinking and Learning* 51, 37-70.

Radford, L. (2003b). On Culture and Mind. A post-Vygotskian Semiotic Perspective, with an Example from Greek Mathematical Thought. In M. Anderson, A. Sáenz-Ludlow, S. Zellweger, & V. V. Cifarelli (Eds.), *Educational Perspectives on Mathematics as Semiosis: From Thinking to Interpreting to Knowing.* Ottawa: Legas Publishing, 49-79.

Radford, L., Demers, S., Guzmán, J., & Cerulli, M. (2003). Calculators, graphs, gestures, and the production meaning. In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27 Conference of* *the international group for the psychology of mathematics education* (PME27 – PMENA25), Vol. 4, 55-62.

Steffe, L. P., & Thompson, P. W. (2000). Interaction or Intersubjectivity? A Reply to Lerman. *Journal for Research in Mathematics Education* 31.2, 191-209.

Vergnaud, G. (1985). Concepts et schèmes dans la théorie opératoire de la représentation. *Psychologie française* 30.3-4, 245-252.

Verillon, P. and Rabardel, P. (1995). Cognition and Artifacts: A Contribution to the Study of Thought in Relation to Instrumented Activity. *European Journal of Psychology of Education* 10, 77-101.

KENNETH RUTHVEN

TOWARDS A NORMAL SCIENCE OF MATHEMATICS EDUCATION?

Abstract: This paper suggests that the first BaCoMET [Basic Components of Mathematics Education for Teachers] project (Christiansen, Howson and Otte 1986) can be seen as an important early attempt to sketch a 'disciplinary matrix' (Kuhn 1962, 1970) for the field of mathematics education. This project brought together representatives of different national traditions of research in mathematics education, with the aim of identifying fundamental ideas which should be given high priority in any teacher education programme. My analysis of the project draws on Kuhn's (1970) prioritisation of different senses of 'paradigm' in relation to the development of 'normal science', and consequently draws out the central part played by clusters of exemplary problems (and solutions) in mediating between symbolic generalisations and practical action.

Key words: didactical research; mathematics education; normal science and scientific paradigms; teacher education; theory-practice mediation.

THE IDEA OF BASIC COMPONENTS OF MATHEMATICS EDUCATION FOR TEACHERS

My first – indirect – contact with Michael Otte was as a newly appointed teacher educator, through reading his report on 'The education and professional life of mathematics teachers' (Otte 1979) in a volume prepared by the ICMI for a UNESCO series on New Trends in Mathematics Teaching (Christiansen and Steiner 1979). From the start – and entirely characteristically as I now recognise – Otte cautioned that "such a report can only ... provide cues for orientation at a general level by furnishing a conceptual framework for the contextual analysis of problems" (108). For Otte, "the special structural problem of the teaching profession" was "that it does not have a basic science such as law for the lawyer, medicine for the physician" (114-115), with the result that "the most central problems for teacher education are undoubtedly those of mediating between theory and practice, [and] between the subject matter and the social and educational sciences" (126). In particular, he argued that "the present situation of the pedagogy of mathematics is characterized by great conceptual deficiencies" at a time when "we are faced with a host of practical problems which can no longer be handled by the conventional inventory of spontaneous principles gained from experience and transmitted by tradition" (127).

Together with two fellow contributors to the UNESCO volume – Bent Christiansen and Geoffrey Howson – Otte set out to remedy this situation by initiating a project aimed at establishing what they termed 'Basic Components of Mathematics Education for Teachers'. The rationale for this project is briefly presented in

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 153 – 158.

the introduction to the resulting 'Perspectives on Mathematics Education' (Christiansen, Howson and Otte 1986). The animating idea was that "there were certain basic, fundamental ... components of the didactics of mathematics which should be given high priority in any teacher education programme" (ix). A 'basic component' was taken as being an aspect of mathematics education which is:

> (i) fundamental in the sense that it plays a decisive part in the functioning of mathematics teachers; (ii) elementary in the sense that it is accessible to intending teachers of mathematics and of immediate interest and value to them; and (iii) exemplary in the sense that it exemplifies important didactical or practical functions of the teacher and their inter-relationships. (x)

Thus the project aimed to produce a text "to convey... such knowledge as would be useful to the teacher in carrying out his functions, i. e. *knowledge for action*" (xi). Nevertheless, in conceptualising the form that such 'knowledge for action' might take, the project group was exercised by "the relationships between, on the one hand, *theoretical knowledge* (scientific theories about subject matter and about didactical concerns), and, on the other, the *know-how* of the practitioner (i. e. the experienced teacher) who is operating and acting in appropriate ways in the classroom" (xi). Moreover, the multi-national project group faced a further challenge in synthesising and refining the diverse knowledge – both theoretical and practical – which individual members brought from different traditions of research, so as to build "knowledge which was in some form 'common' or 'shared' by the group" (xi).

PARADIGM AS METAPHYSICAL FRAME, DISCIPLINARY MATRIX OR EXEMPLARY PROBLEM

This first BaCoMET project was an important attempt to frame a shared perspective on the field across different traditions. A more recent ICMI study, entitled 'What is Research in Mathematics Education and What are Its Results?', has sought to explore – indeed, attempted to resolve – such differences of perspective within the field. Its main conclusion has been summarised by its leaders in the following terms: "[I]n spite of all the differences that divide mathematics education researchers (in terms of theoretical approaches, views on relations between theory and practice, philosophies of mathematics, etc.), they still constitute a community, and it is necessary to search for what constitutes its identity" (Sierpinska and Kilpatrick 1998, xi). The differences manifested in the course of the study were often characterised in terms of alternative – even conflicting – 'paradigms' for research in mathematics education. This issue is discussed most fully by Ernest (1998) who, following Habermas, talks of "multiple research paradigms, each with its own assumptions about knowledge and learning (epistemology), about the world and existence (ontology), and about how knowledge is obtained (methodology)" (77). While this conception has been widely influential in methodological discussions in the social sciences, in my view, it represents an overly rationalistic and foundationalist approach.

I see the alternative perspective offered by Kuhn (1962, 1970) as a more fruitful one for appraising the situation of mathematics education. In the second edition of his work on 'The Structure of Scientific Revolutions', Kuhn (1970) added a substantial postscript in which he sought to tighten his loose central construct of 'paradigm'. In particular, he drew attention to two related – but ultimately distinct – senses of the term; the first more popular, the second more profound:

> On the one hand, ['paradigm'] stands for the entire constellation of beliefs, values, techniques, and so on shared by members of a given community. On the other, it denotes one sort of element in that constellation, the concrete puzzle-solutions which, employed as models or examples, can replace explicit rules as a basis for the solution of the remaining puzzles of normal science ... Philosophically, at least, this second sense of 'paradigm' is the deeper of the two. (175)

It is these shared models or examples which mediate between codified theory, expressed in what Kuhn terms 'symbolic generalisations', and practical tasks of framing and solving problems. This identification of the central part played in scientific thinking by a myriad of problem-solutions points to a much smaller granularity of scientific knowledge than do accounts which conceive such knowledge more exclusively in terms of symbolic generalisations. Kuhn's approach dissolves the idealised epistemological model of theory application, and highlights the crucial part that exemplars play in mediating theoretical constructs.

Accordingly, Kuhn draws attention to the central function of exemplary problems in the formation of disciplinary knowledge:

> The paradigm as shared example is the central element of what I now take to be the most novel and least understood aspect of this book ... Philosophers of science have not ordinarily discussed the [exemplary] problems encountered by a student ... for these are thought to supply only practice in the application of what the student already knows. He cannot, it is said, solve problems at all unless he has first learned the theory and some rules for applying it. Scientific knowledge is embedded in the theory and rules; problems are supplied to gain facility in their application. ... [However] this localization of the cognitive content of science is wrong. ... In the absence of such exemplars, the laws and theories [the student] has previously learned would have little empirical content. (187-8)

Working with exemplars not only gives substance to laws and theories but establishes a fine texture of largely tacit knowledge through which a wide range of situations can ultimately be related to a single idea:

> [The] ability to see a variety of situations as like each other, as subjects for ... [some] symbolic generalization, is ... the main thing a student acquires by doing exemplary problems. ... After he has completed a certain number ... he views the situations which confront him as a scientist in the same gestalt as other members of his specialists' group. For him they are no longer the same situations he had encountered when his training began. He has meanwhile assimilated a time-tested and group-licensed way of seeing. (189)

Such shared disciplinary referents lie at the heart of Kuhn's account of a scientific community, the members of which "have undergone similar educations and professional initiations ... absorb[ing] the same technical literature and draw[ing] many of the same lessons from it" (177).

BUILDING A SHARED DISCIPLINARY MATRIX FOR MATHEMATICS EDUCATION

As a field, mathematics education has barely started to build a shared disciplinary matrix, codified in a standard literature. This involves not so much a search for grand overarching schemes as the development of many more modest and loosely coordinated analytic frameworks, clearly focused on issues of recognised significance, and closely associated with clusters of exemplary problem formulations and solutions. As I have argued elsewhere (Ruthven *et al.* 2002), taking replication and synthesis seriously could play a significant part in such development. From a technical perspective, replication of a study across varied sites not only makes it possible to address issues of generalisability and contextual influence more rigorously, but provides an important mechanism through which theoretical ideas and research tools can be sharpened and refined in action, particularly in response to the operational challenges and cultural differences which arise in translating them between educational sites, phases and systems and between research teams. From a social perspective, the diffusion of research design and instrumentation from one group to others through replication studies not only mediates the development of more strongly shared systems of language and method, but also directs attention to the degree to which carrying through such work calls for recontextualisation rather than straight replication, illuminating contextual influences and cultural differences which tend to be glossed over in current discussion, evaluation and synthesis of research in the field.

Likewise, critical reviews of research on particular topics, informed by appreciation of such contextual influences and cultural differences, could play an important part in development of the field. At present, the synthesis of research receives insufficient attention, perhaps on account of a popular perception of review studies as 'secondary' rather than 'primary' research, but also because of the challenges of carrying through such work rigorously and reflexively. The development of handbooks for the field is making some contribution in this respect, but their mode of production often limits their scope. Viewed in this light, the first BaCoMET project represented an unusually ambitious and sustained attempt at holistic synthesis across different traditions. Contemporary reviewers recognised this. One commented on how the book examines "aspects of mathematics education which are all acutely relevant to anyone engaged in initial or in-service teacher training" (Schwarzenberger 1987, 67); another considered that the book "makes an important contribution to the emerging consensus on what constitutes the discipline Mathematics Education and provides a very broad range of organising ideas and principles which a contemporary teacher education programme needs to address" (Booker 1988, 505); a third expressed the view – as did the others – that the book "will be a fundamental reference, for years to come, for the training of mathematics teachers" (Adda 1988, 108).

A feature which reviewers particularly appreciated in the BaCoMET text was the attention given to examples. Adda found the closing contribution on classroom organisation and dynamics "a revealing chapter based on many illuminating and stimulating examples" (108), while Booker commented of the same chapter how "the fundamental ideas ... are analysed in depth by the provision of a large number of well chosen classroom examples which are returned to on several occasions to reveal the full inter-related aspects" (509). The opening chapter on social norms and external evaluation took a rather different approach:

> In this provision of illustrative lessons, the first chapter differs from much of the remainder of the book as it allows for the practical implementation and testing of the ideas being put forward. The inclusion of several sets of student-teacher directed questions provide for further investigation of these concepts and strategies (Booker, 1988: 505-6)

This is the only case of exemplary tasks being proposed in the text, as opposed to exemplary analyses being offered. Equally, the sense conveyed – as in the text more generally – is of issues being raised for discussion rather than problems posed for solution. Reviewers noted how the style of the text is more discursive than conclusive; with Adda commenting that "this is not a book that deals with categorical assertions ... it sets problems and puts its readers in the position of posing many more" (108); and Booker concluding that the book "provides a very useful framework for organising the discussion of mathematics education with intending teachers and supplies a wealth of ideas on which practising teachers could well reflect" (510). In this respect, then, the text does not employ generalisations and exemplars in the more definitive way implied by Kuhn's analysis of their function within normal science.

The fullest reviewer comment bearing on the relation between theoretical generalisations and exemplary analyses within the BaCoMET text was occasioned by a chapter on 'observing students at work':

> Using compelling, if familiar, examples [the authors] show how students frequently think about mathematical tasks in very personal ways; when this proves unexpectedly useful, the student is said to have unusual insight but when it leads to a misconception, the student is said to be in error. By highlighting the common basis to these two very different outcomes, the authors provide a meaningful context for the errors that students make in building and using their individual conceptions of the mathematics that the teacher is endeavouring to communicate. Practical means of investigating student conceptions are provided and these observations are related to an underlying theory of human information processing. (Booker 1988, 508)

Given the relatively extensive attention that this area had received within the mathematics education research community of the time, it not surprising to find this particular chapter being singled out. Significantly, some of the points made by Booker can be reformulated in terms of the emergence of critical features of a disciplinary matrix. For example, paradigmatic examples *should* be both compelling and familiar to those in the field; indeed these characteristics are related inasmuch as extensive public scrutiny plays a part in identifying particularly powerful examples. Equally, an important function of such exemplars is precisely to motivate and illustrate deeper theorisation of the issue, serving a bridging function between theory building and practical action. In this respect, Adda was more sceptical: while she considered that the chapter "provide[s] many good examples", she found "the argument of the theory stemming from cognitive science ... a little unconvincing, especially as regards the complexity of the reported observations throughout this book"

(107). Once more, the comment is revealing: while the exemplars are found significant and persuasive, the capacity of the imported theory to enhance their analysis is questioned. Again, it is through wide and sustained public scrutiny of this type that scientific norms are established.

SUMMARY AND CONCLUSION

My argument has been not that issues of mathematics education can – or should – be treated wholly in scientific terms, but for the potential contribution of a normal science of mathematics education to the wider human enterprise. I have suggested that the first BaCoMET project can be seen as an important early attempt to sketch a disciplinary matrix for the field. My critical appreciation of this work has been guided by the different senses of 'paradigm' expounded by Kuhn, and notably by the central part that he identifies clusters of exemplary problems and solutions as playing in mediating between symbolic generalisations and practical action.

Faculty of Education, University of Cambridge

REFERENCES

Adda, J. (1988). Review of B. Christiansen, G. Howson, & M. Otte (Eds.) (1986). *Perspectives on Mathematics Education. Educational Studies in Mathematics* 19 (1), 105-108.

Booker, G. (1988). Review of B. Christiansen, G. Howson and M. Otte (Eds.) (1986). *Perspectives on Mathematics Education. Educational Studies in Mathematics* 19 (4), 503-510.

Christiansen, B., & H. G. Steiner (Eds.) (1979). *New Trends in Mathematics Teaching* IV. Paris: UNESCO.

Christiansen, B., G. Howson, & M. Otte (Eds.) (1986). *Perspectives on Mathematics Education.* Dordrecht: Reidel.

Ernest, P. (1998). A postmodern perspective on research in mathematics education. In A. Sierpinska, & J. Kilpatrick (Eds.), *Mathematics Education as a Research Domain: A Search for Identity.* Dordrecht: Kluwer, 71-85.

Kuhn, T.S. (1962, 1970). *The Structure of Scientific Revolutions.* First and Second Editions. Chicago: University of Chicago Press.

Otte, M. (1979). The education and professional life of mathematics teachers. In B. Christiansen, & H. G. Steiner (Eds.), *New Trends in Mathematics Teaching* IV. Paris: UNESCO, 107-133.

Ruthven, K. *et al.* (2002). Reflections on Educational Studies in Mathematics. *Educational Studies in Mathematics* 50 (3), 251-257.

Schwarzenberger, R. (1987). Review of B. Christiansen, G. Howson, & M. Otte (Eds.) (1986) *Perspectives on Mathematics Education. Mathematical Gazette* 71 (455), 67-68.

Sierpinska, A., & J. Kilpatrick (Eds.) (1998). *Mathematics Education as a Research Domain: A Search for Identity.* Dordrecht: Kluwer.

GUY BROUSSEAU

THE STUDY OF THE DIDACTICAL CONDITIONS OF SCHOOL LEARNING IN MATHEMATICS

Abstract: The production, knowledge and learning of mathematics follow a certain combination of "logics": a logic of knowledge, a logic of the subject and a logic of situations. They do not coincide. How do they relate to didactical activity?
Do the epistemological, semiological and psychological approaches suffice as theories of didactical activity? What is the importance of the latter with respect to the first two? Is it necessarily reduced to the description of practices, or to a collection of techniques or technologies?
Knowledge is the most important means and object of the transmission of acquisitions from one generation to another, and also the most powerful means of influencing nature and people. Its production and transmission have become one of the principal activities of all of humanity. Studies of the conditions for creation and transmission of knowledge now have high priority.

Key words: didactical laws, logic of situations, macrodidactique, microdidactique, production of knowledge, theory of situations, transmission of knowledge.

I. INTRODUCTION

Michael Otte has interested himself in numerous subjects, with a remarkable perspicacity and depth. I have had the good fortune to take advantage of his reflections and he has been kind enough to invite me to offer my ideas in the context of the BACOMET project. He has introduced me to numerous eminent colleagues who have caused me to envisage different approaches to mathematical education. He invited me to write an article with him which permitted certain of my ideas to be expressed in English, but which above all showed me his immense culture and his profound knowledge of French literature.

His interest was thus already in linguistics and for instance in the "science of symbols" (Alleau 1982). For my part, grounded in observation of classes and students, I concerned myself with modelling situations, with ergonomic studies and with the production of experimental engineering. The study of semiological /didactical relations brought us together of a moment.

His interest in the logic of s-knowledge[1] (savoirs) can be explained. The culture, science and organization of s-knowledge – even if they involve more individual psychological processes – are among the most powerful and the most economical of the means and tools of c-knowledge. It is a standard practice to try to derive from them all new c-knowledge. The *fragility of knowledge* (Brousseau and Otte 1991) established for a moment by isolated human beings at the whim of fluctuating conditions would seem to condemn their observation to a crippling sterility. The object of this

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 159 – 168.

article is to make a rapid presentation of the studies I have carried out on the logic of didactical situations in mathematics.

II. RESEARCH IN MATHEMATICS EDUCATION

C-Knowledge (connaissances) in Mathematics Education

The democratic sharing of opinions and responsibilities is retreating everywhere in the face of a redistribution of powers which allocates decisions, even the most personal, to opaque economic organisms disguised by powerless politics and increasingly servile media. The reasons invoked publicly are a cosmetic cover-up of conclusions reached by the "specialists." These are in principle supported by scientific study which is assumed to make their assertions verifiable by those who *can* understand them.

Treatment of society's major problems, health, industrial production, the economy, etc. follows this model more and more. Education appears to be an exception, being perhaps the last domain in which the majority of citizens feel they have the right to claim full competence, from being either parents or students or erstwhile students or simply citizens concerned for the instruments of cohesion and progress for their city. The portion of the population currently devoting itself to the creation and diffusion of information or its teaching probably makes up a majority of the city. In all cases it plays the principal role there. In any case, education is the principal and final instrument of political and commercial use of the intangible but essential product, hope. The hope of improving living conditions for the next generation is an inexhaustible commercial lode. Appetites whetted by this huge market, certain economic interests relentlessly attack the vestiges of the traditional educational structures. In France, successive governments and numerous institutions have for thirty years shamelessly exploited the alibi of education and training to satisfy needs or projects that have nothing to do with education. For example, the duration of non-professional education of teachers of primary school has been augmented by six years, education entirely confiscated by the universities in order to augment the number of their students and thus of professors in the fields that interest them. Not only do they jettison from their programs of disciplinary training everything which might relate to the future career of many of their students and exclude everything that might improve the common professional culture of teachers, but they also fight outside of their realm against anything that might constitute a scientific university setting for teaching.

If a number of actors of the diffusion of knowledge and culture wish to see themselves in the role of experts, the scientific field which should validate them still fails to be clear.

It's not for lack of candidates: every discipline, on one account or another, has been presented as a distortion-free theoretical setting, and sometimes as the unique setting for the study of these phenomena. The disciplines concerned as objects of study were long ago disqualified *a priori* by the fact that they are specific, and thus incapable of assuming the unifying role necessary for the "standard" functioning of the institutions (a role comparable to that of other fields of problems). They have thus been supplanted by "more general" disciplines such as psychology, sociology, linguistics or statistics. One can cite by the dozen the branches which have been called on, from philosophy to computer science, from medicine to ecology, from anthropology to economics, not to mention the *ad hoc* sciences and home made theories.

If effectively numerous interesting works on education have issued from reflections based on these sources, their dispersion damages their unity, and the number of specialists perched on this or that twig of their discipline is not sufficient to assume the support expected of the collection. Nonetheless, social and administrative necessities tend to assemble these renegades into a community of "sciences of education" in the confines of institutions of teaching and training. The repertoire of this community is composed of migrant concepts interpreted in a soft and diverse manner better designed to create complicity than cooperation or real debate. But in fact the most obvious efforts are those which tend each time to diversify yet further the approaches, the concepts, the vocabulary and the practices. Thus educational science forms an enormous field of knowledge, but one without structure.

The scientific fragility of researchers in this domain as they face various classic disciplinary institutions is obvious. Moreover, from one country to another the variety of cultures and institutions illustrates perfectly both the efforts at unification demanded or imposed by those who are responsible and the "natural" efforts of diversification resulting from the individual (and individualistic) activities of the actors.

Despite the diversity of the administrative and cultural organizations for teaching, at the elementary level there is a certain uniformity of objectives and practices, which is encouraging.

The relationship between research institutions and those dedicated to teacher training or the effective management of teaching vary from one country to another and are never very clear. In these conditions the teachers are bombarded with injunctions of all sorts, orchestrated or chaotic, in the name of justifications or slogans whose origins they do not know, whose objectives are the subject of all manner of fantasies, but whose real consequences frequently elude them.

Research in mathematics education adheres approximately to this scheme of things, and for the past thirty years on the basis of a mathematical knowledge all the better shared for being more elementary, most research has based itself on studies of psychology, and most reform on "naïve" propositions of mathematicians. The more the knowledge of the processes of learning and teaching grew, the more the enthusiasm for mathematics education of the community of mathematicians diminished, to the point of indifference. For some it turned into downright hostility towards research, curiously, in inverse ratio to their distance from mathematics itself.

Didactique[2] *of mathematics*

In the space of a short article, this picture is lacking in nuances and does not do justice to all those who are making considerable efforts to offset the faults I have pointed out. But it is not the virtues of the actors themselves which command our attention here, but the phenomena they are fighting. Our goal is to identify the uncontrollable effects of various "laws" which confront them, not in general economic, sociological or psychological laws, but didactical laws – those which are specific to the knowledge in question. These laws arise, by whatever name one gives it, from a "science of the specific conditions for the diffusion of knowledge necessary to humans and their institutions". The ambition of limiting the study to "social projects of causing this or that piece of knowledge – constituted or in process of constitution – to be appropriated by this subject or that institution" would seem more limited, but it leads to the examination of the same field. At the same time, the appropriation of a piece of knowledge implies (or even is equivalent to) its re-creation. Studying it requires comparing it with the conditions of its creation (historical) and of its use in various institutions of society.

Referring to Economics, with which it has many things in common, the study of *Didactique* seems to me to need to be divided in two large portions: *microdidactique* and *macrodidactique.*

Microdidactique is concerned with specific minimal conditions which are at the disposal of a teaching organism to "determine" the appearance, appropriation and use of a precise piece of knowledge, perceptible in the behavior of a human student organism. It is micro in the sense that, like microeconomics, "in its abstract formulations it claims to respect the individuality of each piece of goods and each agent."

Macrodidactique is concerned with the partial or global functioning of aggregates of agents or institutions relative to the diffusion of aggregates of pieces of knowledge belonging or connected to the same discipline.

In relationship with *didactical engineering* which tends to produce projects useful to effective teaching, most of the research work published these days in mathematical education concern microdidactique.

Microdidactique

Studies of the *agents* (the students and the teachers) in their characters, their general behaviors (attitudes, learning) and their interactions occupy most of the terrain. These studies are mostly based on methods imported from psychology, clinical in particular, or linguistics, and on statistics about cohorts of students (rarely about cohorts of classes) and on concepts drawn directly from the teachers' practices but rarely subjected to a tight theoretical analysis.

Studies of mathematical c-knowledge taken up from a didactical point of view have become scarce, and are rarely published in journals of the science of education, perhaps for lack of theoretical support for elementary s-knowledge, and certainly for lack of mathematical and epistemological c-knowledge beyond a certain level.

The modeling of human *interactions* recovered part of the older work in artificial intelligence (or formerly interactions with automated systems or computer systems). It offers its means and its methods ... and its scientific ideologies to every kind of domain, including ethnology and *didactique*. This point of view makes it possible to connect certain characteristics of the agents and of knowledge in the interactions, thus by characteristics other than those arising solely from the logic of the subject or that of the knowledge. But the most direct approach takes as its object the cause and the function of the interactions.

It chooses the conditions which determine these interactions and treats them as *forming a system.* In the "theory of situations," each of these systems or conditions can be modeled by a "situation" – a formal game – which makes it possible to describe and justify (for instance with the help of the theory of games) the actions of the agents with their "milieu." But the ecological and anthropological approach of Chevallard, for example, generalizes this approach and also considers ecosystems without agents, for example a "praxeology" of the c-knowledge which permits a piece of s-knowledge to function.

The theory of situations thus brings out the role of another type of logic which integrates and subordinates the two preceding ones. This approach may still appear a bit exotic to some, even though it has been developing for around thirty years and has brought in a number of new and useful concepts. For the past fifteen or so years, only a few sociology researchers (Barwise, Berger and Luckman, Quéré) have taken up the ideas of McHugh and have occupied themselves – independently of *didactique* – with contrasting a logic of situations with that of groups or individuals.

To conclude, microdidactical approaches are numerous and varied, but they restrict themselves to one, or occasionally two, of the three logics: that of the subject, that of s-knowledge or that of situations. Very few of the works really combine them and establish relations among the three.

Macrodidactique

On the other hand, studies of *macrodidactique* are very rare and only mobilize parameters of very little scientific value. Now, the most important difficulties and the majority of those that are encountered by the teaching of mathematics are of macrodidactical nature.

One example may give at least an idea, if not a proof: In the sixties, a study of the processes of multiplication and division as taught in the French schools demonstrated theoretically that an appropriate disposition of the calculations would make it possible to improve decisively and durably the performance of nearly every student. This ergonomic study was supported by the precise measurement of performances of large groups of students. A didactical experiment carried out in the classroom proved irrefutably the validity of the study: the improvements were a little better than predicted. It also showed that the time required for learning could be materially shortened. One could hope to replace nearly two years' worth of daily computational drill by more valuable mathematical activities without losing any of the students' ability to calculate. The publication of this article in the acts of an international colloquium of science of education did not provoke a single question or awaken the faintest interest. Later and repeatedly, supported by this work, the project of reforming the teaching of the processes of calculation was proposed up to the highest levels The reception of the idea became more civil as the age and reputation of its author grew, but not a one of the people asking about it ever envisioned trying the reform. It is easy to understand why changing cultural practices that are the most elementary and the most wide-spread in the whole of a population is a didactical enterprise that appears infinite, no matter how beneficial the aim of it. And in that case, the simple and non-mysterious nature of the suggested modification made it appear trivial and almost ridiculous. What mathematical savant would dare propose such a project to his minister, and what politician would want to stake his reputation on it? Once the microdidactical or pedagogical difficulties are resolved, there remain those which reside in the relationship of society to the piece of knowledge in question and the learning of it. Conceiving of the conditions for a reform of the human methods of calculation used by the population would have arisen from macrodidactical c-knowledge which was absent at the period. This reform seemed useless when calculators were appearing. This reform would have required other means than new objects or procedures of teaching, and other knowledge than that of the psychology of children.

Should we then deduce that the results of microdidactical research, even when solidly established and very rationally connected to their application, are condemned to be useless? No, but it is necessary to apply to the diffusion and to the use of this type of knowledge the same analysis of functionality and application that we apply to mathematics itself, applying the *didactique* of mathematics to the *didactique* of the *didactique* of mathematics!

III. THE LOGIC OF SITUATIONS

Thus the *didactique* of mathematics concentrates on the *study of the conditions of learning and teaching which are specific to the knowledge aimed at*, this study being in relation to the relevance, adequacy, dependability and economy which the knowledge in question procures in these circumstances (**Principle 1**).

The fundamental hypothesis of the theory of situations (Brousseau 1998) is that *the conditions which prevail for setting a piece of knowledge in action do not act independently of each other* (**Principle 2**). (They are never optimal at the extremes of their interval of action) It is therefore necessary to consider them together. They form *systems* which it is convenient to model before calling attention to them. Moreover, the sole means available to teachers is to recreate a set of favorable conditions

The general model maintained is that of the economic theory of games. (**Principle 3**). A subject does what works best for him given his circumstances and his projects. The objective of his reactions or his decisions is to minimize and regulate the perturbations imposed on him

Thus in the *theory of situations* the method of defining a concept C is the following:

"C is the object that resolves the situation S *optimally.*" (**Principle 4**).This mode of definition specifies the categorical definition used in mathematics: "O is the object which satisfies the relationship R(O)" This principle is applied first to mathematical knowledge. A mathematical notion cannot be analyzed in the Theory of Situations in the *didactique* of Mathematics until the moment when it has appeared as the solution to a situation. From this arises the phenomenotechnical and theoretical importance of didactical engineering.

These situations can be composed and decomposed in various ways (**Principle 5**): for example a subject can be an actor simultaneously in several situations with his decisions in one depending on conditions determined in the other, as when a subject acts by exchanging information with a correspondent about his action The observer is part of the system and his own game should be analyzed as a subsystem. The comparison of his "models" with what happens requires a certain engagement on his part and he must verify them by this means.

We have been lead to a coarse classification of these systems of interaction in terms of various criteria: presence or absence of an actor with didactical intentions with regard to the others (didactical or a-didactical situations) and the types of reaction which they produce (actions on the milieu, formulations, assertions, devolution or institutionalization). The use of these types makes it possible to describe and explain a good many phenomena of teaching, but they chiefly serve as an entry point for research on conditions specific to a given piece of knowledge.

I will recall here the paradoxes which led to giving didactical situations a different model from non-didactical situations. They have been widely presented (in particular in our article G. Brousseau and M. Otte, 1991).

Principle 4 aims to establish a correspondence between mathematical knowledge and situations. Clearly any "real" situation mobilizes a great deal of knowledge. The ones which interest us here are those which cannot be resolved except by originating a piece of knowledge, but which can be presented by using more elementary knowledge which should already have been acquired. The prerequisite knowledge is that which permits the subject to learn the rules of the game and imagine some basic strategies (whether or not they actually resolve the situation). The solution knowledge is what the optimal strategy leads to. It is in this sense that the following principles are formulated:

Every mathematical notion has at least one situation that characterizes it. (**Principle 6a**)

Every situation (considered by the Theory of Situations in the Didactique of Mathematics) determines a set of pieces of mathematical knowledge which are indispensable to the invention of its resolution. (**Principle 6b**) Depending on the familiarity of the situation, the forms of knowledge required to solve it will modify themselves in order to reduce the cost of use.

Note that this principle makes it possible to characterize mathematical knowledge following new variables (relevance, adequacy, adaptation, economy, dependability in a field, cost of use or of learning, etc.) and not only from the point of view of validity. It tends to offer a means of producing a hierarchy, at least locally, according to their *complexity*. It has been shown that consideration of situations makes possible a certain liberty in the conception of the didactical articulation of knowledge relative to axiomatic ordering. This possibility is very important for the construction of meaning of the knowledge taught.

We have added to these principles a working hypothesis, a hypothesis which is probably useless for establishing the consistency of the theory but very productive in the search for processes: the collection of situations of a certain type (action, formulation, validation, etc.) relative to the same piece of mathematical knowledge has at least one generator: *the fundamental situation of this knowledge.*

At this point in the exposition one should understand "generator" in the static sense: the schematic description of this situation results in the appearance of a set of specifications (values or intervals) of a set of variables which are satisfied by all situations associated with this knowledge and only by them. But one might understand it in the dynamic sense as "a situation which will cause the others to appear by the *process* which it engages" as soon as this notion of process is introduced.

For the moment, at this stage of the theory, "to learn" is synonymous with "to change strategies in a characteristic situation," but this definition is much too restrictive to account for all the causes of more or less permanent adaptation and to distinguish the causes of adaptations or *causes of c-knowledge* and the reasons to know (their necessity in the s-knowledge) to which they should lead.

Along with the decomposition and composition which fix the synchronic relationships of various situations there appear *diachronic conditions* relative to their succession, articulation in process.

IV. THE LOGIC OF PROCESSES

Separating the pieces of knowledge associated with a situation into base knowledge and resolution knowledge establishes a condition on the temporal succession of situations a subject can approach. These ordering conditions are a great deal more complex but also more flexible and realistic than those used in the classical pedagogical analyses: certain forms of knowledge (implicit models for example) suffice for the introduction of others. They intend to establish under certain conditions the legitimacy of a "functional" and "usable" order aside from the classical axiomatic or rational order.

An initial use of the Theory of Situations in *Didactique* of Mathematics consists for the teacher of arbitrarily choosing the situations she will offer her students and the order in which she will offer them in such a way as to "construct" the knowledge that she wishes to teach while contenting herself with respecting the temporal order expressed above. She will thus be able to use a "curriculum" of non-didactical situations, but the genesis of the knowledge itself will be entirely didactical.

Now, while a situation calls forth a knowledge for solution, and can sometimes provoke and permit its invention, it also generally produces a good many new questions which are the source of the situations for the future. The manner in which an answer produces new questions is an object of study which has been a bit neglected in the research of these past twenty years, but which has a lot of importance in the comprehension of the process of learning and of teaching. It's a matter of a procedure for aggregating situations such as we have described above, but one which requires new conditions. We won't give here the supplementary principles extending the theory of situations into a theory of processes.

A didactical process is a series of didactical situations relative to the same piece of knowledge (object of teaching or of learning) and such that in order for one to succeed, all the previous ones must have succeeded. This definition is suitable for a description after the fact. After the conception of a normative curriculum one uses other terms (program, progression,...) But if the situations are relatively a-didactical it is necessary to consider the local and temporary reasons one has at the end of each situation to consider the next. Thus not only can each situation be proposed thanks to the acquisitions for the preceding ones but it is in addition (more or less) justified by the questions raised by the preceding one. These justifications have two sides, the justifications for the teacher (a step in a curriculum project, for example) but we will interest ourselves here in justifications for the students (intelligibility of the situation, relevance and immediate interest of the questions, possibility of solving, etc.)

Such a process constitutes a "genesis" of a concept or of a notion, that is to say, a construction. It's a matter of a *chronogenesis*, that is of a genesis where the links between the pieces of knowledge are determined by their place in a history and by the relations of causality, of dialectic. The authentic – historical – chronogenesis of a piece of knowledge can only rarely serve as a didactical model because the situations of which it is composed are far too complex, the effective conditions of discovery cannot be reproduced, the motivations and repertoires of the original constructors are often very distant from what the culture remembers of them.

An important part of mathematical work consists of substituting for this frequently chaotic or hesitant chronogenesis of a piece of mathematical knowledge a logical, ergonomic and if possible elegant construction which will simultaneously permit its verification, comprehension and use. This continual effort of reconstruction is of a didactical nature. It ends up with a *topogenesis of mathematics* where every object has a place according to its definition and its properties, in a partial ordering determined by relations of logical necessity and of ergonomy.

In a topogenesis and in a chronogenesis of the same piece of s-knowledge the objects are mathematically equivalent, but their organization, their reciprocal places, their significance and their environment of c-knowledge are different.

The merit to a topogenesis is that it structures knowledge in such a way as to minimize memory, risk of errors, redundancy, effort for communicating,... But it also tends to cause the disappearance of the conditions which made the knowledge necessary and functional. The chronogenesis is always much closer to the real functioning of mathematics and because of that much better motivated and even exciting for the students, but it is also much more random and extravagant in learning time than the topogenesis. And since in the end the students need to leave school with knowledge structured according to the topogenesis of the moment, if one uses a chronogenesis that is too different, one is obliged reorganize it, which takes even more time.

This example gives a clear demonstration that *didactique* does not consist of determining norms, but rather of studying the equilibrium between opposing constraints

CONCLUSION

Our manner of studying the conditions for the functioning and learning of knowledge permits us to place the use of mathematical or psychological or other knowledge which one wants to use in teaching under the control of a theory of *didactique*. It should not be understood that the conditions whose effects we are studying are *obligations* or norms which teaching should realize, or restrictions which condemn us to examine only a part of the effective situations. On the contrary, every situation of real teaching, by the very fact of existing, leads us to think that it should satisfy a certain number of conditions which we ought to be able to study and model in the Theory of Situations in *Didactique* of Mathematics.

Université "Victor Segalen", Bordeaux2

NOTES

Translated by Virginia Warfield

[1] In order to convey the distinctions between the French words "savoirs" and "connaissances," both of which translate to "knowledge," we use the following definitions: C-knowledge means knowledge as a means to make a decision, or understanding in the sense of having a familiar relationship. "C" is initial letter in the Latin word "Conoscere," from which are derived "connaissances" (French), "conocimientos" (Spanish) and "conoscenza" (Italian). S-knowledge means knowledge as a cultural and social means to identify, organise and communicate the C-knowledge. "S" is the initial letter of the Latin word "sapere," from which are derived "savoir" (French), "saber" (Spanish) and "sapienza" (Italian).
[2] Didactique des mathématiques, «Science des conditions spécifiques de la diffusion des connaissances (mathématiques) utiles aux humains».

REFERENCES

Alleau, R. (1982). *La science des symboles*. Paris: Payot.

Aveline, C. (1961). *Le code des Jeux*. Paris: Livre de Poche Hachette.

Brousseau, G. (1998). *Théorie des Situations Didactiques*. Grenoble: La pensée sauvage éditions. (1997 in English: edited by N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, *Theory of didactical situation in mathematics*. dordrecht: Kluwer.)

Brousseau, G., & Otte, M. (1991). Fragility of Knowledge. In A. Bishop, S. Mellin-Olsen, J. Van Dormolen (Eds.), *Mathematical Knowledge: Its growth through teaching*. Dordrecht: Kluwer Academic Publisher.

Brousseau, G., & Warfield, V. (1999). The Case of GAEL. *Journal of mathematical Behavior* 18 (1), 7-52.

Brousseau, G., Brousseau, N., & Warfield, N. (2002). An experiment on the teaching of statistics and probability. *Journal of mathematical Behavior* 20, 363-441

Chomsky, N., & Miller, G. A. (1968). *L'analyse formelle des langues naturelles*. Paris: Gauthier Villars.

Diderot, D. (1773). *Le paradoxe sur le comédien*.

Dokic, J. (1999). L'action située et le principe de Ramsey. In Fornel, M. de, & Quéré, L. (Eds.), *La logique des situations*. Editions de l'EHESS.

Ekeland, I. (1974). *La Théorie des jeux et ses applications à l'économie mathématique*. Paris: P. U. F.

Fornel, M. de, & Quéré, L. (1999). *La logique des situations*. Editions de l'Ecole des Hautes Etudes en Sciences Sociales.

Huizinga, J. (1951 <1938>). *Homo ludens. Essai sur la fonction sociale du jeu*. (Transl. C. Seresia). Paris: Gallimard.

Moulin, H. (1979). *Fondation de la théorie des jeux* (containing R. de Possel: Sur la théorie Mathématique des jeux de Hasard et de réflexion). Paris: Hermann.

ROLAND FISCHER

THE FORMAL, THE SOCIAL AND THE SUBJECTIVE:

Variations on a Theme of Michael Otte[1]

Abstract. With strong reference to his book "Das Formale, das Soziale und das Subjektive" ideas of M. Otte are put into relation with some deliberations of the present author. These concern the role of science in society, general education, the role of mathematics in education, the limits of mathematics and their social relevance.

Key words: complementarity, consciousness of society, diagrammatic thinking, general education, limits of mathematics, logical types, mathematics and organization, sociology of knowledge

Michael Otte has written a great deal, and I have read a great deal of his work. His book *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik*[2] (The Formal, the Social and the Subjective. An Introduction into Philosophy and Didactics of Mathematics) presents a summary of his previous work, a kind of interim précis. I would now like to offer a few variations on his ideas; these will be interpretations of his thoughts, comments and additions, as well as some alternative perspectives.

THE CENTRAL CONCERN

What is important to Michael Otte; what are his concerns? As I understand him, mathematics is not simply a discipline like any other; mathematics reflects central elements of human existence and of society. Contemporary existential problems can be reflected upon, understood and dealt with in terms of mathematics – mathematics not in the traditional sense of applied mathematics, rather that mathematics is used as a medium for reflection, similar to a work of art. This requires overcoming limitations and biases in mathematics with, for example, the principle of complementarity. Many limitations and biases are indeed an inherent feature of mathematics, a condition for its effectiveness. In the area of didactics, however, a more complete, differentiated yet simultaneously comprehensive view is required. In this context I recall a phrase from Arnold Kirsch, who said, "Teachers must draw distinctions where mathematicians draw none."

An example of this, the one with which Michael Otte begins his book, is the *differentiation of logical types*. The linguistic variant of this is that there is an object language and a meta-language. Michael Otte's thesis is that mathematics depends on non-differentiation, on evening-out. Differentiation comes only on subsequent analysis. Otte speaks of a "symmetrizing, non-hierarchical handling" of logical

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 169 – 178.

types[3] (36). To be more concrete, here is an example which in didactics has been frequently analyzed: the handling of variables. This takes the form of a permanent interplay between the variable as a specific object (e. g. as an unknown, but definite number) and as a sign, which undergoes various operations according to determined rules. Underlying this interplay is the complementarity of semantics and syntax, of content and logic. Any resolution that tends to one of these directions has its disadvantages; one loses sight of the essentials. A competent mathematician will avoid this mistake; he intuitively alternates between levels, but in any event he does not routinely reflect upon what he is doing. The problem is first of all one of basic theory: the familiar antinomies; this problem can be solved by sophisticated scrutiny. Second, there arises a didactic problem, namely, how do I teach someone to learn without reflection? Third, the complementarity also relates to a *societal problem.* We live in a formalized society that functions largely according to rules, be they legal-economic or technical. Setting this layer of rules as an absolute leads to dehumanization; on the other hand focusing on the material level, that is, on people, overlooks the necessity for social mechanisms. And although simply dealing with complementarity in an intuitive way is perhaps possible in individual cases, it is difficult to translate into social patterns. That requires conscious, organized reflection. Here is where mathematics could make headway, initially with its ability to allow the layer of rules to assume concrete form; above all however through a more precise scrutiny of the activities of mathematicians themselves and historical developments. However, with its tactic of evening out levels of communication, mathematics can hinder exactly this social learning process.

To put it somewhat more dramatically: Mathematics and its instruction are today at a crossroads. They can either chart a course towards a *mechanistic, unconscious world*; *or* they can help create a *new social consciousness*, a new step toward release from a self-imposed immaturity; a new enlightenment, as it were. Michael Otte's work in my view makes the case for this second path, and as such is a challenge for anyone who has to do with mathematics and its instruction: teachers, teacher-college educators, school curriculum administrators and of course mathematicians themselves.

That actually says everything. I am following a dramatic structure that Michael Otte applies in his own work: to state essentials at the beginning, the rest being variations on the theme, where each individual piece often contains the whole in fractal form. That is how Michael Otte once explained his methods to me, when I told him of my difficulties in reading his work. His writings as a rule do not have the linear-logical structure that we are accustomed to find in mathematics and in the sciences in general.

Nevertheless a conceptual construct does underlie the following remarks: to proceed from the general to the specific. Society-science-education-mathematics are the themes, and that last theme once again subsumes all the others. That is exactly the intended point.

SOCIETY AND SCIENCE

Michael Otte describes society as follows (131): on the one hand there is *variety*; there are many realities, culturally, politically and in other dimensions. The autonomy of subsystems, of regions and of individual people (singleness) has to a large extent been realized. On the other hand there is *uniformity* and connectedness, for example in commercial products (hamburgers, Coca-Cola and jeans), in production methods, in life styles, in the communications media, and so on. There is, then, differentiation as well as integration. Social complexity is a result of the meeting of these two tendencies, at the same time as it intensifies abstraction in self-depiction.

To expand upon this statement, the bonds of society are based upon rules: economics, technology and political-administrative *regulatory mechanisms* are what hold society together. We trust such mechanisms, trust the various "invisible hands," and surrender ourselves to them. They are exactly what subsequently make variety possible on other levels as well; but at the same time this self-surrender also implies a renunciation of any holistic consciousness of society.[4]

According to Michael Otte, the ambivalence between differentiation and integration and the accompanying increase of complexity and abstraction also occurs on the level of *knowledge*. Differentiation and heterogeneity are obvious today. According to Michael Otte the effort towards integration has led to an abstraction, to the theoretization and methodolization of knowledge (131, 149). Scientific knowledge since about the 19th century has been regarded as a "system of instruments for the reconstruction of reality." (131) Its connection with reality is becoming increasingly indirect. Michael Otte regards the interest in epistemology itself as a consequence of the fragmentation of a uniform world view. Through the detachment from a direct connection to reality, which has been pursued most consistently in the field of mathematics, but observed by theorists of science in the other sciences as well, the claim for total coherence becomes possible, even if it is reduced to little more than freedom from contradictions.

Besides the mathematical course in its widest sense, another way to deal with the complexity of knowledge is to attribute it to social complexity, as the *sociology of science* and *sociology of knowledge* attempt to do, from Karl Mannheim to Thomas Kuhn and today above all in science studies. The gain of the sociological approach compared with the philosophical approach is that sociology brings more complexity into play (A. Comte) (399). Whereas the philosopher usually assumes a single subject of cognition, the sociologist accepts a priori the multiplicity of individuals.

Michael Otte, however, adheres rather to a *philosophical course* in ascertaining that science has a special responsibility that concerns its links to society. He writes:

> Science fulfils ... in human evolution a specific function that is essential for that evolution altogether and as an overriding objective. This is meant in terms of the concept of truth or objectivity, as opposed to the particularity and subjectivity of individual interests. (344)

My comment to that statement: I am of the view that science has exhausted its ability to provide truth and objectivity as forces for connectivity in society. Science had supplanted the traditional connection-providing institutions of faith and religion which had operated through the Church and ruling dynasties (the great enlightenment project). But science, in its function for providing social synthesis, has itself now reached a crisis. However, it is often said that we live in a *knowledge-based society*. Knowledge is meant in this sense in terms of particularity. Pieces of knowledge are factors of productivity which can be bartered: knowledge as wares, so to speak. Nevertheless we have simultaneously overcome the knowledge-based society in another way: it is not knowledge which binds us together; it is rather, as mentioned earlier, a rules-bound dependence on one another. To mention the drawback of this situation once again: the whole *does not come into view*; the collective capacity for action exists only in terms of contextual mechanisms, whereby these are set up as absolutes and the resulting problems cannot be addressed.

EDUCATION AND WHOLENESS

Although in the matter of truth/objectivity I am of a differing opinion from Michael Otte, I share his view that the sciences have a contribution to make to overcoming the particularity of society, and hence a contribution to a holistic view. For Michael Otte this is part of what he describes as education. He believes that the sciences (and not only schools) have an *educational responsibility*. Reducing science to research seems to him an inappropriate limitation. He writes, inter alia, "Science = Research + Learning," or, "Science = Productivity + Culture," or again, "Science = Tool + Reflection." (190) And he postulates that learning must be concerned with the *integral whole of society*. On the other hand education also has to do with autonomy, specifically the *autonomy* not only of the individual who is learning, but of the school as an institution as well. Michael Otte formulates it as "respect for the self regulation of learning processes." (132)

One could also say that it is *the exchange between several whole entities* that should be encouraged: that of the individual, that of society, and, so speak as an intermediate and connecting link, that of the learning community. Because a comprehensive wholeness can never be scientifically established, something like a moral component enters into the matter, as Michael Otte observes. (130, 132)

In the implementation of a demanding educational responsibility, a specific entity, namely the *teacher,* has for Michael Otte an outstanding role to play. He describes the teacher as an "*exemplary intellectual,*" who is effective not only through what he does, but for who he is. (161) In this regard Michael Otte considers various concepts of what an intellectual is; in each case the intellectual obligates himself to a holistic social idea, to the "totality of the system of reference," as he calls it. That Michael Otte is an intellectual in the special sense of the word is obvious to me. However, there are not many like him. Thus there arises the problem of the education of such intellectuals, or, put another way, the question of what a *general education* is or ought to be.

My suggestion: instead of being simply an acquisition of general knowledge, the process of general education signifies *the systematic practice of establishing connections and relationships between one branch of knowledge to another.* This is a question of determining the relationships of elements of knowledge to one another, including even contradictions, especially when taking into account differing basic conceptions and views of mankind and of the world: indeed when all the different entities are considered. *Interweaving and contradiction* is the shorthand term I have introduced for this approach. *Social consciousness* in this case cannot (any longer) be understood as fixed, common knowledge; it is rather a process of people constructing holistically interwoven theories of the world (and thereby of themselves) and of their subsequent deconstruction as a result of continuing human development. To be an educated intellectual means taking part in this process. Here questions of relevance and assessment play a significant role. The task of a public education system is a dialectical one: to offer a comprehensive picture and at the same time to enable people to discuss it critically while keeping in mind the question, "what does all this mean to me/us?"[5]

Where for me establishing interweaving and contradiction, construction and deconstruction is mainly a communicational and organizational problem, Michael Otte keeps the individual more sharply in focus. Whereas I take for granted that overcoming social integration that is merely rule-oriented is in fact our mutual objective. And my approach also relies on intellectuals of the cast of Michael Otte. I notice this increasingly with such entirely pragmatic questions as: who should be the recipients of events or publications for which I or the group I work for are responsible? There must be agents for that process of construction and deconstruction, of interweaving and contradiction who transcend the field of science. To regard teachers here as the ideal group is appealing to me and would at the same time have implications for their vocational profile, for their place in society and, finally, for their training.

COMPREHENSIVE DISCIPLINARY DIDACTICS

And so I come to the subject of didactics of disciplines, to which I would like to contribute a thought that for its part also concerns organizational matters. The question is that of a holistically integral relationship to society in the areas of education and science. However, the sciences themselves give the impression of being fragmented – into disciplines. I believe there should be serious thought given to the extent that didactics of disciplines should be interdisciplinary, not simply in terms of establishing links to psychology, pedagogy, sociology, etc., but especially to other parallel disciplines, i. e. to other school subjects. *A comprehensive didactics of related disciplines,* then, *rather than the didactics of separate disciplines.* Disciplinary boundaries perhaps no longer even allow for postulating essential questions. And in strategic terms: The didactics of different disciplines have always received their stimulus as a result of the significance of those disciplines; this has also always been a problem for their independence. Because today so many disciplines no longer have unquestioned relevance in the educational canon, new directions in the self-understanding of didactics should be seriously considered. The concept "didactics of a discipline as an advertisement for that discipline" has in any event always been problematic.

Michael Otte would possibly not agree with the foregoing ideas, because for him *object-orientedness* is an essential condition for scholarly practice. There must be an object and resulting resistive force. He repeatedly warns against the dissolving of science or education into mere abstract communication. And what insures a material object better than establishing its boundaries?

MATHEMATICS AS AN EXEMPLARY GENERALITY

In keeping the universality of science and therefore the totality of society in view, Michael Otte pursues a different course. His concern is that within the individual scientific disciplines a researcher dedicated to particular field of study should always keep sight of general principles that reach beyond his particular field. In Otte's mathematic-didactical and mathematics-philosophical writings one repeatedly comes across broad statements that apply to mathematics, to science in general, indeed to life itself. A few examples: "The hierarchy of logical types is to be noted in life, knowledge and logic; at the same time it cannot be noted." (31) Or: "There exist no generality without activity, movement, change." (79) Or: "Generality is recognizable and explains the particular, the accidental. That which is undefined controls what is defined." (83-85) Or again: "We must determine something definitively, at the same time that we must keep those determinations in perspective. One should value one's own importance, but one should never take oneself seriously." (264) How does one arrive at such insights if one concentrates on mathematics?

I believe (and I think this is also Michael Otte's opinion; I shall however formulate it somewhat more radically) that mathematics has the potential to become *the quintessential scientific field in education*. As such it would be an alternative to the interdisciplinary study that I have just propagated, or at least complement it. That is to say, we mathematicians can offer something that extends far beyond what is usually understood as mathematics, and which has special relevance in modern times. That immodest statement may be less appealing than saying that in the concert of disciplines we also have a contribution to make. However, immodest propositions are urgently necessary. The human need for orientation is very apparent, and if the scientists have nothing to offer it, others will: journalists, sect leaders, and practitioners of the esoteric arts. In passing, the worst offerings don't come from science journalists.

Where, then, are mathematics' unique possibilities? Essentially it is the *objectification of structures*. "Everything is number," said the Pythagoreans; the modern variant of this is "everything is structure". This view includes everything from the composition of matter, the laws of cosmic motion, the depiction of life as a special form of organization, the development of interactive social and economic models, to the perception that matter itself is nothing but structure (e. g. nothing but an asymmetry of space). The specific achievement of mathematics is to turn these structures into objects of observation, of study and, eventually, of manipulation. Visual and material representations provide an important means to accomplish this. Michael Otte speaks in connection with Charles S. Peirce of mathematics as "diagrammatical thinking." (94) He quotes Peirce: "Mathematical thinking means the providing of experiments with diagrams and observing the results of those experiments." (382/3) These objectifications provide the requisite resistivity; they give rise to the autonomy of the object and lend it stability. It is the sort of stability that can generally be produced through textual means, but that does not exist in a purely oral culture; whereas mathematics, with its possibilities for systematic manipulation of symbols, goes beyond written speech.

With diagrams the abstract, the invisible, the intangible, on the other hand the invariant (that which Durkheim describes as "the wisdom of the community" (409/10)) can be permanently preserved. They can be studied and transmitted. In this light the frequently discussed relationship of mathematics to the natural sciences, or more exactly to the material world, is turned on its head: mathematics is applied, material (inanimate) nature. [6]

The simplest type of structure is a *distinction*. "Draw a distinction" is, according to Spencer-Brown the first step in any systematic procedure[7]. The introduction of a sign implies a distinction and/or creates it. In any case the sign thus gains stability. The basic prerequisite for any mathematical process is a distinction, at least that between an empty set and the notion of it. The Cantorian definition of set speaks of "well-differentiated objects" that must be taken as a premise. To this Michael Otte remarks, "in the final analysis mathematics is based upon the possibility to observe distinctions in a spatial-temporal world, and to draw conclusions from it."

LIMITS OF MATHEMATICS

Every possibility has its limits, including that of the objectification of structures. The educational process, especially when it aims for comprehensiveness, must reflect these limits. I would like to point out one limit in particular: *Mathematics does not allow for objects to include the structures that contain them.* To put this more concretely in terms of mathematical function: an element x, to which a function f is applied, and the function f itself lie on two different levels. The element x may not determine the function f. This applies as well when considering functions of several different variables: mathematical processes are applied to arguments that are not determined by those arguments. Michael Otte describes it thus:

> Objects have as it were no influence on 'what happens to them,' in the sense of applying the function. The algebraic function which assigns the element x^2 to an element x establishes itself, it would seem, independently of the nature of the individual argument, designated here as x.

Mathematical relations become independent of the objects they relate.

Carrying the situation into the social realm, one can make the following analogy: x or several x_1, x_2... would represent individual people, f an organizational structure, through which these people are linked with the objective of accomplishing a particular task. The functional value would be the particular output of the organization. In accordance with functional principle, the organization cannot be defined, at least not entirely, by its members. What this means for our administrative organization where the functional concept prevails, is obvious. It is an *undemocratic* feature: the structure does not derive from elements, at least not entirely. This is not meant to be a left-wing criticism of business management (or indeed of jurisprudence); not everything must be democratic. It is, however, a mathematical limit in my view.[8]

But we have yet to come to Michael Otte's main point. In the *historical development of ideas* and in practical interaction with them he sees an *overcoming of the undialectical separation between relations and the objects they relate,* of function and argument. For the development of the concept of function the parallel development of the concept of real numbers, especially the notion of continuity in the field of argument, was an essential condition (403). The differentiation of the concept of function, the emergence of the concept of continuity have been closely linked with the development of the concept of the real number. One could certainly put a set-theoretical concept of function at the beginning of a lecture, or construct real numbers without reference to the concept of function. However, the historical dialectic would then be lost.

Incidentally, without formulating the social analogy explicitly, as I have just done, Michael Otte speaks of an "*equality of objects and relations*," (404). It is however obvious that as a whole he is making reference to social realms, and this concern emerges from other formulations of his as well. He writes, for example, "Coherence, as well as formal consistency, are possibly outmoded requirements and should be replaced by other forms of cooperation." (291)

(In writing this paper, a difference between Michael Otte's thinking and my own became clear to me, which perhaps explains several arguments that he and I have had. I accept mathematical biases (for example the separation of relations and the objects that are related), and so set a mathematical limit and try to overcome this limit by the way mathematics is handled, especially by establishing a social context. Michael Otte on the other hand sees the process of overcoming within mathematics itself; at least when taking its development into consideration and bearing in mind the creative mathematician. I consider this a difference in strategy.)

PRESERVATION OR TRANSCENDENCE

The representation of a structure (function, organization) with mathematical methods in the widest sense can serve two entirely different purposes. The first is in order to preserve this structure, to stabilize it. This is necessary to mathematics; while calculating the value of a function, for example, that function may not be altered. But preservation can also have a social effect, in the sense of preserving or even justifying existing conditions. Second, a mathematical representation can be used to promote discussion about the structure and to change it if necessary. In this way the bounds of mathematics are exceeded in the narrow sense, unless a larger framework is given. (However, a framework in the largest sense, e. g. the set of all sets, cannot be expressed in terms of mathematical concepts.)

If one at least keeps the second course open, then mathematics also has the *potential to transcend itself.* This would be a significant accomplishment that could transform mathematics into a pre-eminent field of study. In this context, I would like to repeat a quotation from Michael Otte already cited earlier: "We must determine something definitively, at the same time that we must keep those determinations in perspective." (264) Mathematics can be useful in both cases. What is particularly novel here, and which mathematics makes possible, is that these determinations and perspectives are raised beyond merely intuitive-individual interplay and allow for social reflection and action by means of objectification. Through the medium of mathematics the broadening of philosophy, which until now philosophy itself has not succeeded in accomplishing, could take place: to establish the transition from the reflection of a single subject and an individual – even an absolute – into a *collective reflection*. Put another way, mathematics could create a synthesis of sociology and philosophy.

CONCLUSION

It may be that some might shudder at the thought of this broad arc; that the mathematics one has learned and loved might be lost. Some may feel that the whole purpose of mathematics as an educational discipline is to set concrete, definable terms against such scholarly-speculative constructs. But I believe that keeping to such a conception of mathematical instruction will leave it no great future, because other fields such computer science have already begun to challenge it. On the other hand, it is also clear that with an alternative programmatic model such the one as I have outlined here, the efforts to come to terms with problems in this area are still not complete. However, it would be wrong to believe that such a model could not be realized. In numerous essays, books and in the dissertations he has advised (more than twenty) Michael Otte has let his philosophical approach assume concrete form. He is, as I have already mentioned, a passionate advocate of a postulate of objectification for the sciences: one must have something material to manipulate; something not subject to the preferences of a human agent and which offers resistive force. He and his students have often taken texts from the *history of mathematics;* through these texts and through consistent and thorough application of dialectical mathematical understanding they have succeeded in establishing a new form of scientific scholarship in the didactics of mathematics that in my opinion is altogether unique. These new directions, or at least the recognition of the superior abilities of their leading proponents (these include Michael Otte's students as well), have won them respect and notable success.

Allow me to set forth this important issue once more. With mathematics unconsciousness of society can in the long term be preserved or transcended. As such mathematics is dangerous, but can also be beneficial. At Plato's Academy of Philosophy it was common knowledge that no one could enter who had not learned mathematics. To ensure an interaction with mathematics that encourages consciousness, I would turn Plato's phrase around: *No one who has not learned philosophy should be allowed to learn mathematics;* otherwise he would be a danger to humanity. And Michael Otte is *the* advocate of the sort of philosophy I mean.

Institut für Mathematik, Universität für Bildungswissenschaften, Klagenfurt

NOTES

1 I am indebted to Jim Edinberg for his understanding translation of the original German text into English.

2 Suhrkamp Taschenbuch 1106, 1994.

3 The page numbers in parentheses here and throughout refer to the above mentioned book "Das Formale, das Soziale und das Subjektive." M. Ottes's texts have been translated by Jim Edinberg.

4 See Fischer, R. (1998). Wissenschaft und Bewußtsein der Gesellschaft. [Science and Consciousness of the Society]. In Gubitzer, L., Pellert, A. (Eds.); *Salbei und Operndujft. Reflexionen über Wissenschaft*, 106-120. Wien: Zeitschrift für Hochschuldidaktik 3.

5 See R. Fischer (2003). Höhere Allgemeinbildung und Bewußtsein der Gesellschaft (Higher General Education and Consciousness of tue Society). *Erziehung und Unterricht* 5-6, 559-566.

6 See Fischer, R. (1999). Mathematik anthropologisch: Materialisierung und Systemhaftigkeit. (Mathematics Anthropologically: Materialization and Systemness). In G. Dressel (Eds): *Mensch – Gesellschaft – Wissenschaft. Versuch einer Reflexiven Historischen Anthropologie*. Innsbruck: Studien Verlag, 153-168.

7 Brown, G. S. (1969). *Laws of Form*, Bantam Books, Toronto/New York/London, 3

8 See Fischer, R. (1999) Technologie, Mathematik und Bewußtsein der Gesellschaft (Technology, Mathematics and Consciousness of tue Society). In G. Kadunz, G. Ossimitz, W. Peschek, E. Schneider, B. Winkelmann (Eds.) *Mathematische Bildung und neue Technologie*. Stuttgart/Leipzig: B.G.Teubner, 85-102.

BERND FICHTNER

REFLECTIVE LEARNING

Problems and Questions Concerning a Current Contextualization of the Vygotskian Approach[1]

Abstract. Both the works of art and the theoretical concepts are forms of knowledge based upon the hypothetical nature of our knowledge about reality. For reflective learning, the importance of works of art and theoretical concepts lies in the fact that they are spaces for development of thinking; they are never objects, results or drilled routines, methods and techniques. They are spaces for development of thinking in a special way: the subject will be able to think himself or herself. Drawing on an empirical research project, we shall inquire into following question: In what way is the sphere of the hypothetical or, the thinkable or the possible a space in which the human being can unfold his existence as a free and active being, potentially infinitely capable of development?

Key words: Activity theory of learning, self-referentiality, work of art

To focus our theme we would like to describe a scene from Lucchino Visconti's marvelous film "Bellissima:"

Anna Magnani sits in her ghetto flat with her view fixed on the shabby screen of an open-air cinema opposite for which she cannot even afford the admission. She watches a scene from "Red River" by Howard Hawks. To the reproaches of her husband, who is only interested in his every day affairs, she answers: "Oh Spartaco" – what a name for a chronically unhappy proletarian always sitting around in his undershirt! – "Oh Spartaco, allow me my dreams."

Certainly, this film is about nothing if not about the destruction of her dreams, although at the same time it deals with the preservation of dignity and with the love experienced by the dreamers. Perhaps it is exactly the artistic quality of the film that allows illusion and disillusionment to coincide in such a way that the human being is saved with respect to both his body and his mind.

The film within the film here is no simple citation; it is rather a key which opens the film itself as a complex system of self-referentiality: In the relationship between the film in the film and the main plot of the film the process of the destruction of the dreams of the principal character is becoming the central theme. The relationship between the film in the film and Visconti's film becomes a means by which the latter is making itself a subject of discussion as a film. Only for the spectator can this self-referentiality become a means of reflection with which he refers Visconti's film to himself.

We now want to use this scene taken from Visconti's film to consult different aspects and dimensions of self-refereniality. We narrow this questioning down to the

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 179 – 190.

content-part and the subject-part of reflective learning. We hope it will thus become clearer what "reflective learning" means.

We will begin with the following historical example: Wilhelm von Humboldt was arguably the first person to introduce a concept of reflective learning about 200 years ago. He did this within the context of his practical administrative work related to education. Our second step will be to consider how and why "self-referentiality" became a fundamental concept in Vygotskij's approach within the political context of the formation of a new society. The third step will be to criticize certain tendencies of the current Vygotskij fashion which is in the process of forfeiting the political core of the cultural-historical paradigm as a science of subjectivity. Deleuze's and Guattari's concept of *désir* enables our critique to regain a conception of the individual as the social subject of his life. Finally, we will return to Visconti's film and outline issues and aspects of a concept of reflective learning in reference to the reflective potential of art.

1. A HISTORICAL PROLOGUE: "LEARNING HOW TO LEARN" IN HUMBOLDT

In 1806 and 1807 Napoleon's troops inflicted a crushing defeat on the Prussian army in the battles of Jena and Auerstaedt. The entire state of Prussia collapsed. This catastrophe illuminated the extensive backwardness of this society on economic, technological, and political/cultural levels.

At the same time, this catastrophe was both the context and the impetus for the Prussian reforms initiated by Stein and Hardenberg. These reforms supplanted the traditional feudal society and as "reforms from above" were geared toward something new that had not existed in Prussia beforehand.

As a part of this movement, the educational reforms aimed at developing an entirely new type of school. The concept of "general education" (*Allgemein-Bildung*) functioned as a political strategy in the development of a general public school for all children. In a politically decisive administrative position, Humboldt organized this educational reform around 200 years ago.

In an extremely concise and precise manner he worked out a new conception of the contents of instruction, a new conception of learning itself, and of the connection between the two.

The contents were limited to instruction in *language* and *mathematics.* "Empirical and historical" subjects such as history, natural history, and geography were to be permitted as soon as they had become a matter of theoretical reflection – which was not the current state of affairs.

Instruction in language included those areas of philology which had already been theoretically and methodically clarified: philosophical grammar, Greek and Latin grammar. – The guiding principle was: "The form of the language as language" should become perceptible in instruction. According to Humboldt, this could be achieved "more easily with a dead language that causes astonishment because of its unfamiliarity than with the living mother tongue."

Instruction in mathematics was to take place in the form of a mathematics characterized by exact logical deductions as taught by Euclid, Lorenz, or according to some other precise conception of mathematics.

The contents of instruction were no longer "objects" in the treatment of which useful skills and abilities were to be learnt as according to the pedagogy of the Enlightenment. Here for the first time, a *theoretical conception of knowledge* displaced knowledge in an immediate practical sense. Instead of being oriented to the "needs of daily life" – as Humboldt described the immediate and pragmatic relation to society – an orientation was established towards knowledge on the highest level of a theoretical generalization. Astonishingly, at the same time this caused a radical focus on the individual, more precisely, on that activity which allows him to realize himself as the subject of his learning.

Humboldt expressed this in the following manner:

> With reference to the contents of instruction, from which all original creative work must always follow, the young person should be made capable of already actually beginning to compile the subject matter to a certain extent and to a further extent of accumulating it as he pleases in the future and of *developing his intellectual-mechanical powers. Thus, he is preoccupied in a twofold manner: with learning, but also with learning how to learn.* (1809, 169-170, my italics – B. F.)

Within the scope of the pedagogy of the Enlightenment, "mechanical skills" were developed – particularly with regard to the technical handling of articles for work, their material prerequisites and means. This accounted, for example, for a large part of the instruction that took place in the industrial schools.

Humboldt's suggestion signalized a fundamental change. Instead of a direct adoption of articles, substances, and knowledge as a finished product, the activity of learning itself became the focus, but not simply as some sort of automatism, activeness, or action.

Here, the characterization of learning as *simultaneously* being an orientation towards the content "from which all original creative work must always follow" and an orientation towards "learning how to learn" as a conscious focus on the learning process itself seems to be of primary importance. *For only in this simultaneous orientation does a simple reproduction of knowledge become replaced by a self-active production of knowledge as a subjective constitution, by learning as learning activity.*

This formulates a conception that deals with the development of individuality by means of acquisition of and access to knowledge at the highest level of its generalization. That is, from a radical, one-sided position general education (*Allgemeine Bildung*) is determined as the sole purpose of instruction (Humboldt maintained, "*Every carpenter should be required to learn Greek*").

There are two dimensions to the solution of the problem of generalization:

- Knowledge at the highest level of generalization is connected to the logic of the process of acquisition itself. – Learning confronts itself as learning how to learn.
- This necessarily requires a generalization at the social level: compulsory public schooling for all pupils.

So much in the way of an outline of Humboldt's concept. The actual implementation of this conception was a failure – not because of its radical, utopian perspective, but, rather, due to the contradictions of bourgeois class society, which developed very rapidly during Prussia's industrialization (Fichtner 1996, 174-194).

2. VYGOTSKIJ: MASTERY OF ONE'S OWN BEHAVIOR AS REVERSED ACTION AND SELF-REFERENTIALITY

Analogous to Humboldt's situation, we find a similarly dramatic socio-political context for the development of the paradigm associated with the cultural-historical school.

We comprehend the cultural-historical school from its historical context as an attempt within the humanities to define the subject in a new way under revolutionary conditions.

This new characterization became necessary as a result of the historic radical change and its social conflict-related pressures. Although the political and social environment of this period was shaped by the upheaval of an entire society and this was considered to be an historic act of self-constitution by a social subject of history, it became evident early on that such catchwords as "re-molding human beings" or "creation of new man" included determinist elements. Above all, it became evident that simply confronting people with objective necessities was not sufficient to change their consciousness.

We consider the research undertaken by the cultural-historical school to be an attempt to overcome both determinism and voluntarism in the formulation of the political aims of this social process of radical change. Categorically, Vygotskij emphasized the fact that the human individual as a subject can be reduced neither to nature nor to society.

Vygotskij's attempt to establish a science of subjectivity was based on a philosophical and methodological premise that could only be formulated in a negative way: as the overthrow of any type of dualism and, in particular, of the dualism between individual and society.

Within this context, the *Theses on Feuerbach* were of considerable significance, especially the third thesis, which stipulates that a change in reality necessarily includes a change in human beings themselves:

> The materialistic doctrine concerning the changing of (men's) circumstances and education forgets that circumstances must be changed by men and that the educator himself must be educated. This doctrine therefore has to divide society into two parts, one of which is superior to society. The coincidence of changing circumstances and human activity or self-change can be comprehended and rationally understood only as revolutionary practice. (Marx 1983, 156)

We consider this coincidence of changing circumstances and changing human activity or self-change to be a general framework of a revolutionary nature. Here, we do not wish to limit the concept of "revolutionary practice" to political activity in the strict sense of the word. We consider human activity in a very general way to be "revolutionary practice" whenever this connection between change in the world and self-change can be presumed.

Falk Seeger (1998) has demonstrated conclusively the central importance of Vygotskij's concept of "self-control or the mastery of one's own behavior" for the entirety of his work. Vygotskij provided a first systematic development of this concept in his study on "The History of the Development of Higher Mental Functions."

Three basic concepts are combined in this approach: "*the concept of higher mental function, the concept of cultural development of behavior, and the concept of mastery of behavior by internal processes.*" (Vygotsky 1997a, 7)

I need not go into detail on the famous metaphors Vygotskij uses in developing his approach: the example of tying a knot in a handkerchief in order to remember something and the no less famous image of Buridan's ass caught between two equally alluring bundles of hay.

As Vygotskij himself described his approach:

> In contrast to Lewin we attempt to provide for the concept of mastery of one's own behavior a completely clear and precisely determined content. We proceed from the fact that the processes of behavior represent the same kind of natural processes subject to the laws of nature as all other processes. Neither is man, subjecting processes of nature to his will and intervening in the course of these processes, an exception in his own behavior. But a basic and very important question arises: how does he represent the mastery of his own behavior to himself? ... We know that the basic law of behavior is the law of stimulus-response; for this reason, we cannot master our behavior in any other way except through appropriate stimulation. The key to mastery of behavior is mastery of stimuli. Thus, mastery of behavior is a mediated process that is always accomplished through certain auxiliary stimuli. (Vygotsky 1997a, 87)

Here we have an explicit formulation of the main issue: how do humans represent self-regulation to themselves?

Vygotskij provides certain clues about how to deal with this issue:

All clues refer to the social nature of this process in which human beings present self-regulation to themselves.

Vygotskij describes these "certain auxiliary stimuli" as "psychological instruments." He does not understand them as a mediator between subject and object. They are exclusively means of the subject's influence on itself. With its help the child organizes, controls and governs its behavior in very different situations. – It does not any longer react to an external stimulus – but creates, constructs its behavior.

Thus Vygotskij characterizes processes as "mediatory activity" and not as mediated. This ability of the human individual to produce his psychological processes as mediatory, mediating activity can for Vygotskij only be explained out of the subject-subject relationships.

Human self-regulation occurs in accordance with the so-called "general law of cultural development." This means that higher mental functions progress from the outside to the inside, from the social level to the individual level. "*Initially the sign is always a means of social connection, a means of affecting others, and only later does it become a means of affecting oneself.*"

Vygotskij concretized this in many examples. They are all to be found in the context of the question: How does its way of thinking change when the child learns to speak; and how does its way of speaking change when it learns to think?

An Example: What are numbers?

The arithmetical idea of numbers is a generalization of numerical attributes of things. Against this the algebraic idea is a generalization of the subjects operation on which the development of the graphical numerical idea of numbers is based on. It is a conscious generalization of the process of reasoning. On this basis the child is able to handle arithmetical ideas more freely.

Vygotskij demonstrates the same phenomenon using the example of grammatical structures:

> I loosen the knot. I do that consciously. However I cannot say how I did it. My conscious action does not come out to be an action which has become conscious, because my attention is focused on the act of loosening, but not on what I am doing. The consciousness always represents some part of reality. Object of my consciousness is the loosening of the knot, the knot and what happens to it; but not the actions I carry out loosening the knot, not what I am doing. This can in particular become the object of the consciousness, then this is the process of becoming conscious. Becoming conscious is the act of consciousness, whose object is the activity of consciousness itself. (Thinking and Speech; German edition: 1964, 168.)

At this point it would be appropriate to discuss the difference between "objective meaning" and objective "sense" in detail along the lines of the explication of this difference as found, above all, in Leont'ev (1981). Furthermore, the concept of "inner language" introduced by Vygotskij in an almost poetic form in the last chapter of his "Thinking and Speech" (Minck's retranslation of 1987) is of considerable significance in this context. Here we also find important reflections on the issue of how humans represent self-regulation to themselves.

We will cut our outline short here and summarize, even if somewhat too hastily: The paradigm of the cultural-historical school aims at establishing the humanities as a science of subjectivity. At the core is a conception of a human who as the subject of his learning process produces his or her uniqueness and unrepeatability not against the society he belongs to but, rather, by means of this society.

This science of subjectivity views humans as individual social beings who attain their autonomy to the extent that they do not simply observe social wealth in objects, but also have their own subjective means of the acquisition and expression of this wealth at their disposal.

This conception of such a science of subjectivity must became opposed to the system and became markedly contradictory to the political and cultural changes in social life as these began to be realized at the outset of the Stalinist era. The scientific category of "personal sense" facilitated, for example, radical criticism of social living conditions. The actual system of their "objective meanings" became less and less transformable into "personal sense." Stalinism placed the responsibility for this on the people themselves. The Paedology Decree issued on July 4th, 1936 made any further work by the cultural-historical school impossible.

3. A CRITIQUE OF CURRENT APPLICATIONS OF THE VYGOTSKIAN APPROACH AND THE CONCEPT OF DESIRE

The current interest in the cultural-historical school in Europe, Latin America, and the U. S. A. is astonishing and makes us somewhat suspicious.

Where does this widespread interest in the work of a Marxist scholar, and Communist of the former Soviet Union at American universities and Brazilian ones (which we are more familiar with) come from?

What difference exists between a subject who was determined to develop and form social life in the Soviet Union of the 1920s and the subject that is engaged in forming our present society?

What is the meaning of formation of identity in a society that is not reconciled to itself, is not identical to itself and as a capitalist society is currently caught up in dramatic changes within the context of globalization?

What does development of the subject mean in an antagonistic society that demands from the individual a balance of forces that is impossible in society itself?

How can a paradigm and its basic concepts that were aimed at making a practical contribution to the development of a society without class differences, without exploitation of humans by humans function in a society that is precisely based upon expansive implementation of capital?

The current reception and further development of the paradigm of the cultural-historical school makes no mention of our reality, its conflicts and contradictions and their significance for the development of subjectivity and identity of children and youths, for their learning and cultural appropriation.

The current reception and further development of the paradigm of the cultural-historical school is far too lacking in mediating factors; it has a peculiarly abstract tendency. The fundamental concepts and strategies are usually not related to our reality in any *concrete* manner.

Currently profound and comprehensive processes of an economic permeation of our society are taking place under the label of "globalization." Subsystems of our society such as public health services, law, sports and, not least, pedagogical institutions are forfeiting their relative autonomy to an ever greater extent. They are degenerating to auxiliary and reinforcing mechanisms of the market. Economy, that is, profit is rapidly and without any noticeable resistance becoming the measure of all things. (Chomsky 1999) All this represent factors of dramatic changes involved in how inner and outer coherence of our society is being produced. Within this context, practically all the traditional forms and functions of culture as a medium of the social lifeworld are in the process of dissolution.

In the present discussion, the fundamental concepts of the cultural-historical school are not related to this "disintegration of the social" and to this dissolution of traditional forms of lifeworld. Since this reality is not thematized, many of the concepts forfeit their methodological potential. They no longer allow deliberation on the fundamental and revolutionary connection between change in the world and self-change of the subjects, between the development of this society and the development of its individuals. Consequently, a greater portion of current cultural-historical research exhibits a pronounced orientation towards superficial craftsmanship, towards technical and methodical optimization of what is already available: for instance, the "Zone of Proximal Development" as a sort of "scaffolding" or "coaching" or as a method for implementing group work in existing forms of instruction.

Our own theoretical and practical work of the last few years has made it increasingly evident that it is practically impossible to use Vygotskij's approach for the education of persons who must adapt to the system of a society based on capitalist alienation and exploitation. This would encompass an alienation from the original intentions of Vygotskij's entire work. We see one possibility of regaining the connection between changes in the world and the self-change of individuals in the works of Deleuze and Guattari. Here, the concept of "desire" plays a significant role[2].

Guattari liberated this concept from the psychoanalytic perspective, which had bound it exclusively to the "libido" as the biological source of unconscious aspirations of humans. In this way, Guattari adopts a position with regard to psychoanalysis similar to that of Vygotskij:

> For me, desire encompasses all the forms of the will to live, to create, to love, to generate a different society, a different perception of the world, other values. Regardless of which dimension of desire one considers, it is never simply a general sort of energy, a vague function of chaos or disorder. (...) Desire is always a way of producing something. For this reason, I find it extremely important to dismantle the classical psychoanalytic conception. I am convinced that there is no biological-genetic process within the child that determines the aim of desire. However small a child may be, it lives out its relationship to the world and its relationships to others in an extremely creative and constructive manner. It is the schematizing of the child's semiotics by the school as a form of power that causes a type of schema of non-differentiation. (1986, 215 ff),

and, we would like to add, just as much so by all the other forms of power within the contexts in which the child lives: in the family, in the mass media, in the totalization of commodity-price relationships within social relationships.

We would like to pose the following general question:

What is the productive, critical-analytical potential of this concept of desire for a re-interpretation of the fundamental concepts of the cultural-historical paradigms?

Is it possible to realistically analyze the origin of the higher mental function in our social reality? If yes - how?

Is it possible to claim this also for the other concepts as for

- the concept of cultural development of behavior
- the concept of mastery of behavior by internal processes
- the concept of personality (Leont'ev)?

Only empirical and above all high-quality research can discover what productive, critical and analytical potential this conception of desire might have.

4. WORKS OF ART AS REVERSED ARTIFACTS AND ART'S POTENTIAL FOR REFLEXIVE LEARNING

At a central sequence in Visconti's film, we note how the protagonist watches a movie and becomes engrossed in her dreams, and how the film as a whole deals with the destruction of her illusions. The film within the film presents the simultaneity of illusion and disillusionment and, for the audience watching Visconti's film, it becomes a means of reflection on the nature of cinema in general and on one's own relation to this medium, etc. The film allows the observer his or her freedom of interpretation and provokes reflections and self-reflections.

This ability to present something and at the same time to thematize the presentation itself as a presentation seems to us to be an indication of the greatness of works of art. Such a presentation is always also the destruction of any unmediated perception of presentations – the destruction of presentation. Such presentations do not represent reality, the world, but, rather, reflect our activity in the world. Representation does not consist of the objects it designates.

This self-referentiality of art has enjoyed a practically inexhaustible variety of forms and possibilities in the course of its history. To make mention of just a few of these forms we would like to note some basic principles and mechanisms involved in perception as aptly described by Falk Seeger. He shows some pictures well known dealing with the "figure-ground-relationship" within the psychology of perception as for instance *Figure 1.*

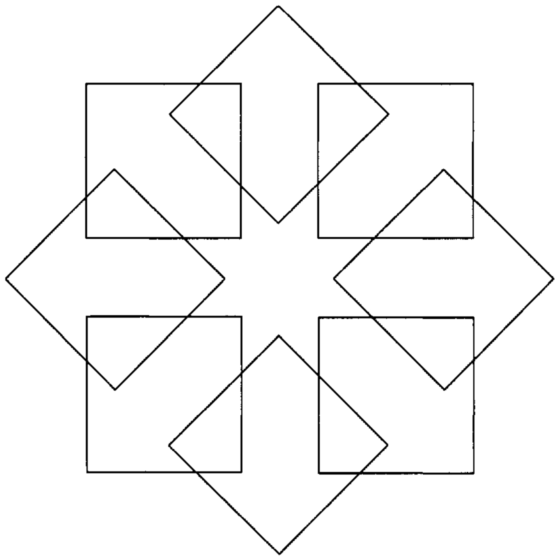


Figure1. A complicated self-referential picture: A Sufi mandala

> If the person, looking at one of those figures, focuses on one of its parts (...) the figure switches after a while: What had been at the forefront now seems to be at the back, and vice versa. (...) The usual semiotic function of the sign as pointing or referring to something else that is not given in the sign or picture is short-circuited because the image refers to itself. As a consequence ,the viewer is "left to her or his own devices," so to speak, and the normally unconscious processes of of perception are made conscious. (...) The key to an understanding of the psychological functioning of those pictures is Vygotsky's idea of "reverse action," of producing an artifact that is operating on the individual, not on the environment. (Seeger 1998, 330)

The here given picture is a "pure case" of self-referentially, it demonstrates the effect of self-referentially and nothing else. In work of arts self-referentially is an origin of an enormous explanatory potential.

Works of art are not primarily objects. They can be made to be such – as a commodity, as a fetish, as a status symbol, etc. Works of art are artifacts that mediate a relationship. What sort of a relationship? Each individual work thematizes art as a relationship. Art defines itself in relationship to what it is not. It is neither a product of labor nor is it nature, but, rather, something that does not exist in this empirical sense: namely, free form and "definite negation" of our accessible, consummate world.

In contrast to our relationship to nature or to work, our relationship to art is one in which we do not *objectify* ourselves and other things, but, rather, as Marx expressed it: one in which man "behaves towards himself as to a universal, and thus, free entity."

A work of art does not represent anything. Nothing else can be put in its place. By contrast, I can present the spoken word "tree" in sign language or in any alphabet used somewhere in the world, and it can always be related to a particular object.

Cezanne's pictures of Mont Saint Victoire do not represent a mountain with the purpose of illustration or documentation. They do not convey any meaning of or symbolize empirical reality – they do not refer to anything beyond themselves. These paintings are art in an immanent sense. They establish a relationship to art that continually asserts itself against their reification. These pictures refer only to themselves. But what differentiates them from the "pure cases" in which pictures thematize perception itself?

We would like to illustrate the particular potential of the self-referentiality of works of art with the following two examples:

Kafka's "The Castle:" The character of the surveyor K. is not based on some historical model and is not conceived of as representative for a figure of some particular social standing. He is a character who is anonymous, even with respect to himself, among other anonymous characters. No one knows anyone else, even though they all meet each other and speak to each other. In "The Castle" everything is narrated just as it happens to K., as he sees and understands it. There is no relationship whatsoever to any reality beyond the confines of the novel. Any concrete form of reality is totally absent. The castle does not symbolize some ruling power of which K. is a subject. And yet power and powerlessness and totalitarian coercion are present in every sentence. In the novels by Zola, the world of social misery, the exploitation of workers is realistically described from the *outside*. Zola's characters were present and still are in a great variety; but they always remain where they are in actuality. Kafka writes *within* the fictional character of K., which only originates while reading and makes reading a formidable experience: I do not read the words, but, rather, the words read me and they determine the rhythm and tone of my reading.

Velázquez's "Las Meninas:" The painter , Velasques himself pauses during his work. His gaze bores into the room where we, the observers, are to be found. The majority of the persons present in the scene on the left side of the easel also concentrate on the space in front of and outside the painting, which is actually our location. In the background a mirror hangs on the wall, but does not reflect the models, the royal couple, but, rather, a part of the picture that Velászquez is painting and that we can only see from behind. The painted frame almost collides aggressively with the surface of the canvas on which the entire scene is presented.

The idea of the reflecting mirror – one of the great themes of European painting – is reversed in this image: painting itself, and not reality, is reflected.

In the background, in the opening leading to another room, we see Jose Nieto de Velázquez, the queen's chamberlain. He is behind the scene in much the same way as we are before it. Velázquez is facing us; he enjoys the privilege of being able to see the picture. All of the lines in perspective converge at his hand. This hand grasps the perspective schema and manipulates the curtain.

As a whole, this work is a painting about painting: its theme is art as a relationship. Velázquez's replies to the question of what artistic representation is in the form of an aporia: He openly reveals his countenance, yet he conceals his work. "Las Meninas" is a picture open to an infinite variety of interpretations. Velázquez permits the observer to have freedom to interpret, but at the same time he forces him or her to meditate on the paradox of representation.

What, now, might be the potential of the self-referentiality of art for reflective learning? It is certainly not a one-to-one correspondence. This would mean to study the system comprised of works of art, their reception, and their effects with reference to the mechanisms involved and to construct a model of reflective learning from these mechanisms. But the explanation of works of art by the proper authorities seems to be more of a dead end.

Art is not didactic, art is not pedagogic, nor is it technical. Works of art are not instruments for practical problem-solving. Works of art mediate a relationship by providing space for the development of thought.

Fachbereich 2 Erziehungswissenschaft, Universität Siegen

NOTES

Translated by Thomas La Presti

[1] More than thirty-five years ago, Michael Otte introduced a group of students that I was a member of to the main figures of the cultural history school (Vygotskij, Leont'ev, Luria, and Dawydow) in a way especially typical of him – by explaining theories as perspectives. The criterion of their appropriation and implementation can be formulated by posing the question: Do they help us to make our experiences capable of development? This essay is an expression of my thanks to him for these insights.

[2] Collaborating in various projects with Maria Benites I'm indebted to her for this perspective on Deleuze and Guattari.

REFERENCES

Chomsky, N. (1999). *Profit over people. Neoliberalism and global order.* New York: Seven Stories Press.

Fichtner, B. (1996). *Lernen und Lerntätigkeit. Phylogenetische, ontogenetische und epistemologische Studien.* Marbug: BdWi-Verlag.

Guattari, F./Rolnik, S. (1986). *Micropolítica. Cartografias do desejo.* Petropolis: Vozes.

Humboldt, W. V. (1980). Werke in fünf Bänden. (Eds. A. Flitner/K. Giel) Vol IV. *Der Königsberger und der Litauische Schulplan.* (1809). Darmstadt: J. G. Cotta'sche Buchhandlung.

Jameson, Fr. (1991). *Postmodernism, or, The cultural logic of late capitalism.* New York: Duke University Press.

Leont'ev. A. N. (1981). *Problems of development of the mind.* Moscow: Progress Publishers.

Marx, K. (1983). Theses on Feuerbach. In: E. Kamenka (Ed.) *The portable Karl Marx.* New York: Penguin Books, 155-158.

Seeger, F. (1998). Representations in the mathematics classrooms. In F. Seeger/J. Voigt/U. Waschescio (Eds.), *The Culture of the Mathematics Classroom.* Cambridge: Cambridge University Press.

Vygotsky, L. S. (1987). Thinking and Speech. In R. W. Rieber, & A. S. Carton (Eds.) *The collected works of L. S. Vygotsky* (Vol. 1. Problems of general psychology, 38-285), New York: Plenum Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Vygotsky, L. S. (1997a). The history of development of higher mental functions. In R. W. Rieber, & A. S. Carton (Eds.) *The collected works of L. S. Vygotsky.* Vol.4. New York: Plenum Press.

Vygotsky, L. S. (1997b). The historical meaning of the crisis in psychology. A methodological investigation. In R. W. Rieber, & A. S. Carton (Eds.) *The collected works of L. S. Vygotsky.* Vol. 4. New York: Plenum Press.

RAINER BROMME

THINKING AND KNOWING ABOUT KNOWLEDGE

A Plea for and Critical Remarks on Psychological Research Programs on Epistemological Beliefs[1]

Abstract. In educational, developmental as well as cognitive psychology several approaches to measure epistemological beliefs have evolved. These approaches focus on learners' and teachers' beliefs about knowledge and epistemological issues. Such beliefs are important because of their impact on learning processes and learning results. Some of the most research approaches and their methods are outlined. Then some difficulties which the research is encountering are discussed. One of such difficulties concerns the issue of domain specificity of epistemological beliefs. This issue is one of the reasons for inconsistent empirical results. It will be suggested that some of these inconsistencies are caused by the epistemology implied in the research approach itself. In particular, the idea that epistemological beliefs can refer to knowledge as a social and cultural entity, is underrated. The paper concludes with the discussion of this argument and of it's consequences for further research.

Key words: epistemological belief, learning, measuring beliefs, dualism, relativism, certainty of knowledge, simplicity of knowledge, division of cognitive labor, domain dependency

1. PRELIMINARY REMARK

Learning a certain school subject always requires the development of an epistemological perspective about the content within the context of a certain domain of knowledge (e. g. mathematics). Teaching a school subject (and therefore selecting it from many other options which could be taught) is justified if that specific content is helpful for the development of a broader perspective on the domain of knowledge. The importance of an epistemological perspective in learning and teaching has always been emphasized in the work of M. Otte (e. g. 1994). It is a main theme in his critical thinking on educational issues. In his work this emphasis on epistemology in its own right is expounded mainly as a meta-perspective, in other words, as comments on the ways researchers, teacher educators and teachers think and argue about knowledge. And it is expounded often with a normative intention, driven by the contention that a more thorough reflection on epistemological issues would improve (mathematical) education. It is rarely based on empirical observations of epistemological thinking among students. But such empirical research approaches are also relevant for philosophical and historical reflection on educational issues. Since any 'having to do' requires the 'being able to do' and since any normative statement on educational issues should be based on a realistic examination of the personal and institutional conditions of acting according such normative statements, philosophical

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 191 – 201.

reflection needs such empirical analyses. Otherwise it would be out touch with its communicability with the social field it refers to.

On the other hand, however, empirical analyses are not beneficial unless they themselves are based on an appropriate epistemological conception. Only then is it possible in the event of contradictory results to decide whether they are merely caused by methodical problems or whether a revision of the theoretical approach is needed. (By the way: Such a self-referential use of a certain epistemological argument is inspired last not least by M. Otte's work, where such patterns of self referential argumentation can be found very often and where the importance of self-reference for the justification of mathematical knowledge has been thoroughly discussed).

2. EPISTEMOLOGICAL BELIEF: A VERY INTENSIVELY-INVESTIGATED OBJECT OF RESEARCH INCLUDING MANY OPEN QUESTIONS

Recently several empirically based approaches to epistemology and epistemological beliefs have evolved. They are put forward by psychologists (within the sub-domains of developmental, educational and cognitive psychology) and by researchers in science, language, history and mathematics education. These approaches are attempting to describe and measure learners' and teachers' beliefs about knowledge and epistemological issues. (Buehl, & Alexander 2001; Duell, & Schommer-Aikens 2001; Hofer, & Pintrich 1997; 2002 have given inspiring overviews of this research).

The development of beliefs about the ontological status and justification of 'knowledge,' knowing and (sometimes) about the acquisition of knowledge is the subject of these research approaches. Some of the researchers focus their studies on the beliefs about academic knowledge, e. g. mathematics (Schoenfeld 1992; de Corte, Op t' Eynde, & Verschaffel 2002), science (Lederman 1992), psychology (Hofer 2000). Others prefer to concentrate on the epistemology underlying the application of personal knowledge in connection with problems in everyday life. Here, general ideas about knowledge and its justification are at the forefront while assignment to a particular domain of school knowledge is not so important for these researchers (Kuhn 1991; King, & Kitchener 1994).

Some of the research approaches will be outlined below (section 3). The paper goes on to deal with the difficulties which the research is encountering at present. As an example of such difficulties the issue of domain specificity of epistemological beliefs will be discussed (section 4). This issue is one of the reasons for inconsistent empirical results. It will be suggested that some of these inconsistencies are caused by the epistemology implied in the research approach itself (section 5). In my view the research is based on an inadequate theory as to what the content of beliefs about epistemology and academic knowledge might be. In particular, the idea that epistemological beliefs can refer to knowledge as a social and cultural entity, is underrated. Put bluntly: the epistemology of the research on the development of epistemological beliefs might fall behind the development of the subjects of their research.

The paper concludes with the discussion of this argument and of it's strategical consequences for further research (section 6).

3. FROM THE BELIEF IN ABSOLUTE KNOWLEDGE TO THE CONTEXTUALIZATION AND DEVELOPMENT OF A PERSONAL POINT OF VIEW: PERRY'S DEVELOPMENT MODEL AND SUBSEQUENT STUDIES

Many recent approaches are rooted in Perry's seminal work on the intellectual and moral development of college students (Perry 1970, for review about his work, see: Moore 2001, Hofer, & Pintrich 1997). Perry, a professor at Harvard University, examined several cohorts of college students by a longitudinal method over a period of 4 years per group. In nondirective interviews he asked students about important events in their (intellectual) development at the college. The results were integrated into a development model of beliefs about views on the making of meaning. Although Perry does not mention any stages, his ideas of development are clearly influenced by Piaget's stage theory and also have a strong affinity with Kohlberg's theory of moral development. Thus he views intellectual development as a result of dealing with new experiences, while at each developmental step a state of equilibrium has to be achieved.

The original model consisted of 9 phases which, however, in subsequent studies were reduced to four stages: Dualism, Multiplicity, Contextual Relativism, and Commitment within Relativism. (The following short description of these stages is based on Hofer, & Pintrich (1997), the references are from page 71 of this text).

Dualism is ... "characterized by a dualistic, absolutist, right, and wrong view of the learner."

It is a view about knowledge where it is not difficult to figure out if elements of knowledge are true – it is simply a case of consulting the authorities.

Multiplicity ... "represents the beginning of the recognition of diversity and uncertainity. Authorities which disagree haven't yet found the right answer, but truth is still knowable."

Later this includes also the assumptions that there may be issues where different opinions are possible and that it is permissible to have these opinions.

Contextual Relativism encompasses the acknowledgment that knowing always includes a personal perspective and that it is based on one's own intellectual activities within a certain context. "At Position 6 (the second sub-stage of this stage, R. B.) individuals perceive knowledge as relative, contingent, and contextual and begin to realize the need to choose and affirm one's own commitments."

Commitment within relativism refers not primarily to knowledge and its justification but primarily to a certain approach to finding a personal position within a context of relativism. Nevertheless it is again the idea of multiplicity, i. e. coping with a multiplicity of values and developing a personally-justified position within such an environment.

The basic idea of the development model refers to the tension between a simple, even naïve, belief in a one-dimensional and external explanation of knowledge and meaning and a relativistic view of knowledge and meaning, in which the individual is not only a passive recipient but also an active constructor of his knowledge. Perry's model and results inevitably turn out, on detailed inspection, to be more complex than can be expressed by this dimension. However, the basic idea can be recognized more easily by taking into account the fact that the starting point for Perry's studies was the research carried out by Adorno, Frenkel-Brunswick, Levinson and Sanford (1950) on the authoritarian personality. This research was already based on a similar idea about closed vs. open-mindedness. Contrary to Adorno et al. Perry and subsequent research on epistemological beliefs see these opinions more as a result of a developmental and learning process than a personality trait.

Perry's development model has inspired a number of subsequent studies which share his idea of cognitive development from 'absolute knowing' to 'contextual knowing.' In different studies the poles of this dimension have different names. Here I will use these, in my view very appropriate, terms for the two poles of development suggested by Baxter Magolda (1987). Some examples of these subsequent studies are mentioned here in brief.

King and Kitchener (1994) have described in a series of studies of 'the ways that people understand the process of knowing' (1994, 13) by asking their research subjects to respond to ill-structured problems. These are problems for which there are no clear-cut, unambiguous solutions, e. g. questions pertaining to the objectivity of media, assumptions about the construction of the Egyptian pyramids or chemical additives in food. The research subjects had to verbalize their thoughts on these problems. The analysis of the interviews was based on categories which were essentially concerned with the tension between absolute and relativistic beliefs. The results confirmed and extended Perry's stage model. King and Kitchener asked over 1700 students and students of varying ages. Thus they were able to demonstrate that the developmental stages do in fact correlate with the age of people and also depend on their degree of education: Older and better educated students and young adults, respectively, more often gave explanations which were classified as quasi-reflective or reflective.

However, subsequent studies also led to a fundamental broadening of the term 'epistemological beliefs.' An example of this are the studies of Schommer (1990, 1994; and also based on her work: Jehng, Johnson, & Anderson 1993). Schommer devised a questionnaire in which subjects had to state their degree of disagreement/agreement to propositions, for example regarding the certainty of knowledge, e. g. 'The only thing that is certain is uncertainty itself,' 'Scientists can ultimately get to the truth.' Since the aim was to obtain empirical evidence for a relationship between epistemological beliefs and learning in school, questions were included which did not directly address the nature of knowledge but for example preferences in study strategies 'When I study I look for single facts' or questions with regard to whether learning success depends on ability or on step by step efforts. Schommer's questionnaire made it possible to study a relatively large sample more economically. It also made possible an empirical examination of the question whether epistemological beliefs vary over different dimensions. This concerns for example the question whether beliefs regarding the *certainty of academic knowledge* co-vary systematically with beliefs regarding the *simplicity of knowledge*, or whether such beliefs develop independently of one another. In a series of studies in which the questionnaire was used (repeatedly revised and supplemented), four factors were found by Schommer (1994; Schommer, Calvert, Gariglietti, & Bajaj 1997) which were independent of one another:

(1) *Simplicity of knowledge*, i. e. knowledge as an entity of unrelated items vs. knowledge as a system of propositions,

(2) *the certainty of knowledge*,

(3) *quick learning* i. e. the ease of knowledge acquisition and

(4) *innate ability*, i. e. degree of individual control over knowledge acquisition.

In addition, Schommer – and others – could establish empircally relationships between some of the just mentioned epistemological beliefs and cognitive processes during learning, for example when learning from texts (Kardash, & Howell 2000).

However, the factor structure as reported by Schommer did not prove very stable, i. e. not all the factors could be replicated in other studies. This was caused, among other things, by questions which were not clearly formulated. In addition, the propositions which have be judged refer partly to the personal knowledge of the subject and partly to the presumed knowledge of experts. Finally, it was heavily criticized that many of M. Schommer's questions and two of the four factors which she had been able to find do not really refer to epistemological issues (Hofer, & Pintrich 1997). This conceptual objection is legitimate and is not seriously denied by the author (Schommer-Aikins 2002). The selection of the items was eclectic in so far as the main objective was – as already mentioned above – to include as many beliefs as possible which could have an effect on learning. Nevertheless, the core theme underlying the issue of epistemological beliefs, as outlined above, can be discerned: in my opinion, each dimension contains the idea of 'absolute knowing' vs. 'contextual knowing.' (Note: This is an analytical conclusion about the semantics of the propositions which make up the individual factors, not an assumption that it would be possible to establish empirically a sort of 'g(eneral)-factor' behind the four dimensions described above).

4. ARE EPISTEMOLOGICAL BELIEFS SPECIFIC FOR THE KNOWLEDGE DOMAIN TO WHICH THEY REFER?

Other researchers doubt whether it is at all possible to define and to study empirically epistemological beliefs independently of the domains to which the knowledge refers. Given the great differences between academic disciplines (in terms of their concepts, methods, languages) it is very plausible that students ideas about knowledge might reflect such differences. Actually this has been demonstrated in an interesting study undertaken by Stodolsky, Salk, and Glaessner (1991). They asked fifth class students about special features of the subjects *mathematics* and *social studies*. In order to make students of this age group think about the nature of these subjects an interesting scenario was developed. Students were asked to imagine a popular extraterrestrial movie character (E.T.) visiting their school and needing some explanations about what went on during school lessons. Stimulated by this external perspective the students reported characteristic differences between the learning and teaching in these two subjects (for example: teaching and learning with algorithms is typical for mathematics). Although the questions did not refer directly to epistemological beliefs, characterizations of the two domains were essentially found to refer to epistemological issues.

Schommer and Walker (1995) have reacted to the criticism on their assumption about domain independent epistemological beliefs. They used one of Schommers's questionnaires originally used without any reference to a certain knowledge domain. Now a student sample was asked to refer to a specific domain (e. g. mathematics) when answering. Then this sample was asked to work on the questionnaire again, now with reference to social sciences. Based on correlations between the answers for the two domains and based on comparison with a control group the authors stated a moderate domain independency of the epistemological belief. However, as explained above, two of the four factors found by Schommer did not concern epistemology beliefs but learning strategies. It is interesting that the specifically epistemological beliefs about *certainty* and *simplicity* of knowledge showed a lesser degree of domain independency than the beliefs about learning strategies. Schommer-Aikins (2002) has, therefore, moderated her claim of domain independence of epistemological beliefs.

Buehl, Alexander, and Murphy (2002) also examined the domain specificity of epistemological beliefs. They started with criticism of Schommer's approach which, in my opinion, reveals a very fundamental problem. They argued that a questionnaire which is based on the assumption of domain independency could not simply be applied to different domains. It has to be assured that all questions are applying equally meaningfully in the different domains. They have therefore devised an instrument which contains only items that can be answered with respect to both mathematics and history. Buehl, Alexander, and Murphy (2002) report that they had, for example, to leave out questions on the perceived meaning of formulae in mathematics as they were unable to formulate corresponding questions for history.

However, such a strategy results in risking the loss of those items which are particularly relevant for eliciting domain specific beliefs. This emerging problem of mixing things which are incompatible can in my opinion only be solved if, on the basis of the specific research question a theoretical decision is made as to what degree incommensurability between domains has to be accepted. Thus, in the light of this, it is not surprising that Buehl, Alexander, and Murphy (2002) only found domain specific differences for some beliefs, e. g. for beliefs concerning the relationship between mathematics and history, respectively, and other subjects (mathematics seems to be seen as more strongly linked with other knowledge domains than history). Clearer empirical evidence for domain specificity of epistemological beliefs was found by Hofer (2000), who examined the beliefs of college beginners with respect to psychology and science.

5. WHAT CAUSES INCONSISTENT RESULTS? THE NECESSITY FOR IMPROVING RESEARCH ASSUMPTIONS ABOUT WHAT THE 'CONTENT' OF EPISTEMOLOGICAL BELIEFS COMPRISES

As has already become clear, there are considerable inconsistencies in the findings (see also Pintrich (2002), Buehl, & Alexander (2001), Schraw (2001) for critical reviews of such inconsistencies). Below is a summary of some of these inconsistencies:

Complexity of epistemological beliefs. The number of dimensions that can be attributed to epistemological beliefs is not at all clear. In a critical review of the present research position Pintrich (2002) states: There is more than one dimension, but less than ten. For example, Hofer (2002) as well as Quin and Alverman (1995) report that '*Certainty of knowledge*' and '*Simplicity of knowledge*' did not occur as separate dimensions, as Schommer had found, but were mixed. This can be the result of a slightly different choice of items (a method-oriented explanation of inconsistencies) or it points to the fact that *certainty* and *simplicity* are two aspects of more complex beliefs about knowledge (a theory-oriented explanation of inconsistencies).

Domain dependency. It can be safely assumed that epistemological beliefs are influenced strongly by the knowledge domains to which the questions in the inventories refer. There are, however, up to a certain degree, domain independent epistemological beliefs. This varies according to the researcher and, consequently, the research instrument.

Age level: Chandler, Hallett and Sokol (2002) have demonstrated that the development from the pole 'absolute knowing' to the pole 'contextual knowing,' as already described by Perry, can be found in all studies regardless of which age levels are examined. However, it is not possible that the same development takes places exactly in the phases a researcher happens to select.

Of course, many concrete causes of such inconsistencies can be found in the methods used. In the recent debate on epistemological beliefs there are much more suggestions about methods than about the actual content of epistemological beliefs. However, as Chandler et al. (2002) argue convincingly, a solution to the problem of inconsistencies cannot be expected from an improvement in methods alone. In the above mentioned critical reviews it has, therefore, been demanded that besides improving research methods, it is necessary to further develop the theoretical assumptions about the ontology of epistemological beliefs

I would therefore like to ask whether the ideas about what the content of epistemological beliefs under discussion by most of the authors in this research area at present do justice to what people think about knowledge and recognition. In other words, I doubt that the 'ontology' of epistemological beliefs which is presupposed in recent research really 'fits' the 'ontology' of the research subjects. It was illustrated at the outset that present research – slightly simplified – focuses on the basic idea of 'absolute truth vs. contextual knowing.' This understanding of epistemology is however too narrow. It is based on a psychologically inspired idea of epistemology as a theory of individual knowledge attainment. (The relationship between Perry and Piaget has already been pointed out). In contrast, modern views on epistemology emphasize the social nature of the generation and justification of knowledge (Fleck 1979; Latour, & Woolgar 1979; Otte 1994). Here it is clear that the generation and use of knowledge is a process which is socially distributed within a cultural context. There is a lot of division of cognitive labor when it comes to the generation, justification and use of knowledge. And it is clear that knowledge is represented not only in individual minds, but also in books, in the Internet, and in social and in technical artefacts. Accordingly, epistemology (note that this notion refers to epistemology as a philosophical discipline) describes processes and problems of *individual* knowledge attainment only in a limited number of cases.

Two examples will demonstrate why division of cognitive labor and the social nature of knowledge are important for the empirical analysis of epistemological beliefs: One of the factors which has been found in many studies of epistemological beliefs concerns the 'sources of knowledge.' It is elicited by the item 'Sometimes you just have to accept answers from the experts even if you don't understand them' and 'If you read something in a textbook for this subject, you can be sure that it's true' (Hofer 2000, 399). Agreement with these items is taken as an indicator for a belief in authorities and little development towards the contextualized pole of knowledge.

This may be the case in certain contexts. But it might also indicate an acknowledgment of the division of cognitive labor. It is unnecessary and impossible in many contexts for individuals to actively gain, structure und justify knowledge themselves, i. e. to think in the way it is described by the positive pole of Perry's sketched development. Instead a stable judgment is needed as to when in a division of labor one can rely on other people and when one has to make a judgment of one's own.

However, contemporary research on epistemological beliefs underestimates the division of cognitive labor. Maybe the research subjects who have to answer the questions about knowledge take the division of labor into account. Consequently this ambiguity of 'expert truth' reveals itself in inconsistent results. It is possible that subjects prefer such absolute truths although they would be prepared to develop a critical position towards the opinions of *certain* experts and *specific* questions.

A different but related example concerns the issue of 'absolute' vs. 'relative,' i. e. 'personal view dependent truth.' The belief that there are absolute, externally proven truths which cannot be personally verified is not always an indication for naïve epistemology (Hammer, & Elby 2002). In fact, it makes sense to assume that the earth is round and not a disc, that Newton's laws are always valid in macrophysics, etc. It is the belief about the social distribution of knowledge (Who states the proposition, when and what for?) which decide if and when a statement should be regarded as true or false.

6. A SUGGESTION FOR A RESEARCH STRATEGY: SEARCHING FOR THE 'DIVISION OF COGNITIVE LABOR'

Of course, it is an open empirical question which age groups and with what education know about the division of cognitive labor. To this extent my criticism implies also a research program. There is already empirical evidence in developmental psychology (Lutz, & Keil 2002), in concept research (Malt 1994) as well as in expert research (Bromme, Rambow, & Nückles 2001) that conceptual knowledge consists not only of knowing about the meaning of concepts but also of assumptions about a division of cognitive labor. Individuals know about other people possibly having different, better, and in other contexts more useful knowledge about concepts than the individual has. Nevertheless, these concepts can be applied in one's own thinking and communication with other people (Bromme 2000). If some of the research subjects know about this division of labor and take this into account in some of their answers one needn't be surprised at the above mentioned inconsistencies of results on epistemological beliefs.

This criticism regarding the implicit epistemology of epistemology research (i. e. its own notion of epistemology) implies not only that the division of cognitive labor should be added as a possible additional variable. I do not suggest a simple addition of further ontological elements. A merely descriptive, so to say 'bottom up' approach to the empirical reconstruction of epistemological knowledge will not be successful. Adding further items in questionnaires will presumably not result in the reduction of inconsistencies. It has to be accompanied by a contextualization of the use of epistemological beliefs. It might be helpful to start with the question: why and with what aim epistemological beliefs emerge and for what are they needed? The question also arises as to what the problems are which people solve by having certain epistemological beliefs. When, i. e. with what actions and in which contexts do problems arise for which the availability of a certain epistemological belief is relevant?

A final example will help to demonstrate this: If a student has to choose between conflicting factual statements made by another student or by a teacher, it makes more sense for him to believe in the teacher's authority. A person who is considering renting a flat which is near a mobile phone mast has to decide whether to believe the previous tenants who complain about headaches, or physicists who claim that there is no health risk. In this case it might make sense not to believe in the physicists' authority. In either case the epistemological belief has nothing to do with knowledge as such, but concerns the relationship between the statement 'proposition X is true' and the social context of who, why and where has said this, and of whom, why and where the truth of this proposition is important.

It therefore makes sense to distinguish between the contexts in which epistemological beliefs are applied. So school and university can be seen as such a context. A personal search for natural scientific or medical knowledge, e. g. on the internet, is a different context. However, since people act in several such contexts at the same time, it appears essential to apply constraints to any proposed extension of the ontology of potential epistemological beliefs by specifying the context in which and aim for which the knowledge is needed about which people have epistemological beliefs.

Fachbereich Psychologie, Universität Münster

NOTES

[1] Thanks to Elmar Stahl for critical remarks and to Ingrid Speight for native speaker advice.

REFERENCES

Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York: Harper & Row.

Baxter Magolda, M. B. (1987). The affective dimension of learning: Faculty-student relationships that enhance intellectual development. *College Student Journal* 21, 46-58.

Bromme, R. (2000). Beyond one's own perspective: The psychology of cognitive interdisciplinarity. In P. Weingart & N. Stehr (Eds.), *Practicing interdisciplinarity*. Toronto: Toronto University Press, 116-133.

Bromme, R., Rambow, R., & Nueckles, M. (2001). Expertise and estimating what other people know: The influence of professional experience and type of knowledge. *Journal of Experimental Psychology: Applied 7*.4, 317-330.

Buehl, M. M., & Alexander, P. A. (2001). Beliefs about academic knowledge. *Educational Psychology Review 13*.4, 385-418.

Buehl, M. M., Alexander, P. A., & Murphy, P. K. (2002). Beliefs about schooled knowledge: Domain specific or domain general? *Contemporary Educational Psychology* 27, 415-449.

Chandler, M. J., Hallett, D., & Sokol, B. W. (2002). Competing claims about competing knowledge claims. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 145-168.

de Corte, E., Op t' Eynde, P., & Verschaffel, L. (2002). "Knowing what to belive:" The relevance of students' mathematical beliefs for mathematical education. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 297-320.

Duell, O. K., & Schommer-Aikins, M. (2001). Measures of people's beliefs about knowledge and learning. *Educational Psychology Review 13*.4, 419-449.

Elby, A., & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education 85,* 554-567.

Fleck, L. (1979 <1935>). *Genesis and development of a scientific fact*. Chicago: London.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 169-190.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology* 25, 378-405.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research 67*.1, 88-140.

Hofer, B. K., & Pintrich, P. R. (Eds.) (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum.

Jehng, J.-C. J., Johnson, S. D., & Anderson, R. C. (1993). Schooling and students' epistemological beliefs about learning. *Contemporary Educational Psychology 18,* 23-35.

Kardash, C. M., & Howell, K. L. (2000). Effects of epistemological beliefs and topic-specific beliefs on undergraduates' cognitive and strategic processing of dual-positional text. *Journal of Educational Psychology 92*.3, 524-535.

King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgement: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco, CA: Jossey-Bass.

King, P. M., & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 37-62.

Kuhn, D. (1991). *Skills of argument*. Cambridge, England: Cambridge University Press.

Latour, B., & Woolgar, S. (1979). *Laboratory Life*. London.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research on Science Teaching* 29, 331-359.

Lutz, D. J. & Keil, F. (2002). Early understanding of the division of cognitive labour. *Child Development, 73*.4, 1073-1084.

Malt, B. (1994). Water is not H_2O. *Cognitive Psychology* 27, 41-70.

Moore, W. S. (2002). Understanding learning in a postmodern world: Reconsidering the Perry scheme of ethical and intellectual development. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 17-36.

Otte, M. (1994). *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik.* Frankfurt: Suhrkamp

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme.* New York, NY: Holt, Rinehart, & Winston.

Pintrich, P. R. (2002). Future challenges and directions for theory and research on personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 389-414.

Qian, G., & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology 87*.2, 282-292.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. New York: Macmilla,n, 334-370.

Schommer, M. (1990). Effects on beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology 82*.3, 498-304.

Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P. A. Alexander (Eds.), *Beliefs about text and about text instruction*. Hillsdale, NJ: Lawrence Erlbaum, 25-39.

Schommer-Aikins, M. (2002). An evolving theoretical framework for an epistemological belief system. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, N. J.: Lawrence Erlbaum, 103-118.

Schommer, M., Calvert, C., Gariglietti, G., & Bajaj, A. (1997). The development of epistemological beliefs among scondary students: A longitudinal study. *Journal of Educational Psychology 89*.1, 37-40.

Schommer, M., & Walker, K. (1995). Are epistemological beliefs similar across domains? *Journal of Educational Psychology 87*.3, 424-432.

Schraw, G. (2001). Current themes and future directions in epistemological research: A commentary. *Educational Psychology Review 13*.4, 451-464.

THOMAS MIES

THE COGNITIVE UNCONSCIOUS

Recalling the History of the Concept and the Problem

Abstract: For a long time the cognitive unconscious was a marginal subject in the epistemology and psychology of the 20th century. In the last decades, however, this situation has dramatically changed under the influence of cognitive psychology and cognitive science. As a contribution to a better understanding of this change, its reasons, and its perspectives, this article tries to recall the history of the cognitive unconscious as a philosophical and psychological concept that does not begin with the rise of cognitive psychology and cognitive science. It links the concept to the philosphical criticism of Descartes and Kant as the most eminent proponents of a philosophy of consciousness. In particular, it links the concept to American pragmatism. By anchoring consciousness in practical and sign-mediated intersubjectivity, Peirce, Dewey, and Mead, at the same time laid the foundations for a pioneering conception of the cognitive unconscious that requires further elaboration and remains a stimulating challenge for current philosophical and psychological research.

Key words: abduction, aesthetics, cognitive science, habit, mind, philosophy of consciousness, pragmatism, psychoanalysis, unconscious, unconscious inference

> I believe that much of early Freudian theory was upside down. At that time many thinkers regarded conscious reason as normal and self-explanatory, while the unconscious was regarded as mysterious, needing proof, and needing explanations. Repression was the explanation, and the unconscious was filled with thoughts which could have been conscious but which repression and dream work has distorted. Today we think of consciousness as mysterious, and of the computational methods of the unconscious, e. g., primary process, as continually active, necessary, and all-embracing. (Bateson 1975, 135 f.)
>
> It is because agents never know completely what they are doing that what they do has more sense than they know (Bourdieu 1990, 69).

INTRODUCTION

The fact that unconscious mental processes have a role in learning and cognizing would seem to be a truism from a present-day point of view. It is all the more surprising then to note that scientific discussion of this "truism" has long been the privilege of eminent individualists, being left without much resonance during the 20th century, a century deemed, according to widespread prejudice, to have been that of the discovery of the unconscious. In truth, discussion of the cognitive unconscious was confined to the margins of scientific debate, to footnotes, and to anecdotic occurrence.

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 203 – 214.

Undoubtedly, psychoanalysis dominated the scientific discourse on the unconscious quite decidedly from its very onset until the 1970s. It is thus suggestive to link the fact stated above to certain basic assumptions of the Freudian conception of the unconscious: This conception can be conceived of as cognitivist inasmuch it refers primarily to a certain, albeit quite specific, class of "notions"; conceiving these notions and the regularities of their linkages, however, not only as a primitive cognition, but rather as a depraved one that is in sharp contrast to productive learning and reasoning. "Primary process" and "secondary process" are mutually exclusive. This is also connected to the fact that psychoanalysis, despite its important contribution to the semiotic turn in philosophy and in the humanities, was long dominated by a concept of the symbol that isolated it from the scientific and philosophical developments initiated by the semiotic turn (cf. Lorenzer 1972). The belief that the unconscious is creative shared by many psychoanalysts – including Freud himself, of course – cannot be founded adequately under these premises of reasoning, and it is no accident, albeit involving considerable inconsistencies in this field as well, that it is being demonstrated almost exclusively in studying artistic production. A conception of the unconscious that declares itself incompetent to such an extent regarding questions of learning psychology and epistemology seems to confirm all reservations and biases that consider any application of psychological concepts and insights to the epistemological field a path of error anyway.

During recent decades, however, there has been a far-reaching turn in evaluating the cognitive unconscious. A marginal theme has evolved into a legitimate and important object of research. The most spectacular aspect of this turn is undoubtedly the growing recognition of cognitive structures and processes that operate without consciousness. The rapid progress of cognitive psychology and cognitive research demonstrates the power of these structures and processes (see, e. g., Kihlstrom 1987; Perrig/Wippich/Perrig-Ciello 1993; Pfeifer/Scheier 1999). How far-reaching the turn is here may be illustrated by the following quote: "Paradoxically, it would seem as if the psychologists were able at present to make clearer statements about unconscious cognitions than about phenomena of consciousness. It is widely recognized in modern research that unconscious cognitions have a role in every kind of information processing, that is also in case of so-called higher forms of coping with the external and the internal world (like reasoning, imagining and remembering)" (Perrig/Wippich/Perrig-Ciello 1993, 218).

In psychoanalysis as well, the assumptions for discussing the topic of the cognitive unconscious have undergone a profound change. The notion that unconscious and conscious thinking are incompatible is being replaced by conceiving the primary and secondary process as a continuum in which the two are recognized to be functionally different forms of the cognitive acquisition of reality (see, e. g., Holt 1967; Matte-Blanco 1975; McKinnon 1979). The psychoanalytic concept of the symbol is liberated from its isolation with regard to the general development of the theories on signs, language, and symbols (cf., specially, Lorenzer 1972, 1977, 1981). This change does not stop at the core region of psychoanalytic practice and theory formation. Recent empirical studies on the process of psychoanalytic therapy no longer assume "that the patient's basic thrust is the wish to gratify unconscious infantile wishes," but rather start "from the radically different assumption ... that patients come to therapy with the conscious and unconscious desire to *master* early conflicts, traumas, and anxieties and with unconscious *plans* as to how this mastery can be achieved." (Eagle 1984, 96 ff.). The author just quoted comments this change of paradigm as follows: "The concept of unconscious plans ... might have seemed untenable some years ago to some sceptical critics. However, under the impact of developments and research in so-called cognitive science and cognitive theory, the idea of complex and purposive unconscious cognitive operations seems commonplace and entirely feasible." (loc. cit., 101). It is thus also not surprising that an increasingly intense cooperation between psychoanalysts and cognitive psychologists, respectively cognitive scientists has developed in the last two decades. (see, e. g., Moser/von Zeppelin 1991; Leuzinger-Bohleber/Schneider/Pfeifer 1992; Shevrin et al. 1996; Bucci 1997)

Further evidence for the turn just sketched may be found in the fact that the analytic philosophy of the mind, which, indeed, has been intimately connected with cognitive science as an interdisciplinary approach, has made energetic efforts during recent decades to integrate the Freudian theory of the unconscious, thus trying to understand irrationality within the context of the cognitive acquisition of the world (see Davidson 1982, 1985; Pears 1984; Cavell 1993). Guiding these efforts is the search for an alternative to Freud's radical way of opposing the pleasure principle and the reality principle: "If instead we assume the position ... first that passions (wishes, and desires) are fully mental and logically inextricable from belief, and second that mental content is constituted out of interrelations between organism and external world, organism and other organisms, then we will expect it to be the case that passions, desires, and interests are also in some general way adapted to external reality, no matter how much they stray here and there" (Cavell 1993, 158).

It is beyond any doubt that psychoanalysis has revolutionized the study of the unconscious theoretically, methodologically, and empirically. The history of science and philosophy, however, also shows that it would be completely erroneous to ascribe the discovery of the unconscious to Freud (see, e. g., Ellenberger 1970; Adler 1988; Lütkehaus 1995). Just as the philosophic and scientific discourse on the unconscious does not begin with Freud, the cognitive unconscious has by no means been scientifically thematized for the first time by cognitive psychology and cognitive science. In view of the revolution just sketched, an effort to recall the history of "the unconscious" might be useful that is not confined to Freudian thinking and to psychology. It might contribute toward a better understanding of the reasons that have led to the present upswing of the cognitive unconscious, and toward placing alternative modes of thinking at the disposal of the present scientific debate which have fallen into oblivion because of the course taken by the development of concepts and theories. What follows is intended to contribute to such an effort of recall. The history of classical psychoanalysis itself provides a negative example for the uses of such an effort of recall. Freud and his disciples, by believing that they could mostly ignore the history of the discovery of the unconscious, in particular its philosophical aspects, as a mere prelude, adopted without reflection decisions about direction that had been made long before them. These decisions drew insurmountable boundaries to any attempt at articulating what was new in their own approach (see Sulloway 1979; Mies/Brandes 1999; Mies/Scholz 2001).

THE CRITIQUE OF THE PHILOSOPHY OF CONSCIOUSNESS CHARLES SANDERS PEIRCE

Our confining the history of the concept here to Peirce and to American pragmatism is more or less arbitrary and not just due to the fact that a comprehensive history of the cognitive unconscious, which would have to begin with Leibniz at the latest, would go beyond the frame of this contribution.[1] The history of the concept refers – however contrarily – from its very outset to the development of the philosophy of consciousness as its determining context of reasoning. Descartes as its founder, and Kant as its most important reformer, thus do not just mark fundamental incisions in the history of the philosophy of consciousness; after them, a fundamental revolution also took place in thinking about the unconscious (see Mies/Brandes 1999; and, specifically for Kant, Marquard 1987). What distinguishes Peirce and other authors belonging to the current of American pragmatism, which founds mind and meaning in practical intersubjectivity, from other philosophical thematizations of the unconscious in the 19th century and at the beginning of the 20th is precisely their radical break with the prerequisites of reasoning assumed by the philosophy of consciousness, without making any concessions to irrationalism, and with the goal of reconstructing the concept of rationality itself.

In Peirce's justification of his own position, Descartes and Kant take central positions in the critical debate. Peirce sharply rejects the methodological postulate of radical doubt set by Descartes, according to which the individual cognizing subject is called to negate all the previous givens of his own reasoning in his own mind to be able to proceed, as a lonesome self-consciousness, to an indubitable basis of cognition upon which the edifice of our knowledge is to be erected completely anew. Peirce considers a doubt with regard to what is considered unproblematic in action to be an intellectualist fiction. For him, the doubt is connected with a discrepancy between our beliefs and our actions, and its aim is to do away this discrepancy. The doubt's point of origin is the divergence between the expectations integrated into a belief, or into a system of beliefs, and the results of the action guided by this belief. The concept of belief links consciousness with action: Belief "is something that ... involves the establishment in our nature of a rule of action, or, say for short, a habit" (Peirce CP 5.397). Elsewhere, Peirce defines belief as "a deliberate, or self-controlled habit" (CP 5.480). The generalization of reasoning is based on the practical generalization in the application of beliefs. The important contribution of the experiment to scientific progress consists in systematically exploiting the orientation of reasoning toward action in order to attain cognitions. Thus, Peirce emphatically stresses Lavoisier's scientific achievement of having carried "his mind into his laboratory, and literally to make of his alembics and cucurbits instruments of reasoning, giving a new conception of reasoning as something which was to be done with one's eyes open, in manipulating real things instead of words and fancies" (CP 5.363). With regard to the dependence upon this orientation toward action, there is continuity between common sense and scientific reasoning, even if the respectively relevant contexts of action and the methods of fixing beliefs may differ.

Just like he destructs Descartes' method, he destructs the result of Descartes' quest for a foundation of all scientific cognition. The reflexive orientation of the cognizing subject toward his or her own thinking in self-awareness is by no means favored epistemologically over the knowledge we have about the social and natural world to which we belong. For Peirce, reflexive self-awareness is the result of a complicated process of learning that integrates experiences with the external world, and in particular with relevant others, and not something given with intuitive certainty that would have to be assumed before any reasoning. Introspection cannot be segregated from questions regarding how dependent our consciousness is on the external world. Just like the consciousness of others, our own consciousness is only accessible if mediated by signs. Without signs, there is no thinking. "Man makes the word, and the word means nothing which the man has not made it mean, and that only to some man. But since man can think only by means of words or other external symbols, these might turn round and say: 'You mean nothing which we have not taught you, and then only so far as you address some word as the interpretant of your thought'" (CP 5.313)

The indissoluble linkage between thinking and sign points to the temporality and unconditional intersubjective character of consciousness. Individual self-consciousness is linked to discovering that one's own opinion deviates from that of others, and thus to discovering the possibility of error, of error that can be clarified and rectified only in communication with others. "Ignorance and error are all that distinguish our private selves from the absolute ego of pure apperception" (CP 5.235). The formulation obviously alludes to Kant who also rejects the Cartesian preference for knowledge generated in reflexive self-reference with regard to *empirical* self-consciousness, but adheres to this preference where the *transcedental* unity of self-consciousness as a condition of the possibility of all knowledge is concerned. A philosophy, however, that considers the attempt to detach consciousness from its embodiment in sign and action to be a false track will not be able to be satisfied with such a transcendental unity of self-consciousness. The pure *ego* of apperception becomes a community's *self*, a community that approaches an understanding of reality by means of practical testing and intersubjective understanding within a process that will never end. The transcendental unity of consciousness becomes intersubjective agreement in the use and interpretation of signs, respective agreement in reasoning and action. Consciousness "is sometimes used to signify the: *I think*, or unity in thought; but the unity is nothing but consistency, or the recognition of it. Consistency belongs to every sign, so far as it is a sign; and therefore every sign, since it signifies primarily that it is a sign, signifies its own consistency. ... But the identity of a man consists in the consistency of what he does and thinks" (CP 5.313 – 315; see, on the reception of Descartes and Kant by Peirce, Apel 1973, 1975).

For Peirce as well, the conscious self is not master in his own house, albeit in quite another sense than that intended by Freud. To form this self logically and temporally presupposes the existence of social systems of signs and actions that are anchored profoundly in a collective history and orient beyond this self toward a collective future. The self is formed in a practical process of learning and acquiring these systems that is the very thing that makes consciousness possible and can be made an object of conscious reflection only in a limited way. Ideas and signs receive their meaning only in the context of habits. Peirce borrows this concept of "habit" from the context of the psychology of association going back to Hume, although not interpreting it like the latter as merely factually sensory, respectively sensori-motor linkage, but rather as embodiment of thoughts, as interpretant of signs, as practical generalization, and thus as a cornerstone of his own theory of meaning. With this, "habit" becomes an epistemological key concept. Habits are embodied processes of logic inference: "A habit arises, when, having had the sensation of performing a certain act, *m*, on several occasions *a*, *b*, *c*, we come to do it upon every occurrence of the general event, *l*, of which *a*, *b* and *c* are special cases. That is to say, by the cognition that

Every case of *a*, *b*, or *c*, is a case of *m*, is determined the cognition that

Every case of *l* is a case of *m*.

Thus the formation of a habit is an induction, and is therefore necessarily connected with attention or abstraction" (CP 5.297). According to the stricter distinction between induction and hypothesis, respectively between inductive and abductive inference, in Peirce's later works, it would probably be more appropriate to grasp the genesis of a habit in the form of an abductive inference.

Now humanity's habits are only in part beliefs, that is, linked to consciousness. Against the definition quoted above, even with regard to beliefs, Peirce is indeed not sure whether they must be bound to consciousness: "Belief is not a momentary mode of consciousness; it is a habit of mind essentially enduring for some time, and mostly (at least) unconscious" (CP 5.417). Peirce tentatively develops a functionalist concept of consciousness that derives it in a way similar to that of deriving doubt from the disturbing confrontation of habits with reality, and whose further elaboration then became a priority task in the development of pragmatism: Consciousness is "symptomatic of the interaction of the outer world - the world of those causes that are exceedingly compulsive upon the modes of consciousness, with general disturbance sometimes amounting to shock, and are acted upon only slightly, and only by a special kind of effort, muscular effort – and of the inner world, apparently derived from the outer, and amenable to direct effort of various kinds with feeble reactions" (CP 5.943).

Peirce works with a concept of meaning and mind that extends far beyond the realm of consciousness, albeit he is as an eminent logician engaged for the concern of an improved conscious control of our reasoning. In his "Lectures on Pragmatism," he concludes: "But the sum of it all is that our logically controlled thoughts compose a small part of the mind, the mere blossom of a vast complexus, which we may call the instinctive mind, in which this man will not say that he has *faith*, because that implies the conceivability of distrust, but upon which he builds as the very fact to which it is the whole business of his logic to be true" (CP 5.212). There are two operations of reasoning whose execution Peirce completely, or at least for an important part, ascribes to the cognitive unconscious of the "instinctive mind," and both are of fundamental importance for human reasoning: perceptual judgments, and abduction. Peirce speaks of the "fineness of subconscious observation," and of the temptation of consciousness toward "breaking down, denying, and pooh-poohing away" this fineness (Peirce RLT 182). He considers this subconscious element of observation to be the "very most important of all the constituents of practical reasoning" (ibid.). Peirce takes up Leibniz's theory of a continuum of representations extending from the quite imperceptible representations to those most coercive to consciousness, albeit in the modification this theory has experienced by Herbarth, and according to which the representations fight for access to consciousness, trying to displace one another. He himself undertook experiments in order to prove that sensual perceptions below the level of consciousness exist, which, nevertheless, are cognitively recorded and enter into perceptual judgments, thus doing pioneer work in the field of studying subliminal perception, a field that has only been discovered as an important field of research by psychoanalysis and cognitive psychology during the last decades (see ibid 312 ff.)

Even with this, however, Peirce still remains within the framework of a psychology of an associative sensualism, and the epistemic significance of the concept of habit remains unexploited. Peirce leaves this framework when he only links the concept of perception to the concept of habit, advocating the thesis "that the conformity of action to general intentions is as much given in perception as is the element of action itself, which cannot really be mentally torn away from such general purposiveness" (CP 5.212). In this context, Peirce develops the conception of the perceptual judgment as an unconscious inference.[2] He justifies this conception by noting that our perceptions as a rule include a classification, speaking of "the interpretativeness of the perceptive judgment" (CP 5.183), which is familiar to every psychologist. The point of his argument, however, is that perceptual judgments must be analyzed logically in the form of the abductive inference, or "that abductive inference shades into perceptual judgment without any sharp line of demarcation between them" (CP 5.181). The cognitive process assumed in the perceptive judgment, however, differs, as a border case of abductive inference, from all other abductive inferences in that it is unconscious and thus exempt from logical control, and through the symptom "that we cannot form the least conception of what it would be to deny the perceptual judgment" (CP 5.186). As a border case of abductive reasoning, the perceptive judgment, too, is bound to signs, in this case, to iconic signs. It is an act of an unconscious cognitive creativity unfolding immediately in sensuality, and refers, as practical generalization, not only to present things but also to future applications in which the field of conscious control then can also extend, proving "that what are really [that is, *fallible*; author's addition] abductions have been mistaken for perceptions (CP 5.188). Such applications, however, presuppose again perception and thus unconscious cognition as the basis of all reasoning.

The thesis suggesting a continuum between perceptive judgment and abductive inference cannot remain without impact on the conception of abduction itself. As it is a leitmotiv of pragmatism to emphasize the orientation of consciousness and of scientific cognition to the future, and to make the idea of the new at home in philosophy, abduction is entitled to take a key role in Peirce's reasoning. "It is the only logical operation which introduces any new idea; ... every single item of scientific theory which stands established today has been due to Abduction" (CP 5.171 ff.). Now it is remarkable that while Peirce distinguishes abduction with regard to its controllability from the perceptual judgment, he emphatically moves it into the neighborhood of the perceptive judgment with regard to the creativity that finds expression in it. He relates abduction to humanity's "faculty of divining the ways of Nature .... An Insight, I call it, because it is to be referred to the same general class of operations to which Perceptive Judgments belong." This faculty resembles instinct too in its small liability to error; "for though it goes wrong oftener than right, yet the relative frequency with which it is right is on the whole the most wonderful thing in our constitution" (CP 5.173). For the new to be able to enter human cognition, it must become present through a reconstruction of sensuality that evades conscious control and forces itself on the consciousness with sudden compulsion: "The abductive suggestion comes to us like a flash. It is an act of *insight*, although of extremely fallible insight. It is true that the different elements of the hypothesis were in our minds before; but it is the idea of putting together what we had never before dreamed of putting together which flashes the new suggestion before our contemplation" (CP 5.181). In Peirce, scientific reasoning unfolds in the context of a sensuality which is characterized by orientation toward action, conveyance by signs, and generalization. Hence, it does not just begin with the cognitive unconscious; it also receives its decisive developmental impulses from it.

PRAGMATISM AND PSYCHOANALYSIS: JOHN DEWEY AND GEORGE HENRY MEAD

Peirce's successors did not agree with important aspects of this theory of the unconscious. What irritated them was the metaphysical embedding of the concept into a theist theory of evolution in which not only the conceptual distinction between instinct and the cognitive unconscious of human reasoning loses its acuteness, but also the distinction between natural law and "habit." This embedding leads to a certain affinity between the unconscious in Peirce and the conception of the unconscious developed by Schelling as well as the German romantic natural philosophy. But this affinity finds its boundary in the semiotic and pragmatic turn initiated by Peirce. After him, it has become a priority task for American pragmatism to develop a concept of communication as a process of understanding between really different subjects, a concept whose difficulty is more played over than solved by Peirce's metaphorical figure of the human as a sign. Peirce's successors struggle with the problem that while Peirce postulated the subordination of the individual consciousness to the consciousness of the community, he did not develop a concept of consciousness that realizes this postulate conceptually. This problem is all the more urgent as William James, who was the first to make pragmatism known to science and to the general public, and without whose support Peirce's philosophical work would have been completely neglected by the contemporary scientific community, indeed tried to found his own individualist variant of pragmatism deviating from Peirce on a psychological study of the "stream of consciousness." The concept of "habit" as embodied generality does not agree well with a psychology of association to which Peirce nevertheless remains confined across wide stretches of his work, and which still exerts a rather unbroken impact on Freud's reasoning as well. Eventually, the reflection on the unconscious in the context of American pragmatism soon fell under the strong influence of psychoanalysis which had been received relatively rapidly and enthusiastically in the United States, to Freud's ambiguous surprise.

The problem situation sketched above reveals why the most eminent pragmatists of the generation after Peirce, George Henry Mead and John Dewey, concentrated rather more on the fields of study concerned with the prerequisites of a theory of the cognitively unconscious: intersubjectivity, consciousness, morals, aesthetics, epistemology, and theory of science. Their analyses attempt to realize the possibility opened up by the sign-theoretical and pragmatic turn in philosophy initiated by Peirce of establishing, respectively elaborating, alternatives to the philosophy of consciousness in all these fields. Because of the push in this direction, the focus was predominantly on a new foundation of the so-called higher cognitive functions, while the cognitive unconscious as a research field of its own was made an explicit topic only marginally, hardly being made a direct object of research. The concept of "habit," the thesis that meaning and mind are more comprehensive than consciousness, and a functionalist conception of consciousness that situates the latter within the context of a practically-sensually mediated unity of subject-object, nonetheless remain guidelines in this. To give only one exemplary quote: "Mind is more than consciousness, because it is the abiding even though changing background of which consciousness is the foreground. Mind changes slowly through the joint tuition of interest and circumstance. Consciousness is always in rapid change, for it marks the place where the formed disposition and the immediate situation touch and interact. It is the continuous readjustment of self and the world in experience. "Consciousness" is the more acute and intense in the degree of the readjustments that are demanded, approaching the nil as the contact is frictionless and interaction fluid. It is turbid when meanings are undergoing reconstruction in an undetermined direction, and becomes clear as a decisive meaning emerges." (Dewey 1934 (1988), 270). One may object that this quote has been taken from the most important contribution of American pragmatism to philosophical aesthetics, thus shifting the subconscious again in the vicinity of creative work, removing it from scientific reasoning. It is, however, the very point of this aesthetics, which has not been published by accident under the title of "art as experience," that it is free from any kind of genius cult, trying to liberate art from its isolation from everyday experience and from scientific reasoning. In doing so – albeit under completely changed social and cultural conditions and in a rather different conceptual and theoretical context – it takes up the object of a "science of the sensual cognition" that was attached to the constitution of aesthetics in the German philosophy of the Enlightenment and was intended to help toward obtaining recognition for the *"fundus animae (Grund der Seele)"* and the subconscious in epistemology.

Mead and Dewey seek the debate with the philosophy of consciousness on the latter's own traditional terrain, hoping, in a second step, to arrive at a convergence with psychoanalysis, which by abandoning this terrain gets itself entangled, in the pragmatists' view, in the pitfalls of this philosophy in an unreflected way. John Dewey has clearly articulated the proximity and distance that simultaneously characterize the relationship of pragmatism after Peirce to psychoanalysis: "The rise at the present time of a clinical psychology which revolts at a traditional and orthodox psychology is a symptom of ethical import. It is a protest against the futility, as a tool of understanding and dealing with human nature in the concrete, of the psychology of conscious sensation, images and ideas .... Every movement of reaction and protest, however, usually accepts some of the basic ideas of the position against which it rebels. So ... the founders of psycho-analysis ... retain the notion of a separate psychic realm or force. They add a statement pointing to facts of the utmost value, and which is equivalent to practical recognition of the dependence of mind upon habit and of habit upon social conditions. This is the statement of the existence and operation of the "unconscious," of complexes due to contacts and conflicts with others, of the social censor. But they still cling to the idea of the separate psychic realm and so, in effect, talk about unconscious consciousness. They get their truths mixed up with the false psychology of original individual consciousness" (Dewey 1922 (1988), 61 ff.).

In the field of psychology, behaviorism and psychoanalysis have pushed this critique to the background, just like the analytic current in Anglo-Saxon definitively seemed to displace pragmatism. Pragmatism itself, however, in no way conceives itself in opposition to these research approaches and tendencies. It is revealing in this context that Mead sees the task of a psychology inspired by pragmatism "to state the whole of human behaviour in scientific terms which would be equally applicable to primitive impulses and to the so-called higher processes and cultural expression" (Mead 1930, 703). He welcomes behaviorism and psychology as first attempts pointing in such a direction, motivating "new methodological approaches" (ibid) beyond themselves. The novelty of these new methodological approaches consists in relating practical generalization and social generalization to one another, and to identify this relation as the source of all processes of generalization.

Neglecting this source has for a long time determined psychological and philosophical reasoning in the 20th century, creating the impression as that pragmatism were no longer more than a marginal phenomenon of the history of psychology and philosophy. In psychology, the result was a behavior that, while being generalized, was not generalizable itself, and an unconscious that stood in contrast to the cognitive appropriation of reality; and in philosophy, a sophistication of the set of logical tools that declared the inquiry into what makes reasoning possible in the first place to be meaningless. This situation has changed fundamentally with the cognitive turn in psychology and the constitution of cognitive science, and it is suggestive that the seeming obsolescence of pragmatism is, in important points, founded in an anticipation of a problem situation for whose reception the time was not yet ripe. Pointing out the problem of the cognitive unconscious is not one of the least important arguments supporting this presumption.

Münster

NOTES

* The quotes from Perrig et al., Bäumler, Baumgarten, Helmholtz, and Schlick were translated for this contribution from the original German or Latin text.

[1] Only such a comprehensive presentation, however, could demonstrate that the concept of the unconscious formulated by Freud and his successors in the 19th century supersedes an older conceptual tradition

beginning with the German Enlightenment in which the term was established for the first time, not only with reference to metaphysics, anthropology, or psychology; but, in the first place, also with reference to epistemology. "What is philosophically typical for the 18th century is thus not rationalism, but the problem of irrationalism seen from the realm of rationalism .... The 19th century does no longer know irrationalism because it is purely irrationalistic itself" (Bäumer 1923/1967, 5). This judgment written in the context of an extensive evaluation of the original contribution made by the German philosophy of Enlightenment, is certainly exaggerated and one-sided, and highly problematic with regard to the concept of rationality assumed as well. With regard to the conception of the unconscious, it highlights, however, an important tendency – especially in German philosophy and psychology (see Marquard 1987; Kaiser-El-Safte 1987). This older conceptual tradition is closely affiliated to the question whether there is a science of sensual cognition, and plays an important role in the constitution of aesthetics as a philosophical discipline (see, for a very instructive contribution, Adler 1988). Baumgarten summarizes the sensual capacities for cognition as 'fundus animae,' as 'bottom of the soul' (*Grund der Seele*). Baumgarten also establishes aesthetics as a philosophical discipline with his definition: "Aesthetics (as theory of the fine arts, as basic epistemology, as the art of fine reasoning, and as the art of thinking analogous to reason) is the science of sensual cognition" (Baumgarten 1750/1758/1988, 3). It is easy to see that the concept of aesthetics that comes to prevail in the 19th century, one which covers nothing but the theory of art and of what is beautiful, represents a reduction compared to Baumgarten's concept of aesthetics. There is no space here to more extensively justify the assumption that the quasi-monopoly to unconscious creativity that psychoanalysis continued to assign to the arts, mirrors this reduction that took place in the transition from Baumgarten to Kant, and within the aesthetics of the 19th century.

[2] It would probably be very instructive to compare this conception of the perceptual judgment as an unconscious inference to the concept of unconscious inference developed by Helmholtz in his "Physiological Optics," which I consider to be the 19th century's most eminent German contribution to an epistemology of the unconscious. Helmholz interestingly tries to underpin this concept with the examples of how the child acquires language and how the artist creates. Later he modifies the term of this controversial concept in order to prevent it being erroneously identified with the irrationalist conception of the unconscious in Schopenhauer, without, however, making any concessions in the matter (see Helmholz 1894, 601 ff, and 1921). Freud stresses this concept as exemplary evidence for the significance of the unconscious, whereas Moritz Schlick strictly rejects it in the ceremony for the 100th anniversary of Helmholtz's birthday, considering it to be at best a *façon de parler* apt to give rise to misunderstandings: "Modern psychology emphatically rejects the concept of unconscious inference because it considers reasoning, the logical process, exclusively as a function of consciousness" (Schlick in Helmholz 1921, 165). Such a stark psychologism, which scarcely conceals the autor's dependence upon the premises of the philosophy of consciousness and betrays a one-sided view of the state of the debate in the psychology of his time, is most astonishing in treating a question that is not only of psychological but also of epistemological import.

REFERENCES

Adler, H. (1988). Fundus Animae – der Grund der Seele. Zur Gnoseologie des Dunklen in der Aufklärung. *DVjs 62,* 197-220.

Apel, K. O. (1973). Von Kant zu Peirce: Die semiotiche Transformation der Transzendentalen Logik. In: Ders.: *Transformation der Philosophie Bd. II. Das Apriori der Kommunikationsgemeinschaft.* Frankfurt/M.

Apel, K. O. (1975). *Der Denkweg von Charles S. Peirce.* Frankfurt/M.

Bateson, G. (1975). *Steps to an Ecology of Mind. Essays in Anthropology, Psychiatry, Evolution and Epistemology.* New York.

Baumgarten, A. G. (1988) (1750/1758). Theoretische Ästhetik. Die grundlegenden Abschnitte aus der "Aesthetica." Hamburg.

Bäumler, A. (1967) (1923). *Das Irrationalitätsproblem in der Ästhetik und Logik des 18. Jahrhunderts bis zur Kritik der Urteilskraft.* Darmstadt.

Bourdieu, P. (1990). *The Logic of Practice.* Cambridge.

Bucci, W. (1997). *Psychoanalysis and Cognitive Science.* New York.

Cavell, M. (1993). *The Psychoanalytic Mind – From Freud to Philosophy.* Cambridge/London.

Davidson, D. (1982). Paradoxes of Irrationality. In: R. Wollheim & J. Hopkins (Eds.): *Philosophical Essays on Freud.* New York/Cambridge.

Davidson, D. (1985). Incoherence and Irrationality. *Dialectica* 39, 245-354.

Dewey, J. (1922) (1988). *Human Nature and Conduct. The Middle Works, Vol. 14.* Carbondale/Edwardsville.

Dewey, J. (1934) (1988). *Art as Experience. The Later Works, Vol. 10.* Carbondale/Edwardsville .

Eagle, M. N. (1984). *Recent Developments in Psychoanalysis. A Critical Evaluation.* New York

Ellenberger, H. E. (1970). *The Discovery of the Unconscious. The History and Evolution of Dynmamic Psychiatry.* New York/London.

Helmholtz, H. v. (1894). *Handbuch der Physiologischen Optik.* Berlin.

Helmholtz, H. v. (1921). Die Tatsachen in der Wahrnehmung. In: Ders.: *Schriften zur Erkenntnistheorie.* Berlin.

Holt, R. R. (Ed.) (1967). *Motives and Thought.* New York.

Kaiser-El-Safti, M. (1987). *Der Nachdenker. Die Entstehung der Metapsychologie Freuds in ihrer Abhängigkeit von Schopenhauer und Nietzsche.* Bonn.

Kihlstrom, J. F. (1987). The Cognitive Unconscious. *Science* 237, 1445-1452.

Leuzinger-Bohleber, M., Schneider, H. & Pfeifer, R. (Eds.), 1992: *"Two Butterflies on My Head ...." Psychoanalysis in the Interdisciplinary Scientific Dialogue.* Berlin/Heidelberg.

Lorenzer, A. (1972). *Kritik des psychoanalytischen Symbolbegriffs.* Frankfurt/M.

Lorenzer, A. (1977). *Sprachspiel und Interaktionsformen.* Frankfurt/M.

Lorenzer, A. (1981). *Das Konzil der Buchhalter.* Frankfurt/M.

Lütkehaus, L. (Ed.) (1995). *Tiefenphilosophie. Texte zur Entdeckung des Unbewussten vor Freud.* Frankfurt/M.

Marquard, O. (1987). *Transzendentaler Idealismus, Romantische Naturphilosophie, Psychoanalyse.* Köln.

Matte-Blanco, I. (1975). *The unconscious as infinite sets. An essay in bi-logic.* London.

McKinnon, J. (1979). Two semantic forms: Neuropsychological and psychoanalytic descriptions. *Psychoanalysis and Contemporary Thought* 2, 25-76.

Mead, G. H. (1930). Cooley's Contribution to American Social Thought. *The American Journal of Sociology* XXXV, 693-706.

Mies, Th./Brandes, H. (1999). Unbewußte, das. In: H. J. Sandkühler: *Enzyklopädie Philosophie* Bd. 2. Hamburg, 1657-1665.

Mies, Th./Scholz, R. (2001). Die Theorie des Unbewussten in Psychoanalyse und Gruppenanalyse. *Arbeitshefte Gruppenanalyse* 16, 24-42.

Pears, D., 1984: *Motivated Irrationality.* Oxford.

Moser, U. & Zeppelin, I. v. (Eds.) (1991). *Cognitve-Affective Processes. New Ways of Psychoanalytic Modeling.* Berlin/Heidelberg.

Peirce (CP) (1992). *Collected Papers of Charles Sanders Peirce* (Volumes I – VI, ed. by Charles Hartshorne and Paul Weiss, 1931-1935, Volumes VII – VIII, ed. by Arthur W. Burks, 1958, quotations according to volume and paragraph). Cambridge, Mass.

Peirce (RLT). *Ch. S. Peirce, Reasoning and the Logic of Things* (The Cambridge Conferences Lectures of 1898). Ed. by K.L. Ketner. Cambridge/London.

Perrig, W. J., Wippich, W. Perrig-Chiello, P. (1993). *Unbewusste Informationsverarbeitung.* Bern/Göttingen/Toronto/Seattle.

Pfeifer, R. & Scheier, V. (1999). *Understanding Intelligence.* Cambridge.

Shevrin, H., Bond, J. A., Brakel, L. A. W., Hertel, R. K. & Williams, W. J. (1996). *Conscious and Unconscious Processes. Psychodynamic, Cognitive, and Neurophysiological Convergences.* New York/ London.

Sulloway, F. J. (1979). *Freud, Biologist of the Mind. Beyond the Psychoanalytic Legend.* New York.

HANS NIELS JAHNKE

HILBERT, WEYL, AND THE PHILOSOPHY OF MATHEMATICS[1]

Abstract. Starting from a critical discussion of P. Forman's thesis about the influence of pessimistic and romantic attitudes on the development of mathematics and physics during the Weimar republic the paper investigates the relation between D. Hibert's and H. Weyl's positions in the foundations crisis of mathematics. H. Weyl's statement that the theoretical concepts of mathematics "are interwoven with the history of thinking and shall never be laid down as a dead final result" is seen as an attractive frame for a future philosophy of mathematics.

Key words: context of science, cultural impredicative definitions, foundations crisis, mathematics as a cultural system Weyl's interpretation of Hilbert's formalistic program.

1. FOUNDATIONS CRISIS AND CULTURE OF SCIENCE AT THE TURN OF THE 20TH CENTURY

Any effort to study the foundations crisis and Hilbert's program today calls for explanation. Hilbert's program as an attempt to achieve a definitely valid foundation of mathematics failed, as Gödel's theorems of 1931 have shown. As a field of research, however, Hilbert's proof theory is still worked on. Philosophically, things do not seem to be promising. Hilbert's program appears to state that mathematics is nothing but a *theory of formal systems* which have no meaning (Curry 1951). Hence, there is a criticism as old as this program that it is an instance of philosophical resignation and signifies a definite abandonment of any epistemological reflection about mathematics (Becker 1927, 32). G. Kreisel speaks of a formalistic-positivistic doctrine (Kreisel 1970). Indeed, a close tie of the philosophy of mathematics to mathematical logic has emerged during recent decades, and for many authors, the task of the philosophy of mathematics is reduced to paraphrase results of mathematical logic in everyday language.

In the following, I should like to show that these conclusions do not follow from Hilbert's position, taking up the interpretation of Hilbert's views offered by Hermann Weyl in the 1920ies for the first time in (Weyl 1924). This is to say that, for me, Hilbert's program and his philosophical interpretation do not seem to have been sufficiently treated and that the attempt is promising to make it fruitful for the philosophy of mathematics. The formalist-positivist doctrine is a later interpretation but by no means Hilbert's own conception. It is surprising that even such philosophers who disagree with the close tie of the philosophy of mathematics to mathematical logic have taken little or no notice of Weyl's considerations, and I know of

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 215 – 228.

no serious attempt to develop these considerations further. The philosophy of Ernst Cassirer might be an exception to this statement. However, it seems that Weyl and Cassirer had consciously maintained a certain distance to each other (see Cassirer 1929, 448 ff).

In order to understand Hilbert's program out of its own, it would seem necessary to recall some of the basic features of the contemporary situation of the sciences. To gain access to this matter, I should like to advance the following idea. The first paradox was discovered in 1895 by G. Cantor (in regard to the history of the foundations crisis see, e. g., Thiel 1972). Russel's paradox was published in 1903. The fundamental dispute or crisis, however, did not take place before 1918. This raises a problem. Obviously, the paradoxes were not of the kind to have a great initial impact on mathematicians. Rather, they were taken to be problems in a special discipline, set theory and logic, which should and could be solved there and did not really concern mathematics as such. Besides: that what is considered the solution of the foundations crisis until today, Zermelo's axiomatization of set theory, had existed since 1908 (Zermelo 1908). Zermelo's axiomatization performs just what is intended to do. It permits to operate safely with sets without running the risk of paradoxes as far as they are known today. But if this is the case, what was then the reason why there was such a heated foundations dispute after 1918 which was actually taken for a crisis by many mathematicians, and which, on a personal level, climaxed in Brouwer's dismissal from the editorship of the *Mathematische Annalen?*

This time lag can be explained by assuming a general change of attitude with many mathematicians which caused that phenomena were thought to be of general meaning which had at first been considered to be isolated. But what can have occasioned such a change of attitude? Two complexes of causes could be taken into consideration here:

- the simultaneous crisis in physics
- changes of the cultural and social situation of the sciences at this time.

Indeed, the American historian of science P. Forman, whose work was mainly on the history of Quantum physics, has attempted to show that there were widespread sentiments of cultural pessimism, romanticism and hostility towards science in the Weimar Republic because of the lost war (Forman 1971). As a case in point, he quotes the philosophy of life and the youth movement. According to Forman, Oswald Spengler's book *The Decline of the West* was the most important expression of these attitudes (Spengler 1918). Representatives of the theoretical disciplines (theoretical physics and mathematics) were among the very people who gave way to these sentiments in their self-representation. This development was quite comparable to the modern debate about problems of environmental pollution and nuclear energy. As a result of this active adaptation to pessimistic attitudes, Forman says, the general validity of the principle of causality had been questioned even before the proper formulation of quantum physics in 1925 - 27 without compelling theoretical or empirical reasons. The general tendency was to flirt with contemporary trends and to try to deny a mechanistic *Weltbild* which left no room for the autonomy of the living individual and for the freedom of man.

Forman says, that the rise of intuitionistic positions must be seen in this context as an attempt to present, and do mathematics in a *humanly conceivable way,* that is intuitively and without the monsters of actual-infinite sets of arbitrary cardinality.

To prove his theses, Forman provides impressive material. For our purposes, Spengler's book is important insofar as Hilbert reacted to it quite strongly. Spengler had advanced a historico-philosophical theory according to which various cultures could be distinguished typologically. They experience rise, climax and decline like organic beings. The transition of a culture into the stage of civilization was the onset of the stage of decline. Hence, the highly civilized West was in a state of decline. Typologically, he distinguished the ontologically meaningful mathematics of the Greek from the Cartesian-rationalist mathematics of modern times. According to Spengler, the latter had entered a stage of internal exhaustion.

In a 1930 lecture with the title *Naturerkennen und Logik* which he held on the occasion of being given honorary citizenship in Königsberg and which was also emitted by radio, Hilbert objected incisively against all those who "had given themselves up to a reactionary and fruitless mania of doubt" and who today "prophesy the downfall of culture with a philosophical face and in a superior tone and please themselves in the *Ignorabimus*." To this, he opposed the slogan: "We must know, we shall know" (Hilbert 1930, 87).

That Hilbert thought such polemics to be necessary can well be interpreted as a symptom of a widespread tendency towards cultural pessimism. His term *Ignorabimus*, however, was not aimed at Spengler, but at the physiologist E. DuBois-Reymond, who, at the close of the 19th century, had been convinced that there was an absolute limit to cognition (Du Bois-Reymond 1872). He had coined the popular phrase of *Ignorabimus* ("We shall not know"). *Ignorabimus*, at the time, was a reflection of the widespread feeling that the natural sciences had reached their limits.

This shows, however, that the tendencies critical of the sciences of that time were not caused by Spengler's *Decline of the West.* They were broader and more manifold in their motives. If one thinks this over in its consequences, one is led toward a critical evaluation of Forman's thesis. It is evidently too simplistic to see the cause for the debate on the strict limits of the principle of causality and the questioning of the classical foundations of mathematics in pessimistic tendencies of the German scientific community after the lost World War, or even to their having read a bestseller of cultural pessimism. In Hermann Weyl's case, in particular, this would lead to a quite inappropriate view of things. Forman claims in all seriousness that Weyl's temporary partisanship for intuitionism and his doubts with regard to strict determinism in physics were due to the impression Spengler's book had had on him. Besides the fact that Weyl's programmatic book on intuitionism *Das Kontinuum* was written in 1917 before Spengler's "Decline" had been published, Forman does not ask whether Weyl has had scientific reasons for his critical view of the principle of causality. Whether physics would be able to maintain a purely deterministic view of natural events had been questionable ever since Boltzmann's work an thermodynamics, and had become an unavoidable problem by Planck's discovery of the *Wirkungsquantum.* Also, the intuitionistic criticism of the foundations of mathematics was not only a question of sentiments, but marked a scientific problem: the foundations crisis, I should like to show, has had a result which was important for mathematics and for the philosophy of mathematics.

Thus, my assessment of Forman's work is ambivalent. For me, it is one of Forman's big merits to have inquired into connections between the developments in theoretical physics, the foundations crisis in mathematics, and the general culture of science. There are linkages which, as far as I know, nobody has as yet studied. Also, the significance of culture as a mediator between the individual disciplines (where else could mediations originate otherwise?) and between science and its applications is as yet far too removed from the attention of the historians and philosophers of science. Forman, however, works with a model of the connection between science and culture which falls short because it offers little opportunity of establishing a link between the intradisciplinary logic of the thing and the mediating role of culture. This is how he is able to paint a picture in which the revolutions of physics and mathematics at the beginning of our century are represented as merely dependent on philosophical trends of fashion.

To put it briefly: I accept the broad frame of Forman's work, but I cannot accept its concrete results.

2. IMPREDICATIVE DEFINITIONS

In the following, I should like to show that what I call the broad frame is quite useful for understanding the foundations crisis, and that Hilbert and Weyl, in particular, did not consider the problem of foundations a mere matter of logic, but saw it in a larger context. To do this, I shall proceed in three steps: 1. I shall sketch an analysis of the paradoxes given by H. Poincaré and B. Russell which motivated Weyl's conversion to intuitionism (Weyl I); 2. I shall briefly present Hilbert's program; and 3. I shall describe H. Weyl's interpretation of Hilbert's program (Weyl II) which, in my opinion, is a convincing alternative both to intuitionism and to the "formalist-positivist doctrine".

Poincaré's and Russel's explanation for the set theoretic paradoxes is that they are based on a hidden circle of reasoning. The cause for the paradoxes be the occurrence of so-called *impredicative definitions* (this term was coined in (Poincaré 1909). The concept of the set of all sets indeed implies that this set contains itself and even its power set as an element. This contradicts an intuitive understanding according to which the set-element relationship results in a hierarchical stratification of the objects which causes a set to be on a higher level than its elements. Russel found out that all paradoxes have one feature in common which could be called *self-reference* or *reflexiveness*.

> ... all our contradictions have in common the assumption of a totality such that, if it were legitimate, it would at once be enlarged by new members defined in terms of itself. This leads us to the rule: 'Whatever involves all of a collection must not be one of the collection,' or, conversely: If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total (Russell 1908, 38).

Thus, Russell established the so-called *vicious circle principle*. "No total can contain elements which are defined by this total itself" (l. c.)

What does this precisely mean? We shall discuss this in a first step using Cantor's diagonal proof for the non-denumerability of the real numbers. The relevance of this example is shown by the fact that the diagonal method does not only prove the non-denumerability of **R** but that Cantor's whole transfinite arithmetic essentially depends on this method. The proof that the cardinal number of the power set of a set M is genuinely larger than the cardinal number of M uses a generalization of the diagonal reasoning. Gödel's construction of an improvable, but nevertheless true theorem of arithmetic essentially uses a diagonal construction, too.

Let us recall what Cantor did. He assumes a denumeration of the real numbers by infinite decimal fractions

$$\begin{aligned} r_1 &= 0, a_{11} a_{12} a_{13} \cdots \\ r_2 &= 0, a_{21} a_{22} a_{23} \cdots \\ r_3 &= 0, a_{31} a_{32} a_{33} \cdots \\ &\ldots \end{aligned} \tag{1}$$

Then, a decimal number

$$r = 0, b_1 b_2 b_3 \cdots \tag{2}$$

can be defined with

$$b_n \neq a_{nn} \tag{3}$$

for all n. The new number r does certainly not occur in the given denumeration (Cantor 1874).

What can be deduced from this proof? Cantor says: people assumed that there is a denumeration of the real numbers and derived a contradiction from that by constructing another real number which, contrary to the assumption, does not belong to this denumeration. Hence, there are non-denumerably many real numbers. To this, Poincaré and the constructivists objected: If one assumes that all numbers of **R** are contained in the denumeration, then the additionally defined element r is defined in a circular way (by an impredicative definition) and thus inadmissible. For r is defined with reference to the set of all real numbers. This set, however, contains r itself, and thus r, finally, is defined dependent on itself: $r := f(r)$, circle! The diagonal method corresponds exactly to Russell's description where a universe is assumed which, if it were legitimate, would at once be enlarged by new elements which are defined with its help. Hence, this contradicts the rule he established: if a set which represents a total contains elements which can only be defined with the help of this total, then this said set is no total.

There exists, however, an interpretation of the diagonal method which is acceptable in the constructivist sense. According to that interpretation, Cantor's proof says that additional real numbers can be constructively defined for any denumerable set of real numbers. Here, the additional elements are not defined impredicatively, because they did not belong to the original set. Thus, in the constructivist sense, Cantor's proof can be used to prove the existence of transcendental numbers. The usual argument saying that the set of algebraic numbers is denumerable, that of the real numbers non-denumerable, and hence there are transcendent numbers, however, must not be used in this way. Instead, the proof must be shaped so as to effectively construct, for an effective denumeration of the algebraic numbers, a transcendental number according to the diagonal method which does not belong to the algebraic numbers. In this version, the method is constructive. According to a strictly determined rule, one obtains step by step the decimal representation of a real number which can be proved to be different from all algebraic numbers (Kaufmann 1930).

The interdiction of impredicative definitions does not only strongly intervene into Cantor's transfinite arithmetic, but it was quickly realized that it puts extensive parts of classical mathematics in question. This can be easily shown with the example of the existence of the supremum of a bounded set of real numbers. In order to obtain the proof that any such set has a supremum, this supremum is defined by a condition which necessarily implies a quantification over all real numbers and thus also over the sought supremum itself. The latter is thus impredicatively defined.

In his book *Das Kontinuum* of 1918, Weyl quoted this example as the crucial motivation for his conversion to constructivism.

Weyl commented this difficulty as follows:

> The vicious circle concealed by the foggy character of the common concepts of set and function, which we are pointing out here, is by no means a formal error in the structure of analysis which can be easily corrected. The insight that it is of fundamental importance is something which cannot be conveyed to the reader with few words. The more distinctly one calls the logical network of analysis to mind, the clearer it becomes that, with the present mode of foundation, so-to-say every cell of the huge organism is permeated by this toxic contradiction. (Weyl 1918, 23)

P. Lorenzen, whose constructivist foundation of analysis is essentially an elaboration of Weyl's approach, has formulated a weaker theorem at this point which postulates that a supremum exists only for such sets which can be replaced, with regard to the property of having a supremum, by a definite set of rational numbers (bounded sets with definite left class) (Lorenzen 1965, 66).

Four solutions have been given for the problem of impredicative definitions: 1. Gödel: He says that the impredicative definitions compel to a platonic philosophy of mathematics. If we assume that the set of all subsets of the natural numbers exists independent of definitory predicates, then definitions of this type are not circular, since they can be understood to be selections or characterizations of a certain set out of a universe of sets (Gödel 1944). This is the view taken by most mathematicians. But there is one thing it shows: without assuming independent objects, mathematics cannot be founded. 2. Russell: Together with Whitehead, he developed the so-called theory of types in which the predicates are ordered into a hierarchy of types and quantification is only permitted over predicates of lower levels. In order to save the usual modes of deduction of analysis, Russell introduced a so-called axiom of reducibility which requires that there be an extensional predicative definition for every impredicative one (Russell 1908). This requirement, however, is an assumption which is legitimized only by the purpose which in fact contradicts the logicist structure of mathematics.

The third solution was that of intuitionism, and the fourth was Hilbert's proof theory. Intuitionism shall only be briefly dealt with here. It is governed by the basic attitude expressed by Poincaré as follows: Every theorem in mathematics must be verifiable. As soon as I propose a theorem, I claim that all verifications which could be tried must be successful; even if they are beyond a person's forces. Verifications, however, can only refer to finite numbers, and hence all theorems about infinite sets are nothing but abbreviated claims about finite numbers (Poincaré 1909, 138 ff.). In intuitionism, the sequence of natural numbers is thus the basic starting point of mathematics given in inner intuition, not as something ready-made, but as some process of growth. For Brouwer, this entailed that some laws of classical logic which are valid for finite sets loose their meaning for infinite ones (A well readable exposition of the basic ideas of intuitionism is Bouwer 1912).

It is well known that this conception not only compels us to abandon important theorems of classical mathematics, but also entails enormous difficulties for the practice of mathematical deduction. I should like to point out an aspect which was important to Hermann Weyl. In the intuitionist view, the object of mathematics are sequences which are given by a law and thus can basically be seen, and other, so-called *"free choice, becoming sequences"* which can be taken into account only as far as they have been "actually realized" and numerical values are known.

It is frequently overlooked that this can be interpreted to mean that the object of mathematics is conceived of as a network of determinism and indeterminism. Brouwer did not do so. H. Weyl, however, has pointed out analogies between the intuitionistic conception and contemporary physics. Thus, in a paper which Forman interpreted to be proof of Weyl's conversion to Spengler's cultural pessimism, Weyl tried to argue that Brouwer's *"free choice, becoming sequences"* present a remarkable analogy to a process taking place in quantum jumps and that the intutionistic conception of the continuum is thus much closer to contemporary physics than the traditional view (Weyl 1923). The continuum of intuitionism is no set of unequivocally determined states, but an inward process of growth. As far as I know, however, he never went beyond such hints.

3. HILBERT'S PROOF THEORY AND WEYL'S INTERPRETATION

Let us proceed to Hilbert's proof theory. The idea of proof theory is mentioned for the first time in (Hilbert 1918). Then in the course of the 1920s, Hilbert outlined his conception in a series of lectures and papers (Hilbert 1922, 1923, 1925, 1928, 1931).

Against Weyl, Hilbert insisted that philosophy had to adapt to mathematics, and not vice versa. He wrote:

> ... in my opinion, [Weyl] would have had to recognize, just because he arrived at a circle, that his point of view and hence the constructive principle in his own version and application are impracticable and that the path to analysis is not reachable from his position. (Hilbert 1922, 158)

Let us first recall the basic ideas. In his first step, Hilbert says that the usual rules of inference are legitimate only for finite sets. Logical deduction, in which the signs used to have an objective meaning, is thus possible only for finite object fields, it proceeds arithmetically and combinatorially.

As a necessary prerequisite for any kind of deduction, we may thus assume certain extralogical, discrete objects which

> are intuitively present as an immediate experience before all thought. If the logical deduction is to be certain, these objects must be totally visible in all their parts, and their perception, their distinction, their sequence is immediately visualized together with the objects themselves as something which cannot be reduced to something else (Hilbert 1925, 171).

Among such objects, Hilbert counts the natural numbers, more precisely, finite combinations of strokes which can be established, analysed and compared, and which can be controlled combinatorially. Thus, we obtain an elementary part of arithmetics which only encompasses verifiable statements about finite sets of natural numbers, but not about all natural numbers. Operations in this field are purely intuitive and hence need no axiomatic basis.

The foundation of classical mathematics is then achieved by adding, to these (number) signs, further logical signs, mathematical signs, and different kinds of letters. The use of these additional signs is fixed by rules (axioms). This means that the finite arithmetic which was used to begin with is supplemented by further elements as ideal elements. Such an addition is possible if it can be done consistently, and this means it must be shown that eliminating these ideal elements does not lead to contradictions in the traditional field of finite arithmetics.

If all propositions which make up mathematics are transformed into formulae, mathematics proper becomes a stock of formulae, and by marking certain of these formulae as axioms, it becomes possible to transform the entire mathematics into an operating with signs. Material deduction is replaced by external action according to rules. This enables us to change the totality of mathematics into a controllable calculus. In particular, Hilbert expected that the consistency of mathematics could be proved by formalization. A proof would have to show that the formula $1 = 0$ can never be deduced if operating is done according to the rules.

Finite mathematics plays a double role in this approach. On the one hand, it is the materially certain part of mathematics with which all mathematical deduction must begin, and on the other hand it is the theory which permits to analyse the operations with the signs, and thus is metamathematics.

We refrain from discussing the details and just remark that Hilbert needs one single axiom which is similar to the axiom of choice known from set theory in order to found the transfinite modes of deduction which were the starting point of the conflict. He requires the existence of a function $\in$ which selects an element from each set, and this function is to be specified by the following condition:

$$A(a) \rightarrow A(\in A) \qquad (4)$$

This axiom enables us in particular to formulate the rules of deduction for the tertium non datur for infinite sets and to prove their consistency (Hilbert 1928, 67 ff.).

In order to get an idea of this Hilbertian mathematics, H. Weyl's analogy may be useful (Weyl 1924, 147 ff.). He compared Hilbert's mathematics with a game of chess. The signs are the chessmen, a position on the board is a formula, the starting position are the axioms, and the chess rules are the rules according to which formulae are derived from formulae. A position in chess conforming to the rules is one which has developed from the starting position by application of the rules. To this corresponds, in Hilbert's mathematics, a formula which can be proved. Just as one can show that a position in chess in which ten queens of the same colour appear is impossible, it should be demonstrable for arithmetic that the formula $1 = 0$ cannot be derived.

What is the matter with this Hilbertian conception? Is it true that it transforms mathematics into a meaningless game, or, to state it with more caution, that mathematics has definitely been freed of all ontological ties? What about the object of pure mathematics according to Hilbert? This question shall be treated in the last part of this paper by presenting an interpretation of Hilbert's position given by H. Weyl. It was first published in the *Mathematische Zeitschrift* in 1924 under the title *Randbemerkungen zu Hauptproblemen der Mathematik* and was later repeated in the famous *Diskussionsbemerkungen zu dem zweiten Hilbertschen Vortrag über die Grundlagen der Mathematik* in the mathematical seminar of Hamburg university in 1928. With this interpretation, Weyl withdraw from intuitionism, and in my opinion the paper contains some remarkable thoughts (Weyl 1929).

Hilbert himself had justified the transformation of mathematics into a game of formulae by pointing to the *method of ideal elements*. Its essence is the introduction of new elements into mathematics which cannot be interpreted in the frame of the traditional theory and which are defined only by certain rules. For mathematics, this was a current and successfully applied method in the 19th century. Hence, Hilbert said, it is "by no means reasonable" to make the requirement that "every formula must be interpretable for itself." A new point of view was that Hilbert claimed this for physics as well. In physics, too, it is impossible to interpret every concept and every formula in an empirical way. The experiment controls only certain combinations and deductions of the physical laws, and the physicist does not require that all his concepts should be empirically interpretable (Hilbert 1928, 79).

This is from where Weyl's interpretation of Hilbert went on. It had hitherto been a view shared by all mathematicians, he says, that mathematics is a system of intuitive, reasonable, cognizable truths. The foundations crisis now had shown that this position could no longer be maintained. Brouwer, in particular, had made it clear that mathematics had far surpassed the Limits of intuitive thinking. This was his historical merit. Hilbert's achievement, however, was to have seen that classical mathematics must be interpreted anew for this reason. This was an insight which merited to be fixed as a thesis. Together with Weyl, we may say that the prevalence of the Hilbertian conception "means a decisive defeat of the philosophical attitude of pure phenomenology" (Weyl 1928, 88). For Weyl, any conception is "phenomenological" which strives to interpret mathematical concepts and formulae totally by phenomena intuitively given which exist independent of and anterior to any mathematical theory; of course, this is opposed he has intuitionism in mind.

Here, Weyl said, is an exact analogy to the natural sciences. There, too, the theoretical concepts and propositions could not be empirically interpreted and justified for themselves. In contrast to phenomenal cognition which merely stated what was given in intuition, and in which every judgement had its own sense which could be fully ascertained in intuition, the situation in theoretical physics was quite different. Here, the individual propositions were not empirically interpreted and tested each for itself, but the entire theoretical system was put into question by confronting it with experience.

This shall be demonstrated with an elementary consideration. In the history of physics, the status of the formula

$$F = m \cdot a \tag{5}$$

has been controversial since Newton. If it is a law of nature, the quantities F, m and a must be empirically defined. For the acceleration a, this is certainly the case, it can be determined by measuring lengths and time. For the inert mass m, however, the situation is more complicated. Indeed, there is no general method of measuring masses known which does not assume, in one or another form, the validity of Newton's first law. For the measuring of forces, even additional special laws must be assumed to be valid beforehand. Here again, it can be said that the situation is hardly compatible with the traditional understanding of definitions. There is a circular reasoning here which bears a certain similarity to impredicative definitions.

One possibility of solving this difficulty which is applied in modern epistemology with formal methods amounts to conceive of the Newtonian first law simultaneously as of an implicit definition and a law. m and F are measured and mutually determined within a complicated network of applications (see Sneed 1971 and Stegmüller 1973). Then, however, a theory can no longer be considered as a system of propositions, but rather as a pair formed by a so-called structural kernel K and a set I of intended applications: (K, I).

In this way, Newton's first law becomes a general scheme which is applied in a manifold way in order to develop methods for measuring m and F. This could be expressed as follows: the theoretical concepts of m and a have no previously given empirical content, but they produce methods of verifying them by application. For this state of affairs, it is clear that Newton's First law cannot be falsified by a single empirical observation. This is the very reason why the individual propositions do not form the empirical content in classical mechanics, but rather, as Weyl expressed it, only the entire system as a whole can be accepted or rejected

Hence, our result is that mathematics and physics use circular definitions of their fundamental concepts in quite a similar way. This is the core of the analogy between the ideal elements in mathematics and die fundamental theoretical concepts in physics (see Jahnke 1978, chapt. III for a detailed exposition of this analogy).

In this, Weyl sees a possibility of giving meaning to Hilbert's theoretical mathematics, as opposed to Brouwer's intuitive mathematics, which we shall again describe in Weyl's words and designate as a further thesis.

> One finds the beyond to which the symbols of mathematics are related by allowing mathematics to fuse completely with physics and assuming that the mathematical concepts of number, function, etc. (or the Hilbertian symbols) participate basically in the same way in the theoretical construction of the real world as the concepts of energy, gravitation, electron do. (Weyl 1934, 150)

The question is: What then is the justification of an independent pure mathematics? Or does Weyl give up this independence?

Some indication may be got from other conclusions Weyl drew: While the intuitive phenomenological cognition, for instance the intuition of finite combinatorial facts, is subject to error, but unchangeable in its essence, this is not true for the higher, transphenomenal mathematical constructs. The criterion of consistency is necessary for these, but it is not sufficient. Since aspects of fruitfulness and value are important here, they share the destiny of all other theoretical insights: they are *subject to change.*

As the relationship between the theoretical constructs and empirical reality is mediated and not unequivocal, the selection of these constructs is influenced by factors which are beyond our conscious control and only so much can be said about them that they correspond to our needs by contributing to an interpretation of reality. Weyl says that the meaning of theoretical creation is as obscure for us as that of creative art. If the theoretical concepts are to serve human understanding, they are linked to the respective way humans historically and concretely confront their own reality. Weyl has formulated this in a statement which we should like to maintain as another thesis: the theoretical concepts *"are interwoven with the history of thinking and shall never be laid down as a dead 'final' result"* (l. c.)

In this way, mathematics is conceived of as a cultural system subjected to change.

I should like to summarize the results of these considerations in four theses which at the same time sketch a general program of historical and didactical research.

- The attraction of Weyl's conception, in my opinion, lies in his conclusion that mathematics can no longer be founded as an independent system, but only within the context of men's effort to interpret and control their natural and social environment. At the same time, mathematics is not reduced to its applications. Rather, the autonomy of theory is emphasized. Theories are complicated, self-referring systems. Mathematics, in particular, introduces a specific element of the formal, of the consistent and of the transparent into the interpretation of reality. This is what Weyl means with his analogy to "creative art."
- Mathematics, since Hilbert, is neither without ontology, i. e. not simply a theory of formal systems, nor must, conversely, every concept, every formula or every proposition be verifiable in an empirical-intuitive way as Poincaré

demanded: rather, the decisive insight is that *every theory only creates its respective own methods and processes of verification.*

- Gödel's escape into Platonism does not seem compelling to me. Rather, a conclusion from the second thesis is that the way mathematics refers to applications changes from theory to theory, from application to application. Hence, this relationship can and must be historically investigated if we intend to develop an appropriate understanding of mathematics. Thus, there can be no philosophy of mathematics which describes the reference of mathematics to reality a priori in a valid way.
- As the connection between theory and application is ambiguous, theories will never be justifiable from their applications alone. This is why we have to assume with Weyl that the historical reconstruction of the objectiveness of mathematics must inevitably also investigate the way it is integrated into a culture. Mathematics is a cultural system. Conversely, however, it is also true that the cultural side of mathematical thinking attains its realistic status only if mathematics is understood to be a part of our confrontation with reality. Mathematics is a human construction (man's symbolic construction, as Weyl says), which is only understood in an appropriate perspective and in its entire deep structure if its tension with the natural and social reality is not lost out of view.

POSTSCRIPT 2003

Since 1990 quite a few historical studies appeared which are relevant to the theme of the present paper. Especially, I would like to refer to Skúli Sigurdson's PhD thesis of 1992 *Hermann Weyl, Mathematics and Physics, 1900 – 1927* and to the volume *Hermann Weyl's Raum-Zeit-Materie and a General Introduction to his Scientific Work*, edited by Erhard Scholz (2001). This volume contains a chapter by Robert Coleman and Herbert Korté with among others a detailed exposition of Weyl's approach to the foundations of analysis. A critical examination of Forman's paper from the point of view of a historian of physics is Hendry (1980). I was not aware of the latter paper when I wrote the original version of the present study.

Also, D. Hilbert's lectures *Natur und mathematisches Erkennen*, edited by D. Rowe in 1992, showed very clearly that his general views on mathematics were, in regard to the close relation of mathematics and physics, similar to that of Weyl and cannot be qualified as formalist. (cf. Rowe 1997)

In 1994, Michael Otte published *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik,* a book in which he argued that mathematics as a cultural and social enterprise can be understood adequately only by taking into account its reference to applications ("Gegenständlichkeit mathematischer Erkenntnis").

Universität Duisburg-Essen

NOTES

[1] This paper is the English version of (Jahnke 1990). I would like to thank Abe Shenitzer, Toronto, for his generous help in polishing the translation.

REFERENCES

Becker, O. (1927). Mathematische Existenz. Untersuchungen zur Logik und Ontologie mathematischer Phänomene. In *Jahrbuch für Philosophie und phänomenologische Forschung B*. Halle: Niemeyer. Quoted according to the reprint Tübingen 1973.

Brouwer, L. E. J. (1912). Intuitionisme en formalisme. Quoted according to the English translation. In P. Benaceraff, & H. Putnam (Ed.), *Philosophy of mathematics*, 2nd ed. (1983), Cambridge, 77-89.

Cantor, G. (1874). Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen. *Journal für die reine und angewandte Mathematik* 77, 258-262.

Cassirer, E. (1929). *Philosophie der symbolischen Formen. Dritter Teil: Phänomenologie der Erkenntnis.* Berlin: B. Cassirer. Quoted according to the reprint of the 2nd ed. Darmstadt: Wissenschaftliche Buchgesellschaft 1977.

Curry, H. B. (1951). *Outlines of a Formalist Philosophy of Mathematics*. Amsterdam: North-Holland Publ.

Dieudonné, J. (1970 ff.): Entry "Weyl, Hermann". In: C. C.Gillispie (Ed.), *Dictionary of Scientific Biography*, Vol.14, New York: Scribner, 281-285.

Du Bois Reymond, E. (1872). Über die Grenzen des Naturerkennens. In S. Wollgast (Ed.), E. Du Bois-Reymond, *Vorträge über Philosophie und Gesellschaft* (1974). Hamburg: Meiner, 54-77.

Forman, P. (1971). Weimar Culture, Causality and Quantum Theory. 1918 – 1927: Adaptation by German Physicists and Mathematicians to a Hostile Intellectual Environment. *Historical Studies in the Physical Sciences* 3, Berkeley, California (u.a.): Univ. of California Press, 1-115.

Gödel, K. (1944). Russell's Mathematical Logic. In P. A. Schilpp (Ed.), *The Philosophy of Bertrand Russell*. New York: Northwestern Univ., 125-153. Quoted according to P. Benaceraff, & H. Putnam (Ed.), Philosophy of mathematics. Selected readings. 2nd ed. (1983). Cambridge, 447-469.

Hendry, J. (1980). Weimer culture and quantum causality. *History of Science 18*: 155-180. German Translation in von Meyenn, K. (Ed.), Quantenmechanik und Weimarer Republik (1994). Braunschweig: Vieweg, 201-230.

Hilbert, D. (1918). Axiomatisches Denken. *Mathematische Annalen* 78, 405-415. Quoted according to: Hilbertiana (1964). Darmstadt: Wissenschaftliche Buchgesellschaft, 1-12.

Hilbert, D. (1922). Neubegründung der Mathematik. Erste Mitteilung. *Abhandlungen des mathematischen Seminars zu Hamburg* 1, 157-177. Quoted according to: Hilbertiana (1964). Darmstadt: Wissenschaftliche Buchgesellschaft, 12-32.

Hilbert, D. (1922/1992). *Natur und mathematisches Erkennen, Vorlesungen, gehalten 1919 – 1920 in Göttingen*. Ed. by David E. Rowe. Basel: Birkhäuser.

Hilbert, D. (1925). *Über das Unendliche*. Quoted according to: Hilbertiana (1964). Darmstadt: Wissenschaftliche Buchgesellschaft.

Hilbert, D. (1928). *Die Grundlagen der Mathematik.* Abhandlungen des mathematischen Seminars zu Hamburg 6, 65-85.

Hilbert, D. (1930). *Naturerkennen und Logik. Naturwissenschaften*, 959-963. Quoted according to: Hilbert, Gesammelte Abhandlungen III, 378-387.

Hilbert, D. (1923). Die logischen Grundlagen der Mathematik. *Mathematische Annalen* 88, 151-165. Zitierte Ausgabe: Hilbertiana (1964). Darmstadt: Wissenschaftliche Buchgesellschaft, 33-46.

Jahnke, H. N. (1978). Zum Verhältnis von Wissensentwicklung und Begründung in der Mathematik - Beweisen als didaktisches Problem. *Materialien und Studien des IDM* 10. Bielefeld.

Jahnke, H. N. (1990). Hilbert, Weyl und die Philosophie der Mathematik. *Mathematische Semesterberichte* 37, 157-179.

Kaufmann, F. (1930). *Das Unendliche in der Mathematik und seine Ausschaltung. Eine Untersuchung über die Grundlagen der Mathematik.* Wien: Deuticke. Quoted according to: C. Thiel (Ed.), Erkenntnistheoretische Grundlagen der Mathematik (1982). Hildesheim: Gerstenberg, 242-274.

Kreisel, G. (1970). The formalist-positivist doctrine of mathematical precision in the light of experience. *L'âge de la science* 3, H.1, 17-46. Quoted according to the German translation. In M.Otte (Ed.), Mathematiker über die Mathematik (1974). Berlin-Heidelberg-NewYork, 65-123.

Lorenzen, P. (1965). *Differential und Integral. Eine konstruktive Einführung in die klassische Analysis.* Frankfurt (Main): Akad. Verl.-Ges.

Otte, M. (1994). *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik.* Frankfurt/Main: Suhrkamp.

Poincaré, H. (1909). La Logique de l'Infini. *Revue de Metaphysique et de Morale* 17, 461-482. Quoted according to: H.Poincaré, Letzte Gedanken (1913). Leipzig, 99-143.

Rowe, D. (1997). Perspective on Hilbert. *Perspectives on Science* 5, 523-570.

Russell, B. (1908). Mathematical Logic as Based an the Theory of Types. *American Journal of Mathematics* 30, 222-262. Quoted according to the German translation. In B.Russell (Ed. J. Sinnreich), Die Philosophie des logischen Atomismus (1976. München, 23-65.

Sigurdson, S. (1992). *Hermann Weyl, Mathematics and Physics, 1900 – 1927* PhD Dissertation Harvard University.

Scholz, E (Ed.). (2001). Hermann Weyl's Raum-Zeit-Materie and a General Introduction to his Scientific Work. *DMV Seminar Band* 30. Basel/Boston/Berlin: Birkhäuser.

Sneed, J.D. (1971). *The logical structure of mathematical physics.* Dordrecht: Reidel.

Spengler, O. (1918). *Der Untergang des Abendlandes. Umrisse einer Morphologie der Weltgeschichte.* Vol. l: *Gestalt und Wirklichkeit.* München.

Stegmüller, W. (1973). Theorienstrukturen und Theoriendynamik. *Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, Vol. II, 2. Halbband. Berlin-Heidelberg-NewYork: Springer.

Thiel, Ch. (1972). *Grundlagenkrise und Grundlagenstreit.* Meisenheim am Glan: Hain.

Weyl, H. (1918). *Das Kontinuum.* Berlin. Quoted according to the reprint. In H. Weyl, Das Kontinuum und andere Monographien (1973). New York: Chelsea Publ. Co.

Weyl, H. (1920). Das Verhältnis der kausalen zur statistischen Betrachtungsweise in der Physik. *Schweizerische Medizinische Wochenschrift* (1920). Quoted according to: H. Weyl, Gesammelte Abhandlungen II, 113-122.

Weyl, H. (1924). Randbemerkungen zu Hauptproblemen der Mathematik. *Mathematische Zeitschrift* 20, 131-150.

Weyl, H. (1928): Diskussionsbemerkungen zu dem zweiten Hilbertschen Vortrag über die Grundlagen der Mathematik. In *Abhandlungen des mathematischen Seminars zu Hamburg* 6, 86-88. Quoted according to: H. Weyl, Gesammelte Abhandlungen III, 147-149.

Zermelo, E. (1908). Untersuchungen über die Grundlagen der Mengenlehre I. *Mathematische Annalen*, 261-281.

THOMAS MORMANN

MATHEMATICAL METAPHORS IN NATORP'S NEO-KANTIAN EPISTEMOLOGY AND PHILOSOPHY OF SCIENCE

Abstract. A basic thesis of Neokantian epistemology and philosophy of science contends that the knowing subject and the object to be known are only abstractions. What really exists, is the relation between both. For the elucidation of this "knowledge relation" ("Erkenntnisrelation") the Neokantians of the Marburg school used a variety of mathematical metaphors. In this contribution I'd like to reconsider some of these metaphors proposed by Paul Natorp one of the leading members of the Marburg school. It is shown that Natorp's metaphors are not unrelated to those used in some currents of contemporary epistemology and philosophy of science.

Key words: mathematics, metaphor, Neokantianism, Marburg school, Natorp.

1. INTRODUCTION

Since some time "postpositivist" philosophy of science has become interested in its history and evolution. In order to understand science, not only history of science but also history of philosophy of science has become an important topic for philosophy of science. As a result of this attitude Neokantian philosophy is being re-evaluated as a hitherto unduly neglected source of philosophy of science and epistemology. For instance, the investigations of Coffa, Friedman and others have shown that Neokantian philosophy played an eminent role for the emergence of the Logical Empiricism of the Vienna Circle. This holds in particular for the Marburg School, whose most important members were Cohen, Natorp and Cassirer (cf. Coffa 1991, Friedman 1999, 2000). It goes without saying that a short paper like this is not the appropriate place to present a detailed account of the Neokantian philosophy of science and its relation to modern philosophy. The aim of this contribution is more modest. Provisionally accepting Rorty's thesis that "it is pictures rather than propositions, metaphors rather than statements, which determine most of our philosophical convictions" (Rorty 1979, 12), I want to take a shortcut reconsidering some of the core metaphors that guided Neokantian epistemology and philosophy of science.[1]

Remarkably, the guiding metaphors of Neokantian epistemology and philosophy of science have their origin in science itself, in particular in mathematics.[2] This points at a rather complex relation between science and philosophy of science that is not adequately described by the standard 2-level account according to which philosophy of science is a sort of metascience dealing with the sciences as its object.

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 229 – 239.

Dealing with Natorp's metaphors, I'd like to show two things: first, the Neokantian metaphors are surprisingly modern. They may still deserve to be taken into consideration by contemporary philosophy of science. Secondly, a closer look at the metaphorical apparatus of a gone-by philosophical stance may help sharpen our own sensitivity for the often murky metaphorical ground on which many of our own basic philosophical convictions are based.

More precisely I want to concentrate on some metaphors that Natorp used for the elucidation of a basic thesis of Neokantian epistemology put forward by virtually all authors of the Marburg school and most other Neokantians. According to this thesis the true issue of epistemology is neither the knowing subject nor the known object, but the "knowledge relation" ("Erkenntnisrelation") by which subject and object are related. Subject and object are mere abstractions. Hence, strictly speaking, only the "knowledge relation" exists (cf. Cassirer 1910, Rickert 1915, Natorp 1903, 1912). The K-relation, as I want to call it, has a privileged status with respect to its relata, to wit, the knowing subject on the one hand, and the known (or knowable) object on the other. Rival epistemological approaches such as empiricism, positivism, and non-critical versions of idealism like Hegelianism, are accused by the Neokantians to commit a reductive fallacy falling back on some apparently simpler "monistic" position that eliminates the K-relation in favour of one of its relata. In the end, all these positions are claimed to be unable to characterize the true nature of science as an ongoing process of knowledge acquisition.[3]

For the elucidation of the K-relation, Neokantian philosophy used a variety of pictures, analogues, and metaphors. For the Marburg School the paradigm of knowledge was scientific knowledge, more precisely, mathematics and mathematical physics. Hence it is not surprising that in the Marburg account mathematical metaphors played an important role. Maybe the first of these guiding metaphors was due to Hermann Cohen, the founder of the school. According to him, the essence of the formation of mathematiced empirical science was to be found in the concept of the infinitesimal (Cohen 1863).[4] Cohen's mathematical erudition was not very profound, and he presented this thesis in a rather obscure way. Hence, his account did not gain much real influence, even among the members of the Marburg school. Cassirer's "functional approach" of critical idealism became better known one or two generations later. In *Substance and Function* (1910) Cassirer put forward a "functional" or "relational" account of scientific concepts in which he contended that the essence of the modern science resided in the concept of mathematical function.

Cassirer's "function" was by no means the only mathematical metaphor that guided Neokantian epistemology. In this paper I'd like to consider some of the lesser known mainly due to Natorp that served as guiding lines for the Marburg Neokantianism in general. In his lifetime Natorp was one of the most influential members of the Marburg school. Before Cassirer became prominent he was a kind of official spokesman of the Marburg school whose sober and relatively accessible treatises (compared with the writings of Cohen) taught generations of students the basics of the school's doctrines (cf. Natorp 1903, 1910, 1929).[5]

Natorp took his metaphors seriously. For him, they were more than embroideries, rather, he used them as "intuition pumps" to develop his account of scientific knowledge. He attempted to draw contentful conclusions from them, considering them as models that could be used for the description of the sciences, their methods and development. True, Natorp's metaphors are no longer ours, and sometimes they appear strange and contrived. Nevertheless, even contemporary epistemology and philosophy of science can hardly be said to be an area free of metaphors, as will be briefly discussed in the last section.

The outline of this paper is as follows. In section 2 the Neokantian transformation of the Kant's original epistemological position is discussed.[6] This sets the stage for the detailed analysis of some of Natorp's core metaphors in section 3. In particular, we will deal with his "equational account" of knowledge according to which knowing (cognizing) may be characterized as an activity analogous to solving a mathematical equation. The paper concludes with some general remarks on the problematic of metaphors in philosophy comparing Neokantianism with some post-positivist authors.

2. THE NEOKANTIAN REFORMULATION OF KANT'S EPISTEMOLOGY

The Neokantian approach to epistemology and philosophy aimed to be faithful to the spirit but not to the letter of Kant's philosophy. For Natorp this meant to restitute the "transcendental method" as the true core of the Kantian approach, and to give up all of ingredients of Kant's system that did not sit well with that method. The transcendental method deals with the problem of the possibility of experience. The NeoKantians interpreted Kant as contending that the object of experience is determined by the laws and methods of the knowing subject. Thereby the object no longer is something given ("gegeben") but something "posed" ("aufgegeben") (cf. Kinkel 1923, 405). Conceiving Neokantian philosophy as based on the transcendental method has two implications:

(i) Philosophy recognizes the historical, societal and scientific context in which it exists. It is aware that it is rooted in the specific theoretical and practical experiences of its time and refuses to build up "high towers of metaphysical speculations" (cf. Natorp 1912, 195, Kinkel 1923, 402/403).

(ii) Philosophy accepts the facts of science, morality, art and religion. The task of philosophy is to carry out a *deductio iuris* of these facts, i. e., it has to provide a kind of "logical analysis" which shows the reasons why these facts are possible thereby revealing what is the "quid iuris" of them. In still other words, and going beyond the epistemological sphere, philosophy has to show the lawfulness and reasonableness of the cultural achievements of mankind.

Thereby the philosophy of critical idealism is lead to a "genetic" epistemology and theory of science that regards the ongoing process of scientific and cultural creation as essential, not its temporary results. These are to be considered as being of secondary importance. As Natorp put it with respect to scientific knowledge: knowledge is always "becoming" and is never "closed" or "finished." There never is something "given" that is not transformed in the ongoing and strictly speaking infinite process of cognition. The "fact of science" is, according to Natorp to be understood as a "fact of becoming" ("Werdefaktum").

The rejection of a non-conceptual given in any form brings the Marburg brand of Neokantianism in open conflict with some of the corner-stones of Kant's epistemology, to wit, the dualism of "scheme" and "intuition," and related dualisms such as that of "spontaneity" and "receptivity" of thinking: "Maintaining this dualism of epistemic factors (receptivity and spontaneity, T. M.) is virtually impossible if one takes serious the core idea of the transcendental method." (Natorp 1912, 9).

Subscribing to a "genetic" account of knowledge that emphasises the process character of knowledge gives the K-relation priority over its relata, to wit, the knowing subject and the object of knowledge. Both are constituted in the ongoing process of knowledge. Taken for themselves they are just abstractions from the more basic K-relation. Although it may sometimes be expedient to treat the subject of knowledge and the object of knowledge separately this separation is to be considered as a methodological device by which one may distinguish between two complementary accounts: one in which the object occupies centre stage, and one which emphasizes the role of the cognizing subject. Speaking in a Kantian framework, object-oriented accounts emphasize the role of receptivity of cognition, in particular perception, while subject-oriented, epistemic account are inclined to lay stress upon the constructive aspects of cognition. According to the Neokantian doctrine both accounts are mistaken. For the Neokantianism, ontology and epistemology are two sides of the same coin. Ontology without epistemology would be some kind of magic, which leaves unexplained how knowledge gets access to its object, while epistemology without ontology would be without content, since it denies the objectual character of cognition. Expressed in Kantian language, object-oriented approaches tend to emphasize the receptivity of cognition. According to them, cognition is essentially a passive and receptive behaviour. The thinking mind is confronted with something outside and independent of the sphere of reason. Ignoring more subtle differences this amounts to some kind of "copy-theory" or "mirror-theory" of knowledge. Subject-oriented approaches, on the other hand, emphasize the spontaneity of cognition. According to them, cognizing is essentially to be considered as a creative activity. Such a conception does not admit a "given" as a mind-independent presupposition of the cognizing process. Rather, the given ("das Gegebene") is to be conceived of as the product ("das Ergebnis") of the immanent determination of thought. Thereby, subject-oriented approaches are in danger of underestimating the resisting power of the real world in favor of the unrestricted creative power of the knowing mind. According to Natorp, employing the "transcendental method" as a guide-line, critical idealism overcomes the shortcomings and deficits of both the subject-oriented and the object-oriented accounts.

3. NATORP'S MATHEMATICAL METAPHORS

Natorp's metaphorical frame for elucidating the "relational" account of Marburg Neokantian epistemology and philosophy of science was based on two groups of metaphors, one taken from algebra and the other taken from geometry. Let us begin with his basic algebraic metaphor. According to it, knowing as the determination of the object of knowledge ("Erkenntnisgegenstand") is analogous to the process of solving a numerical equation. In order to be specific, the reader may have in mind a numerical equation like $x^2 + 2x + 1 = 0$. In other words, the object of knowledge may be considered as the "x of the K-equation:"

> If the object is to be the x of the equation of knowledge, it has to be completely determined by the perspective of knowledge, although it is that what one is looking for. In the same way as the X, Y etc. of an equation have meaning only for and in the equation, due to the meaning of the equation itself, ... the X of knowledge becomes meaningful only in the context of the inquiry. (Natorp 1910, 39)

Hence, for Natorp, as for all his fellow-philosophers of the Marburg school, the object of knowledge was not an unproblematic starting point of the ongoing process of scientific investigations, but rather as its limit.[7] The object was a problem to be solved. In various versions this equational account of knowledge can be found in virtually all of Natorp's epistemological writings. For instance, in his *Philosophischer Propädeutik* (Natorp 1903), which may be considered as a compendium of the basic doctrines of the Marburg school, he maintained that the equational metaphor expresses "the very idea of the critical or transcendental method of philosophy" (ibidem, § 7, 10). Against a one-sided and naive realism, the Critical Idealism of the Marburg School insisted that the object of knowledge was not to be considered as "given" ("gegeben") but as a problem "posed" ("aufgegeben"[8]) to the scientific investigation as suggested by the equational metaphor of knowledge quoted above. Being engaged in a solution of an equation, at the same time one does "have" and does "not have" the object represented by "x":[9] On the one hand, one does have the object, since x occurs in an equation that (hopefully) determines it completely, on the other hand, one does not have the object, since one does not know the precise value of x. In a sense, the equation promises to deliver the object but has not yet delivered it, since also the problem-solver, i. e., the scientist has to fulfil his part of the contract.

In order to bring to the fore more clearly the philosophical content of the equational metaphor it is expedient to dwell upon the mathematical or logical form of equations in some more detail. This is in line with Natorp's own approach. An equation in the sense of Natorp has the general form $F(x) = 0$. Strictly speaking, this formula is not an assertion that can be evaluated to be true or false. In order to render the formula a proposition the free variable x has to be it has to be bounded by a quantifier. It is sufficient to consider the existential quantifier $\exists x$ (there is at least one x).[10] Thereby we obtain $\exists x(F(x) = 0)$. In other words, Natorp's equational model of inquiry amounts to the introduction of variables and quantifiers. The introduction of quantifiers is tantamount to entering the realm of ontology. According to him, the objects a theory is referring to are just the values of its quantified variables (Quine 1976, § 26). I do not assert that Natorp had a clear idea of the concepts of variable, range, and quantification in the sense of modern logic. But at least his equational model may be considered as an implicit and informal precursor of Quine's thesis that ontological questions appear when one has to consider quantified theoretical statements whose parameters are determined by appropriate theoretical premises and whose "solutions" - if there are any - may be conceived as the objects the theory is referring to. Numerical equations such as $\exists x\ (F(x) = 0)$ may be considered as a kind of simplified model for them. Thereby, Quine's slogan "To be is to be the value of a variable" may be translated in Natorp's terms as the thesis that the object of knowledge exists exactly if it can be conceived as a "root" of a valid K-equation.

Conceiving Natorp's "K-equations" as quantified sentences, it is natural to ask on what sort of quantification they are based: substitutional, objectual quantification, or perhaps some intermediate form. According to the substitutional conception a variable is nothing but a slot in which one may insert just any constant. Such variables do not contend to refer to objects as their values. In the objectual interpretation the variable refers to some entities as its values, and one need not be able to characterize them by a name or a description (cf. Quine 1976, § 26). As Quine points out the substitutional and the objectual interpretation of variables are opposite to each other. In the following I'd like to consider substitutional variables and objectual variables as the two extreme poles of a spectrum. I will argue that such a "variable conception" of variables fits the dynamics of the process-oriented Neokantian account best. The dynamic of the object's development in the ongoing knowledge process may be described as an ontological move that starts from the substitutional pole and advances towards the objectual pole. To be specific, let us consider the equation $x^2 + 1 = 0$ to be interpreted as the task of determining the truth-value of the proposition

$$(*) \qquad \exists x\, (x^2 + 1 = 0).$$

Whether this proposition is true or not, depends on the range V over which the variable x is running. If one assumes that V is the domain of real numbers $\mathbf{R}$, there is no object in this domain which satisfies this equation. In this situation the inquirer has two options: either he sticks to the traditionally established domain of number objects, considering therefore (*) as false, or he attempts to enlarge the range V in such a way that the equation (*) may come out as true for some object of the new domain. As is well-known modern mathematics has chosen the latter option by accepting "imaginary" numbers $\pm\, i := \pm \sqrt{-1}$ as solutions. Without doubt, this outcome will have pleased Neokantian epistemology which always sympathized with conceptual progress of the sciences, in particular mathematics.

As is indicated already by their traditional name the ontological status of the new "imaginary" numbers $+ i$ and $-i$, and more generally of complex numbers $a + ib$, was at first considered as rather dubious. Imaginary numbers were considered as mere ficticious (but useful) constructs. They were something like theoretical terms (cf. Carnap 1974) by which the theory of numerical equations could achieve a greater unity and coherence. For instance, admitting complex numbers one could assume that every quadratic equation $x^2 + ax + b$ always had two formal solutions even if these solutions did not always define real numbers. In this stage, complex number objects had a purely substitutional character. It took some time before these constructs were recognized as genuine mathematical objects having the same ontological status as that of the familiar "real" numbers. An important step on this road to full recognition was the insight that the fundamental theorem of algebra, according to which every equation of n^{th} degree has n (possibly complex) solutions, was valid only for the enlarged domain *C* of complex numbers. Another argument for their growing ontological respectability offered Gauss's representation of complex numbers as points of the Euclidean plane. Summarizing we may say that in the course of the historical and conceptual development of mathematics the ontological status of the "imaginary" substitutions changed: they got rid of their purely instrumental status and gained recognition as fully accepted mathematical entities.[11]

Natorp's attempt to explicate objectual knowledge with the metaphorical K-equation can be conceived as an intuitive generalization of Hilbert's program of the constitution of mathematical objects by implicit definitions (cf. Hilbert 1899). In Hilbert's *Foundations* geometric objects such as points, lines, and planes are defined by implicit axioms which stipulate that certain relations exist between them. Outside the system, it does not make much sense to speak of points. Inside the system, for the determination of a point as an object of Euclidean geometry, it is necessary to determine all other kinds of geometrical objects as well. Something is a point in the context of Euclidean geometry, if and only if it fits into the relational structure of Euclidean geometry. In the metaphorical language of Natorp's K-equation this fitting may be expressed as the assertion that the conceptual object "point" may be considered as a solution of a structural K-equation. For the objects of modern structural mathematics this account has some plausibility, it appears more problematic for the objects of empirical science, at least from a modern point of view. From a Neo-kantian stance, things may have looked different. In contrast to modern philosophy of science the NeoKantian philosophy of science assumed that there is a profound similarity between mathematics and mature empirical science such as physics (cf. Cassirer 1910). For the philosophers of the Marburg school it even became difficult to draw a line between the two kinds of knowledge. Of course, they could not deny that there is a difference: otherwise they could be accused of succumbing to an unrestricted Hegelian rationalism that neglected the object of knowledge in favour of an unrestricted conceptual activity of the knowing subject. This objection also threatened Natorp's equational model: it might have been plausible to assert that a point as an object of geometry can be considered as the "solution" of some "relational equation." It is harder to understand how this approach can work for the objects of empirical science. Physical objects such as "atoms," "electrons" or "quarks" do not go into the framework of a physical theory without remainder. In this respect mathematical and physical theories are essentially different. Natorp did not ignore this fact, and complemented his equational account in such a way that it no longer fell a prey to this objection. Elaborating the equational model he pointed out that the object of knowledge – as a solution of the K-equation – was not simply a problem but an infinite task ("unendliche Aufgabe") that could be solved in finite time only approximately. Otherwise, the knowing subject would possess completely the empirical object to be known which would amount to an Hegelian rationalism that Natorp strictly rejected:

> Although we conceive, similarly as Hegel does, the object of knowledge (= *X*) only in relation to the functions of knowledge itself, and consider it ... as the *X* of the equation of knowledge, ... we understand that this "equation" is of such a kind that it leads to an infinite calculation. This means that the *X* is never fully determined by the equation's parameters *A*, *B*, *C* etc. Moreover, the series of the parameters *A*, *B*, ... is to be thought

> not as closed but may be extended indefinitely. In contrast, Hegel allows that the irrational is completely dissolvable in the rational, to wit, the lawlike determinations of thought. (Natorp1912, 19-20)

The metaphor of the K-equation is flexible enough to incorporate "infinite calculation" and approximative solvability. Natorp's "infinite calculation" already occurs in rather elementary examples: consider an equation like $x^2 - 2 = 0$ having only irrational solutions, in our case $+\sqrt{2}$ and $-\sqrt{2}$. The effective calculation of the decimal series of these numbers is a "supertask" and cannot be carried out by a finite subject in finite time. Every effective solution remains approximative.[12]

Another more sophistated example of an equation that leads to an "infinite calculation" is provided by recursive equations such as the one that is used for the calculations of the Fibonacchi numbers: $x_0 = 0$, $x_1 = 1$, $x_{n+2} = x_{n+1} + x_n$. In this way, one may define an infinite K-equation in the sense of Natorp as a series (e_n) of equations in which the parameters of the n^{th} equation are calculated as solutions of the previous equations. These examples should suffice to make clear the point Natorp wanted to make. In order to take into account the undeniable fact that the empirical realm does not go into the domain of conceptual activity of the thinking subject without remainder, the inexhaustibility of the empirical object is re-interpreted as the impossibility for the knowing subject to obtain complete knowledge of the object to be known in finite time. If this can be considered as an acceptable substitute of the inexhaustibility of the empirical object is not to be discussed here. At least, the philosophers of Marburg school believed to have countered successfully the objection that their account of the "methodically progressing" scientific knowledge was just a disguised version of Hegel's absolute knowledge.[13] For them, absolute knowledge was not something that we, as finite creatures, could ever aspire to get. Rather, the object as fully known was "the point at infinity which can never be reached but which is nothing but another expression for the always identical direction of the infinite, infinite road of knowledge." (Natorp 1910, 34). Here, then, we are entering the realm of geometric metaphors the philosophers of the Marburg School used to elucidate the unending quest for scientific knowledge. For them, the "illusion of the point at infinity" was an argument against the realist conception of knowledge according to which cognizing was to be conceived as an activity directed to some goal located outside the K-relation. Not so, they claimed, the point of infinity is an illusion caused by misunderstanding the methodological unity that intrinsically constitutes the uncompleatable object of scientific knowledge.

Summarizing we may say that Natorp's epistemology is characterized by a net of tightly interrelated metaphors and analogues mainly taken from algebra and geometry. These metaphors were designed to defend the epistemology of Neokantianism against two complementary threats: on the one hand, the critical philosophy of Natorp's neokantianism is directed against a "dogmatic" epistemology that assumes some kind of non-conceptual given as a base of knowledge. On the other hand, it is directed against a Hegelian conception of knowledge that hands over the objectual part of the knowledge relation without rest to the free-wheeling conceptual activity of the knowing subject.

4. CONCLUDING REMARKS

Once upon a time Berkeley admonished philosopher's to keep away from metaphors: "a metaphoribus autem abstinendum philosopho" but few philosophers have followed his advice. In particular, in the realm of epistemology and philosophy of science the use of metaphors is flourishing as the following brief list suffices to show:

(i) In *Conjectures and Refutations* (Popper 1963) Popper proposed to base the theory of truth approximation of theories on the spatial metaphor that "truth [is] located somewhere in a kind of metrical or at least topological space ..." (232). More precisely, he pleaded to conceptualize the notion of truthlikeness as a distance from truth.

(ii) Probably the most influential metaphor dealing with matters epistemological in the last decades has been Rorty's "mirroring metaphor" in *The Mirror of Nature* (Rorty 1989). More precisely, Rorty blames the so called representationalists as being captivated by the profoundly misleading mirroring metaphor.

(iii) In *Evidence and Inquiry* (Haack 1993) the author bases her "foundherentist" epistemology on the metaphor of the "crossword puzzle." It is not difficult to show that this metaphor has some similarity with Natorp's K-equation. Or, the other way round, Natorp's may be characterized as a foundherentist account *avant la lettre.*

(iv) McDowell's *Mind and World* (McDowell 1994) is thoroughly informed by spatial metaphors dealing with the topography of the "space of concepts" and the "space of reasons".

I think it would be too simple to dismiss all these approaches simply because they heavily depend on metaphors. The philosophical and linguistic investigations of the last decades have shown that, pace Berkeley, metaphors may well be cognitively meaningful and legitimate in philosophy and even in science (cf. Steinhart 2001). This does not mean that metaphorical assertions are exempt of criticism. Some may be better than others. The metaphors that frame Natorp's epistemology and philosophy of science are no longer ours, and his account of science has many features that appear to be obsolete from a contemporary perspective. Nevertheless, it may still be interesting to take notice of his metaphorical framework not the least as a means to better understand our own metaphorical presuppositions.

Department of Logic and Philosophy of Science, University of the Basque Country UPV/EHU, Donostia-San Sebastián, Spain

NOTES

[1] For the following nothing depends on the term "metaphor". Instead of "metaphor"one may use terms such as "analogue," "picture," or "model." The only point I want to insist on is that "metaphors" are more than rhetorical ornaments but play an important cognitive role. For a modern account of the "logic of metaphors," see Steinhart 2001.

[2] Still, metaphors are assumed to be grounded in informal and common sense experiences. For philosophical purposes, other kinds of metaphors that may be called "theory-constitutive" (Steinhart 2001, 7) may be more interesting. For a thorough discussion of this kind of metaphors, the reader may consult Steinhart's book.

[3] As a modern analogue of this epistemological debate one may consider the discussion of a viable "middle way" between coherentism and foundationalism (cf. Haack 1993, McDowell 1994).

[4] For Cohen, the key for understanding the applicability of mathematics to empirical science was the concept of the infinitesimal. He rightly considered standard logic as useless for this endeavour and set about formulating a "transcendental logic" to achieve this (cf.(Cohen 1968, 43 ff).

[5] According to Carnap's own testimony, Natorp was the Neokantian who had had the greatest influence on him.

[6] I think it is still necessary to emphasize that Neokantian epistemology can in no way be characterized as an epigonal rehearsal of Kant's account. Quite the contrary, the various Neokantian schools profoundly modified the very foundations of the Kantian edifice.

[7] How the concept of "limit" is to be understood precisely, will be dealt with later in more detail.

[8] The Marburg school, in particular Natorp, made a lot of this intricate relation between "gegeben" and "aufgegeben." For them, it was more than just a pun depending on a contingent linguistic feature of German.

[9] As is shown by the discussions to be found in Sellars and McDowell, the problem of the given is still on the agenda of contemporary philosophy (cf. Sellars 1956, McDowell 1994).

[10] Analogous considerations obtain for the universally quantified assertion $\forall x\ (F(x) = 0)$.

[11] Using the Kantian distinction between receptivity and spontaneity, one may say that the substitutional conception of variables gives spontaneity an important role: according to this approach the possible values of variables are certain symbolic constructs, whose invention takes place in the sphere of spontaneity. If these constructs turn out to be successful they are "reified," and the "hypothetical" or "fictitious" roots of the knowledge equation obtain the status of fully recognized scientific objects.

[12] Natorp's concept of approximation may be said to be based on somewhat old-fashioned idea of "external" approximation as one may call it: considering the decimal approximation of □2 we may conceive it as a converging series 1, 1.4, 1.41, 1.414 of rational numbers converging to the limiting point $\sqrt{2}$. Then clearly $\sqrt{2}$ is not among the elements of this series. Hence, against his intentions, Natorp's model suggests that the object of knowledge remains outside the approximation process. Later, Cassirer took up the analogue of numerical approximation to construe an analogy that fitted much better the basic idea of Neokantian epistemology. Cassirer based his considerations on what may be called "internal approximation." According to this modern concept the converging Cauchy series (a_n) is itself a representant of its limit $\sqrt{2}$. Using this conception of a limit of a convergent series one obtains a really compelling mathematical example for the basic Neokantian claim that "the road is the end," and this is what Natorp intended.

[13] For instance, the Neokantian Siegfried Marck belonging to the South-West school of Neokantianism, considered Natorp's attempt to avoid the Scylla of Hegelianism as unsuccessful. According to him, the alleged unity of science and philosophy, and the continuity between science, philosophy, and life as propagated by the Marburg School lead to an egalitarian "methodologism" by which the critical character of philosophy was abandoned (cf. Marck 1913, 386).

REFERENCES

Carnap, R. (1974). *An Introduction into the Philosophy of Science*. New York: Basic Books.

Cassirer, E. (1928). *Zur Theorie des Begriffs, Kant-Studien* 33, 129-136.

Coffa, A. (1991). *The Semantic Tradition from Kant to Carnap. To the Vienna Station* Cambridge: Cambridge University Press.

Cohen, H., (1968 <1863>). *Das Prinzip der Infinitesimalmethode und seine Geschichte.* Frankfurt/Main: Suhrkamp Verlag.

Friedman, M. (1999). *Reconsidering Logical Positivism.* Cambridge: Cambridge University Press.

Friedman, M. (2000). *A Parting of the Ways. Carnap, Cassirer, and Heidegger*: Chicago and La Salle: Open Court.

Haack, S. (1993). Evidence and Inquiry. *Towards Reconstruction in Epistemology*. Oxford: Blackwell.

Kinkel, W. (1923). Paul Natorp und der kritische Idealismus, *Kant-Studien* 28, 398-418.

Marck, S. (1913). Die Lehre vom erkennenden Subjekt in der Marburger Schule. *Logos* 4, 364-386.
McDowell, J. (1994). *Mind and World.* Cambridge MA: Harvard University Press.
Natorp, P. (1903). Philosophische *Propädeutik (Allgemeine Einleitung in die Philosophie und Anfangsgründe der Logik, Ethik und Pyschologie*). Marburg: N. G. Elwertsche Verlagsbuchhandlung.
Natorp, P. (1910). *Die Logischen Grundlagen der exakten Wissenschaften.* Leipzig: Teubner.
Natorp, P. (1912). Kant und die Marburger Schule, *Kant-Studien* 17, 193-221.
Natorp, P. (1929). *Philosophie. Ihr Problem und ihre Probleme, Einführung in den kritischen Idealismus,.* Göttingen: Vandenhoek und Ruprecht.
Niiniluoto, I. (1987). *Truthlikeness.* Dordrecht: Reidel.
Ortony, A. (1979). *Metaphor and Thought.* Cambridge: Cambridge University Press.
Popper, K. (1963). *Conjectures and Refutations.* London: Routledge and Kegan Paul.
Quine, W. V. O. (1976). *The Roots of Reference.* Cambridge/MA: Harvard University Press.
Rorty, R. (1979). *Philosophy and the Mirror of Nature.* Princeton: Princeton University Press.
Rickert, H. (1915). *Der Gegenstand der Erkenntnis. Einführung in die Transzendentalphilosophie* (3. ed.). Tübingen: J. C. B. Mohr (Paul Siebeck).
Steinhart, E. Ch. (2001). *The Logic of Metaphor. Analogous Parts of Possible Worlds.* Synthese Library vol. 229, Dordrecht: Kluwer.

KARL-NORBERT IHMIG

NEWTON'S PROGRAM OF MATHEMATIZING NATURE

Abstract. When Newton started his lectures on optics in 1669 as the follower of Isaac Barrow on the Lucasian chair at Cambridge he intended to develop the theory of colors as a mathematical theory of physical objects. It is in connection with this first attempt of mathematizing nature that we encounter most of the central questions with which Newton's subsequent interpreter were occupied. Why did Newton oppose to hypothetical physics? What did he mean when he contended that his principles of physics are "deduced from phenomena?" How is the relation between inductive and deductive inferences to be conceived within his methodological approach; What are Newton's sources of the methods of analysis and synthesis (or resolution and composition) that play an essential role in his investigations of colors and that also paved the way to the theory of universal gravitation. The paper attempts to discuss Newton's methodological presuppositions primarily from the perspective of his early optical studies.

Key words: analysis-synthesis, history of science, mathematic, Newton.

1. THE ORIGIN OF THE MATHEMATIZATION PROGRAM IN NEWTON'S LECTURES ON OPTICS

The success of Newton's Theory of Gravitation, which prevailed in the 18^{th} century against Descartes' mechanistic cosmology, was based last not least on the fact that he was able to found it on a successful application of mathematics to natural phenomena. Relatively little is known, however, about the background, goals, and foundations of this program. This is mainly due to Newton's reserve, both in the *Principia* and in his other principal work *Opticks*, with regard to treating such comprehensive questions. The absence of unequivocal methodological considerations leaves many questions open, at the same time provoking manifold speculations, in particular, regarding whether Newton's discoveries were more or less accidental, or whether they were the outcome of a deliberate methodological approach. While there is agreement among historians and theorists of science that linking experiment and mathematical calculus led to Newton's epoch-making discoveries, how this linkage was effected remains obscure.

One of Newton's most remarkable observations with regard to methodology is found in the "Scholium Generale" he added to the *Principia's* second edition of 1713. This is where he says:

> But hitherto I have not been able to discover the cause of those properties of gravity from phenomena, and I frame no hypotheses; for whatever is not deduced from the phenomena is to be called an hypothesis; and hypotheses, whether metaphysical or physical, whether of occult qualities or mechanical, have no place in experimental philosophy.[1]

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 241 – 261.

The strict rejection of hypotheses in experimental philosophy, however, hardly seemed to agree with how Newton practiced science. As hypotheses can be found in all editions of the *Principia*, some authors have tried to determine the meaning of the concept of hypotheses and its various forms of use more precisely.[2] This has made it possible to separate "good" from "bad" hypotheses, thus confining Newton's rejection of hypotheses merely to the bad ones.

It remains open, however, whether there is a connection between Newton's rejection of hypotheses and his program of mathematizing nature. Moreover, his statement of 1713 does not convey anything new. Already, more than 40 years before, quite similar remarks of Newton can be found in the context of the dispute about his theory of colors. The publication of this first scientific work with the title *New Theory about Light and Colors* in 1672, was followed by a debate with Hooke, Pardies, and Huygens within which Newton expressed his rejection of hypotheses several times. At least here, his methodological assumptions show remarkable continuity. This raises the question whether he rejected hypotheses on the basis of a more comprehensive methodology connected with his own program of mathematization.[3] In any case, a closer look shows that he stressed his intention to conceive his theory of colors as a mathematical natural science already in 1672, considering it *for this reason* to be a certain theory in contrast to merely probable hypotheses.

Thus, the *New Theory about Light and Colors* says:

> A naturalist would scearce expect to see ye science of those [colors] become mathematicall, & yet I dare affirm that there is as much certainty in it as in any other part of Opticks. For what I shall tell concerning them is not an Hypothesis but most rigid consequence, not conjectured by barely inferring 'tis thus because not otherwise or because it satisfies all phaenomena (the Philosophers universall Topick), but evinced by ye mediation of experiments concluding directly & wthout any suspicion of doubt.[4]

That this is not an accidental observation becomes clear when one considers that the *New Theory* was written on the basis of Newton's lectures over optics. In the fall of 1669, Newton succeeded Isaac Barrow's tenure of the Lucasian Chair at the University of Cambridge. The first topic he chose for his lectures was optics. This gave him the opportunity of publicly presenting his own considerations on the theory of colors to an audience – however small. He began his lectures in January 1670, closing them in the fall of 1672.[5] In the third lesson, he summarized the foundations of his theory of colors in the form of four propositions that were specifically not to be hypothetical ones valid only with some probability, but rather experimentally *proven propositions*: "It is affirmed that these propositions are to be treated not hypothetically and probably, but by experiments or demonstratively."[6]

It is probably no accident that he refers to the idea of a mathematical treatment of the theory of color precisely in this context:

> Thus although colors may belong to physics, the science of them must nevertheless be considered mathematical, insofar as they are treated by mathematical reasoning. Indeed, since an exact science of them seems to be one of the most difficult that philosophy is in need of, I hope to show – as it were, by my example – how valuable mathematics is in natural philosophy. I therefore urge geometers to investigate nature more rigorously, and those devoted to natural science to learn geometry first. Hence the former shall not entirely spend their time in speculations of no value to human life, nor shall the latter,

> while working assiduously with an absurd method, perpetually fail to reach their goal. But truly with the help of philosophical geometers and geometrical philosophers, instead of the conjectures and probabilities that are being blazoned about everywhere, we shall finally achieve a natural science supported by the greatest evidence.[7]

The following can be inferred from these findings: First, it is clear that Newton had explicitly formulated his program of a mathematization of nature already at the end of 1669 and the beginning of 1670. Second, this program pursues the goal of making certain of the foundation of the natural sciences in such a way that they are distinguished from merely probably hypotheses. Third, Newton, at this point, understands "hypotheses" to be physical principles or theorems that form the foundations of a theory while not being certain, holding only with a certain probability.[8]

2. A SCIENCE OF NATURE WITHOUT HYPOTHESES?

Newton's controversy with Hooke, Pardies, and Huygens following the publication of the *New Theory about Light and Colors*, is often interpreted as a dispute between the corpuscular versus wave theory of light. Two things, however, speak against this: First, the three scholars just named did not propagate a wave theory of light in the modern sense. Although their theories were not based on the behavior of individual corpuscles, they were based on that of corpuscle *streams,* which were interpreted as periodically compressing and extending longitudinal waves.[9] Second, Newton declared repeatedly that his own theory did not depend on a definite mechanist hypothesis, and, for that reason, behaved *neutrally* toward hypotheses of that kind.[10] Rather, the emphasis in the dispute between Newton and Hooke is on the following questions of *methodological* import: How extensive can the certainty about objects of nature be at all? Are there limitations of principle lying in the nature of the object? Hooke, indeed, mainly challenges Newton's claim that the latter's theory is not only a case of a merely probable hypothesis, but rather a strictly derived consequence "evinced by ye mediation of experiments concluding directly & wthout any suspicion of doubt." Whereas Hooke does not cast any doubt on the results of Newton's experiments, he objects:

> Nor would I be understood to have said all this against his theory as it is an hypothesis; for I doe most Readily agree with him in every part thereof, and esteem it very subtill and ingenious, and capable of salving all the phaenomena of colors; but I cannot think it to be the only hypothesis; not soe certain as mathematicall Demonstrations.[11]

Hooke points out the difference of principle between the certainty of mathematical proofs and the epistemic status of theories of natural science. According to this view, it is reasonable that principles of natural science can be expected only to "save the phenomena." Hence, the approach of the natural sciences must in the first place consist in sketching hypotheses and in testing them experimentally:

> I see noe reason why Mr. N. should make soe confident a conclusion that he to whome he writ did see how much it was besides the busness in hand to Dispute about hypotheses. For I judge there is noething conduces soe much to the advancement of Philosophy as the examining of hypotheses by experiments & the inquiry into

> Experiments by hypotheses, and I have the Authority of the Incomparable Verulam to warrant me.[12]

Whereas Hooke refers to Bacon in this context, the attitude he expresses here reflects an understanding of natural science theories that was very widespread in the 17th century. Descartes and Huygens, for instance, were eminent representatives of this interpretation of theory. In one of his letters to Mersenne, Descartes writes: "But to require of me Geometrical demonstrations in a matter which depends on Physics is to demand that I achieve impossible things."[13] In other words: for reasons of principle, certain proofs can be present only in mathematics, and – one would like to add – in metaphysics, but not in physics. In the realm of the natural sciences, the researcher had no other choice than to try to explain given effects by an assumed cause, then prove the correctness of the cause by the effects. Descartes uses a metaphor, among other things, to prove his view. Just like two clocks that are externally exactly similar and indicate the hours in exactly the same precise way while being completely different in their internal mechanism, in an analogous way, it should be assumed with regard to natural events that God is able to produce the visible phenomena in quite different ways. As insight into God's way of acting is inaccessible to the human mind, it is sufficient, he says, to assume one possible kind of mechanism as a hypothesis, and to show that its consequences agree with the phenomena. That it is impossible to conclude, vice versa, that the hypothesis itself is true, if it conforms to the phenomena seems to have been clear to Descartes. If a large number of phenomena, however, can be inferred from the respective hypothesis, it would be difficult to imagine "that so many things should be consistent with one another, if they were false."[14] Natural science cognition can only possess a hypothetical status, for reasons of principle.

Already at the beginning of his career, Newton seems to oppose such an interpretation of theory, saying that it would lead only to "conjectures and probabilities that are being blazoned about everywhere," without ever being able to arrive at any point to certain insights. "You see therefore how much it is besides the businesse in hand to dispute about *Hypotheses*."[15] What is it that Newton has to set against this, and what is the role of his own program of mathematization in this? Some first indications may be found in his answer to Pardies of May 5th, 1672, in which he treats the relationship between his own theory of colors and possible more far-reaching hypotheses:

> For the best and safest method of philosophizing seems to be, first to inquire diligently into the properties of things, and establishing those properties by experiments and then to proceed more slowly to hypotheses for the explanation of them. For hypotheses should be subservient only in explaining the properties of things, but not assumed in determining them; unless so far as they may furnish experiments. For if the possibility of hypotheses is to be the test of the truth and reality of things, I see not how certainty can be obtained in any science; since numerous hypotheses may be devised, which shall seem to overcome new difficulties. Hence it has been here thought necessary to lay aside all hypotheses, as foreign to the purpose, that the force of the objection should be abstractedly considered, and receive a more full and general answer.[16]

According to this description given by Newton himself, the approach in the natural sciences encompasses three steps: The first consists in carefully studying and determining the properties of things. The second step is about corroborating and affirming these properties found by experiments. The third step concerns a gradual progress from the theory confined to a certain limited range of phenomena toward more general hypotheses in order to explain these properties. This shows that Newton's intention's by no means to banish hypotheses from natural science altogether. He calls attention, however, to an epistemic distinction between the properties established experimentally beforehand and the hypotheses possibly explaining them. These properties indeed represent *limiting conditions* for possible hypotheses in the sense that all hypotheses not agreeing with these properties must be rejected. This serves to exclude the possibility that these properties could be doubted or refuted on the basis of contradicting hypotheses alone. The mere possibility of hypotheses cannot decide on the truth of things, for one can imagine many hypotheses that may cause a variety of additional problems on their part. Nevertheless, Newton realizes that when one has "directly derived" a property or a principle from the phenomena, the latter are not yet explained by this. While it is possible to explain some phenomena on the basis of this property, this property itself requires an explanation. Such an "explanation of the explanation," however, is subject to the same conditions as the original explanation itself. In the final consequence, it also has to be "deduced from the phenomena." The hypothetical status of the explanation thus refers only to a provisional state that must be overcome by additional research. The difference compared with Hooke's conception is that Newton deems this tentative character to be surmountable in principle.

Newton himself designed a model for a possible hypothesis that might explain the properties of light in this sense. In December 1675, he presented his treatise, *An Hypothesis explaining the Properties of Light discoursed in my severall Papers* to the Royal Society. He explicates his motive with the fact that his own way of speaking in a very *abstract* sense of light and the colors, that means, to abstract from more particular assumptions about the nature of light and the cause of colors, has not been understood by a larger public. This is the reason why he has decided, so-to-say for *didactical* reasons, to *illustrate* his own theory with a hypothesis:

> And therefore because I have observed the heads of some great virtuoso's to run much upon Hypotheses, as if my discourses wanted an Hypothesis to explain by, & found, that some when I could not make them take my meaning, when I spake of the nature of light & colors abstractedly, have readily apprehended it when I illustrated my Discourse by an hypothesis.[17]

In this context, he places the greatest emphasis on pointing out that this is only a possible model that, while in agreement with his own theory of light, is not being presented with a claim to truth: "This I thought fitt to Expresse, that no man may confound this with my other discourses, or measure the certainty of one by the other."[18] It must thus be noted that he does not make a claim to certainty for the correctness of the "ether hypothesis." It does not meet the criterion raised by him of having been "deduced from the phenomena."

3. ONE UNIFORM METHOD IN BOTH THE "OPTICKS" AND IN THE "PRINCIPIA"?

With regard to the certainty of the experimentally established properties of natural objects, two questions arise: Is it possible to find indications for a scientific method in Newton that might serve to justify the claim to certainty? What is the share of mathematics in this? If we begin with the last question, we must observe that the idea of basing the certainty of knowledge in the realm of natural sciences on the application of mathematics was not new. Newton himself quotes a number of examples of a successful mathematical treatment of a physical science: astronomy, geography, navigation, optics, or mechanics.[19] Well-known precursors and contemporaries of Newton like Kepler, Galileo, Descartes, or Huygens shared this view. In this connection, the following observation is revealing: In his lectures on optics, Newton criticizes three conceptions of color theory: (1) the peripatetic conception going back to Aristotle, (2) the conception of geometric optics, and (3) the mechanist conceptions (which were also directed against the Aristotelian tradition).[20]

Now precisely the geometric optics was based on applying mathematics to the physical object of light. The mechanist philosophers were also much in favor of mathematizing nature. Hence, if Newton criticizes these conceptions, opposing them with a mathematical philosophy of nature of his own, this can only mean that he intended to establish them on a new foundation. What is new in Newton's program of a mathematical philosophy of nature is that he does not share an assumption made by a number of representatives of the idea of mathematizing nature at the time. It is the assumption that nature is constituted a priori according to mathematical laws. This approach can be found, for instance, in Kepler, or in Galileo. Newton, in contrast, holds the view that mathematical determinations in themselves are of no importance for natural science. In order to convey an importance for objects or processes of natural science, it is necessary to establish physical properties of these objects or processes beforehand on which the application of mathematical concepts or structures can be based. In Newton's eyes, a mathematical philosophy of nature is thus essentially based on physical principles whose correspondence to the mathematical principles of a theory developed independently of experience is by no means established, but must yet laboriously be shown by experimental methods. Only after the possibility of such an assignation has been created, can conclusions for the field of physical objects be drawn from the mathematical structures and calculations. This means that it is only then we are justified in explaining the relations of real physical objects or processes on the basis of the relations between mathematical quantities. That Newton held this view already at the beginning of the 1670s can be seen from his answer to Hooke dated June 11th, 1672:

> I said indeed that the *Science of Colors was Mathematicall & as certain as any other part of Optiques*; but who knows not that Optiques & many other Mathematicall Sciences depend as well on Physicall Principles as on Mathematicall Demonstrations: And the absolute certainty of a Science cannot exceed the certainty of its Principles. Now the evidence by wch I asserted the Propositions of colors is in the next words expressed to be from *Experiments* & so but *Physicall*: Whence the Propositions themselves can be esteemed no more then *Physicall Principles* of a Science.[21]

The problem of founding a mathematical science of natural objects thus shifts to the problem of founding the *physical* principles. In a letter to Oldenburg, then president of the Royal Society, dated June 25th, 1672, Newton explains the methodological maxims to be applied:

> You know the proper Method for inquiring after the properties of things is to deduce them from Experiments. And I told you that the Theory wch I propounded was evinced to me, *not by inferring tis thus because not otherwise*, that is not by deducing it onely from a confutation of contrary suppositions, but *by deriving it from Experiments concluding positively & directly*.[22]

What does it mean to "deduce" something from the experiments, respectively from the phenomena? Further evidence on Newton's scientific method can be found in the second English edition of *Opticks* of 1717:

> As in Mathematicks, so in Natural Philosophy, the Investigation of difficult Things by the Method of Analysis, ought ever to precede the Method of Composition. This Analysis consists in making Experiments and Observations, and in drawing general Conclusions from them by Induction, and admitting of no Objections against the Conclusions, but such as are taken from Experiments, or other certain Truths. For Hypotheses are not to be regarded in experimental Philosophy. [...] By this way of Analysis we may proceed from Compounds to Ingredients, and from Motions to the Forces producing them; and in general, from Effects to their Causes, and from particular Causes to more general ones, till the Argument end in the most general. This is the Method of Analysis: And the Synthesis consists in assuming the Causes discover'd, and establish'd as Principles, and by them explaining the Phaenomena proceeding from them, and proving the Explanations.[23]

It is remarkable that Newton establishes an analogy here between the methods in mathematics and those in natural philosophy; the method of mathematics serving as a *model* whose application he recommends for studies of natural science as well. At the same time, the quote reveals that the method it mentions is subdivided into two partial steps. He calls these the "method of analysis," and "method of composition," or method of "synthesis." The crucial thing in applying these methods is succession. Analysis must always precede composition or synthesis. What is meant by method of analysis and by method of composition or synthesis, which obviously have their origin in mathematics, and which Newton would like to transfer to studies of natural science? He explicates the method of analysis in immediate reference to facts of nature, without mentioning their origin in mathematics in detail.

Two examples are quoted to illustrate the method of analysis. These are how to proceed from the composite to its components, and from the motions to the forces. It is not difficult to recognize that these examples obviously correspond to Newton's own research on optics and mechanics. In optics, he attempts to explain the phenomena of white light assuming that the latter is composed of (homogeneous) monochromatic rays. In mechanics, he derived the force of gravitation from the falling motion of heavy bodies on earth and from the movements of the moon, the planets, and their satellites. Generally speaking, the analytic method leads from the effects to the causes. Once the causes have been found and securely established, they can be used as a basis for the respective theory. The explanations for the various phenomena the theory is concerned with are then derived from them. This is the task of synthesis, or of the method of composition.

Additional references to the methods of analysis and synthesis are also found in the *Principia*. Thus, Newton writes in his preface to the first edition of 1687: "for the whole burden of philosophy seems to consist in this – from the phenomena of motions to investigate the forces of nature, and from these forces to demonstrate the other phenomena".[24] Despite the fact that he does not explicitly mention analysis and synthesis there, the following observation of Roger Cotes, who edited the second edition of the *Principia* of 1713, and wrote a preface to it, reveals that Newton's description points precisely to these partial steps. In his preface, Cotes emphasizes the method of experimental philosophy as a model for the entire research into nature:

> They proceed therefore in a twofold method, synthetical and analytical. From some select phenomena they deduce by analysis the forces of nature and the more simple laws of forces; and from thence by synthesis show the constitution of the rest. This is that incomparably best way of philosophizing, which our renowned author most justly embraced in preference to the rest, and thought alone worthy to be cultivated and adorned by his excellent labors.[25]

Accordingly, analysis and synthesis are explicated as partial steps of an integral method both for the objects of mechanics and of optics. In this respect, they do not differ for the two disciplines. Beyond that, Newton points out that he has applied both methods in optics, in particular in the first two books.[26] Now almost all parts of these first two books stem from observations and experiments he himself had made during the first half of the 1660s, and which he had already published from the end of the 1660s to the mid-1670s.[27] From this, it may be concluded that his methodological maxims show a remarkable continuity not only with respect to the diverse topic areas, but also with respect to his own intellectual development. This can be seen in particular from his early lectures on optics, which were much more strongly oriented toward a mathematical treatment of the theory of colors than his *Opticks*. In both their versions, they contain a separate mathematical part (analogous to Books I and II of the later *Principia*), which is entirely absent in the *Opticks*. Hence, it may be expected that the connection between his program of developing a mathematical science of nature and the methods of analysis and synthesis is much more pointedly present in his earlier lectures than in his later works. The following section will serve to begin clarifying the origin and the significance of the mathematical methods of analysis and synthesis Newton applied in his studies of nature. The last section is intended to make clear from the example of his lectures on optics in which form these methods can be transferred to facts of natural science.

4. THE ORIGIN OF THE ANALYTIC AND SYNTHETIC METHOD IN MATHEMATICS

Convincing evidence for the fact that the *Synagoge* of the Greek mathematician Pappus of Alexandria is the source of Newton's methods of analysis and synthesis is to be found in drafts for a planned preface to the *Principia's* third edition. They presumably go back to the years 1716 – 1718, dealing in particular with the methods of analysis and synthesis. It says there: "The ancients treated geometrical matters by a dual method, namely analysis and synthesis, or resolution and composition, as is clear from Pappus."[28] Besides this general clue to Pappus, we can find additional documents showing that Newton closely studied both the corresponding definition of the methods of analysis and synthesis at the beginning of book VII of Pappus' *Synagoge* (respectively Latin *Collectio*) and their applications to concrete geometric constructions. During the first half of the 1690s, Newton intended to write a more extensive work on geometry. A substantial introduction and parts of the first book have been preserved of this obviously never completed work. One of the chapters of the first book bears the heading "De Compositione & Resolutione Veterum Geometrarum." This is where we find some direct quotes from the introduction to the seventh book of the *Synagoge*.[29]

Whiteside has documented Newton's strong interest in Pappus already toward the end of the 1670s.[30] Most probably, however, Newton's familiarity with Pappus' text goes back even farther. Thus, a mathematical manuscript by Newton has been preserved whose reasons of production are unknown.[31] It originated between 1667 and 1670. Probably, this fragment was produced in the context of struggling with the methods of analysis and synthesis, as Newton explicitly mentions analysis at one point in his proofs.[32] There are several additional indirect indications for Newton's early interest in Pappus. In 1646, Frans v. Schooten edited Vieta's works. There is proof that Newton made excerpts of this edition that also contains Vieta's work *In Artem Analyticem Isagoge* from the year 1591. The *Isagoge* is intended as an introduction to Vieta's new algebraic letter calculus, and the author treats the methods of analysis and synthesis extensively at the beginning. Vieta explicates analysis as a mathematical method for finding truth that is valid for both geometric and arithmetic objects.[33] At roughly the same time as Vieta's writings, Newton also read Descartes' *Geometry,* which had been published in a Latin translation, together with an extensive commentary, by v. Schooten in 1661.[34] Descartes refers to Pappus several times in this volume. Hence, it is not improbable that Newton became acquainted with Pappus' theory of method already in the mid-1660s. This period would concur rather precisely with the "annus mirabilis" of 1666, the year in which Newton succeeded in making his decisive discoveries in the fields of mechanics, optics, and mathematics during his sojourn in his Woolsthorpe home.

The question now is how Newton interprets the methods of analysis and synthesis with regard to their contents. He explicitly quotes Pappus' description of these from Commandinus' Latin translation, the quote being in English:

> Resolution, accordingly, is the route from the required as it were granted through what thereupon follows in consequence to something granted in the composition. For in resolution, putting what is sought as done, we consider what chances to ensue, and then again its antecedent, proceeding in this way till we alight upon something already known or numbered among the principles. And this type of procedure we call *resolution*, it being as it were a reverse *solution*. In composition, however, putting as now done what we last assumed in the resolution and here, according to their nature, ordering as antecedents what were before consequences, we in the end, by mutually compounding them, attain what is required. And this method is called *composition*.[35]

While refraining to discuss here details of the interpretation of Pappus' description of analysis, which raises a number of philological problems,[36] we should like to point out one of the principal difficulties presented by the traditional Greek text. Jones translates it as follows:

> Now, analysis is the path from what one is seeking, as if it were established, by way of its consequences, to something that is established by synthesis. That is to say, in analysis we assume what is sought as if it has been achieved, and look for the thing from which it follows, and again what comes before that, until by regressing in this way we come upon some one of the things that are already known, or that occupy the rank of a first principle.[37]

It would seem that there is an obvious contradiction here. To begin with, analysis is described as a path that assumes what is sought as if it were given, drawing conclusions from that. The next sentence, however, which begins as if it were an explication of the preceding sentence, talks about the path of analysis in the sense of recourse to the conditions *from* which what is sought follows. If one draws on Commandinus' translation for comparison, it is seen that he obviously tried to smoothen this contradiction somewhat inasmuch as he says in the second sentence: "we consider what chances to ensue." His continuation, however "and then again its antecedent," does not quite seem to fit that.

How did Newton respond to this difficulty of interpretation? Surprisingly, he did not respond at all. Either he had not become aware of the contradiction, as it had already been softened in Commandinus' translation, or Newton associated a content-related interpretation with the methods of analysis and synthesis in which the contradiction does not occur in the first place. Newton comments this passage as follows:

> By these words you should understand that a general method for solving problems was known to the ancients, and that this method consisted in its greatest part in the resolved locus, proceeding by means of resolution and composition jointly: by resolution composition is attained and in the fullness of composition all that is geometrical is accomplished; solution is, however, the opposite of resolution in that it may not be had till all trace of resolution be removed from start to finish by means of a full and perfect composition.[38]

To begin with, it is notable that Newton understands analysis and synthesis, respectively resolutio and compositio, as a *general method of problem solving* that need not be confined to the object of geometrical problems and theorems. It is also remarkable that he does not conceive of the methods of resolution and composition as directions opposed to one another. Instead, he opposes resolutio *and* compositio *jointly* to solutio:

> Solution therefore differs from resolution and composition inasmuch as the latter are means and procedures for discovering through the resolved locus, the former is the enddiscovery at which the whole process terminates. Resolution and solution are two totally opposite extremes, and the expert makes his way in stages from one to the other by composition.[39]

He considers both resolutio and compositio to be means and methods of discovery, opposing them to the final solution of a problem. Thus, the above quote says that Pappus' method contains an approach "by means of resolution and composition jointly." In an earlier version of this text, Newton specifically pointed out that resolutio and compositio were occasionally intermingled:

> Solution is contrary to resolution: the expert proceeds from beginning to end and from resolution to solution by means of composition. Resolution can be intermingled with composition, but with solution not at all.[40]

Further evidence for Newton's interpretation of Pappus' methods of resolutio and compositio can be found in the drafts for a preface to the *Principia's* third edition:

> The analysis of the ancient geometers seems to have consisted in the deduction of consequences from givens until the thing sought should result. [...] Where, however, the thing sought would not easily ensue from the givens, they either looked for lemmas or porisms through which some new given might be gatherable, or assumed unknowns as givens so that thereby they might gather some given as though it were unknown, [so that they see what might ensue] and at length by inverting the sequence of argument deduce the thing sought from whatever relationship between the givens and the sought.[41]

This quote reveals that Newton initially ascribes a certain direction to the analysis of the geometers of antiquity, namely, the derivation of consequences from the "given," respectively from the "data." What is meant by a given? Givens are quantities, relations between quantities, points, lines, areas, and so forth, as well as relationships between such objects that are established by the principles or axioms of geometry and by the prescribed constructions. Analysis investigates which further relations can be derived from this in order to find the sought, which consists, in the case of geometric problems, in certain objects to be constructed that shall satisfy certain conditions. If the conclusions from the initial construction lead to the desired result, that is to say, to a theorem already known, or if they lead to an assumed principle, analysis is accomplished. If one encounters difficulties on this path, however, one is compelled to draw on auxiliary theorems or auxiliary constructions. One of their properties is that they introduce new objects, or relations between objects, into the problem that are based on relations between what is given ("data") and what is sought ("quaesita"). In that case, what is sought is assumed as given, and further conclusions are drawn in connection with the initial construction and the axioms from these relations, until one encounters a known theorem or principle. If this is the case, then what is sought can be deduced from these relations by inverting the argumentation.

Obviously, Newton counts this deduction, which runs in the inverse direction, that is, from the principles and axioms to the determined construction, as still belonging to analysis. This would mean that analysis, according to Newton's interpretation, includes two directions: one from the initial construction to the principles, and another back from the principles to this construction. This would also explain why Newton holds the view that resolutio and compositio are frequently intermingled. "Deduction," in this context, obviously possesses another meaning, as a syllogistic derivation of propositions. For it is always possible that many different relations can be derived from a given construction containing determinate relations between the objects given and the objects sought. From these different possibilities, it is not clear beforehand which of them is conducive to the goal, that means, which makes it possible to determine the objects sought on condition of the objects given. In that case, one would be compelled to first examine each of them separately, a task which may require further auxiliary constructions. Possibly, further intermediate steps would be necessary in the course of these studies that sometimes involve an inversion of the direction of argumentation. Hence, the result is a complex, branched-out procedure that leads eventually to the determination of what is sought or to the solution of the problem. Only after this has been achieved, is one in a position to know which of these branches one must follow in the inverted direction in order to receive an unequivocal determination of what is sought from what is given.

What is the function of synthesis in this interpretation? Newton answers this question as follows:

> The ancients used in mathematical matters to practice a dual method, analysis and synthesis, or composition and resolution. Through analysis they discovered propositions, and through synthesis they demonstrated them once found – and when these were not yet demonstrated they did not admit them into geometry; for geometry's title to praise lay in the utter certainty of its matters. And on that account I have in the books which follow composed the propositions found out by analysis in order to render them absolutely certain and so, because of their certainty, worthy to be admitted into geometry.[42]

It is only synthesis that conveys the necessary certainty to the discovery of analysis that Newton understands as an "ars inveniendi," and provides the "solution" or "demonstration" of the theorems or problem solutions discovered. Whereas analysis, for Newton, is no method that can be applied mechanically, but rather depends on the skill and experience of the mathematical expert, synthesis must meet the requirement of being understood by all who are sufficiently familiar with the basics of geometry.[43]

5. THE APPLICATION OF ANALYSIS AND SYNTHESIS TO NATURAL SCIENCE

How can the methods of analysis and synthesis be transferred to facts of natural science?[44] After a brief look back to the concept of Newton's program of mathematization, three central points can be retained: First, the application of mathematics to objects of nature is based on *physical* principles, that must be ascertained experimentally. Physical meaning is not due to the mathematical determinations in themselves, but only in connection with these physical principles. Second these principles have the character of physical properties of natural objects that can be drawn on as causes to explain a certain range of phenomena. Nothing has been said with this about the causes of these principles themselves, that is, about the causes of the causes. For these, it is also true that they must be "deduced from the phenomena." Third, Newton understands the path leading from the effects to the causes as analysis, adhering to the ancient geometers in this. What follows intends to use Newton's early writings on optics to elaborate the parallels between mathematical and philosophical analysis which are significant for Newton's experimental method.

What is the analogy with the method of geometric analysis and synthesis, in Newton's methodological approach in the natural sciences? What is the "given," and what is the "sought?" Let us treat the "given" first. Referring to the example of optics, one is confronted with what Newton describes as "the celebrated Phaenomena of Colors" in his *New Theory about Light and Colors*. These phenomena stand as the "given" at the outset of Newton's considerations. They have not been observed accidentally, but rather are produced by means of a well-devised experimental setup, and form the beginning both of his first lectures on optics and of his *New Theory about Light and Colors*.[45] Sunlight falls through a shutter's round aperture into a darkened room. A prism is placed immediately behind the aperture. The sunlight is led through the prism, and projected on the opposing wall of the room or on a screen. The screen shows an elongated, oval image that is several times longer than broad. In addition, one beholds colors that follow one another from one end of the image to the other in the order of red, yellow, green, blue, and violet. Obviously this spectrum of colors, in connection with the experiment's setup, must be considered to be the "given." What is remarkable in this context is that both the given of geometry and the given of natural science are nothing one can simply find and observe passively, but that both cases are about something intentionally produced and constructed.

What now is the "sought?" In geometry, what is sought is a possibility of constructing (provided it is a "problematical analysis") quantities or relations between quantities *depending on* "given" quantities or relations. Similarly, a physical property is sought in the natural sciences: one that proves to be the determining factor for the production of certain phenomena or processes of nature. How does one find such a property or such a principle? One does this by deriving conclusions from the phenomena, in connection with a tentatively assumed principle, until one meets a principle or theorem already known or possibly even a contradiction.[46] Newton indicates such a principle that has hitherto been assumed to be valid, namely, the *geometric* sinus law of refraction based on the *physical* assumption that the white sunlight is homogeneous. This means that *all the* rays of the white sunlight having the same angle of incidence on a refracting plane will have the same angle of reflection. In his next step, Newton draws conclusions from the phenomenon of the elongated spectrum of colors in connection with this assumed physical principle. In doing so, however, he is confronted by a *contradiction*. For he computes precisely *one* position of the prism for which the *sum* of the two angles of incident at the prism's two sides that refract the rays is exactly equal for *all* rays.[47] If the assumed physical principle of homogeneity of the white sunlight were valid, the image should appear circular *precisely in this case*. After Newton had set up his experiment precisely according to these prerequisites, he had to note, however, that the image did not appear circular, but oval. Hence, the physical principle assumed by geometric optics contradicts the phenomena, and must therefore be rejected.[48]

Hence, the search for a physical principle that is in harmony with the phenomena must be continued. The phenomenon of the elongated form of the colors spectrum remains the point of departure. Newton's new assumed principle is that the white sunlight is composed of different (homogeneous) rays having different degrees of refrangibility. He now derives conclusions from this phenomenon in connection with this new principle. He begins by describing the elongated image as a plane geometrical figure that is delimited below and above by two parallel straight lines and at the edges by two semicircles. This is a conscious idealization; for Newton was aware of the fact that the oval's edges are not delimited by exact mathematical lines in reality. Pursuing this, he develops four additional assumptions that are of purely ideal, theoretical nature and by which the spectrum's form can be explained on the basis of assuming the principle of the heterogeneity of the white sunlight. These assumptions include already proven geometrical theorems, as well as infinitesimal and continuity assumptions.[49] One could compare them to the lemmata and porismata that are to assist the geometer in the analysis of difficult problems and that have been collected in the "locus resolutus."

(1) Rays of white sunlight, which are equally refracted in case of equal angle of incidence when passing the prism, yield an approximately circular image if considered separately. (2) If one now imagines additional rays of sunlight, which have equal refraction compared to one another, but which differ in their refraction from the rays of sunlight considered before, these will again yield a circular image that, however, will occupy a place within the oval different from that of the first image. (3) In a further step, Newton extends his reflections to infinitely many rays. Imagine infinitely many other rays whose quantity of refraction is continually larger or smaller than that of the preceding one. These will thus describe an infinite number of circles that fill out the spectrum in its length. The first three steps thus explain the elongated form of the spectrum image and its lateral delimitation by semicircles. (4) As the circular images of the separate rays all possess approximately the same extent, the lines delimitating the spectrum above and below are approximate straight lines and run parallel to one another. This would also explain the delimitation of the image above and below.

These theoretically derived relationships between the given and the sought assumed to be given are subjected to another experimental test in what follows. This is done by examining the relationship between the spectrum's breadth and length in more detail. This relationship proves to be the decisive relation on which Newton's theoretical explanation is based. In his first experiment, he had established that the relation between the image's breadth and length never went below 1 : 4. If one reduces the aperture for the light by a fifth, the relation grows to 1 : 13.5. Corresponding observations can be made if one increases the distance between prism and screen or inserts a convex lens into the experimental setup. In his *Opticks*, Newton later describes experiments in which he was even able to attain relationships of 1 : 60 or 1 : 70.[50]

If one summarizes Newton's comprehensive conduct of proof, it becomes clear that in order to test and confirm the new physical principle, the application of the method of analysis involves a multitude of experiments mutually supporting and supplementing one another. The claim to certainty for this principle is finally based on the presence of a lower boundary for the mathematical relation between the spectrum's breadth and length that proves invariant under all possible variations of the experimental conditions. For Newton, this invariance of the *mathematical* relation is an indication that this cannot be a phenomenon dependent on random circumstances. He obviously judges the presence of such an invariant mathematical regularity to be a clue to the existence of a real physical property. For he considers it to be proven that the property of the different refrangibility of the rays emerges from a "previous disposition of the rays" and is based on certain laws.[51]

What can be inferred from this for the application of the analytic method in the natural sciences? One of the first things to be remarked is that both deductive and inductive elements appear in it, and they are closely interwoven. Newton systematically varies the experimental conditions to establish functional dependencies of the phenomena on the factors relevant for their production. In doing so, he recurs to auxiliary assumptions like, for instance, the principle of continuity, considerations of limiting values, or theorems of geometrical optics. To these assumptions one can attribute in an analogous way the function that the lemmata and porisma had within the analytical method of the ancient geometers. The goal of analysis in natural science consists in discovering and confirming constant quantitative relations that indicate an invariant physical property. Inductive generalization, which belongs to the method of analysis, has the character of exploring constant functional dependencies between factors by means of continuous variation. This obviously cannot be done with a single experiment. Rather, Newton can be observed to repeat experiments to investigate different experimental conditions separately. Techniques of eliminating disturbing factors or increasing observed weak effects are of particular importance here. Where Newton describes *one* experiment, it is always a case of a multitude of systematically connected experiments that are derived from *one principal experiment.*

All the aspects listed here must be taken into account if one wishes to understand on what Newton bases his claim that his theory's principles were not hypotheses, but rather "deduced from the experiments." "Deduction" must not be equated here with a syllogistic derivation, but implies both an inductive generalization by means of variation and a critical examination of the regularities thus established by experimental *tests.* This means that analysis does not represent a one-way street in the sense of moving from the phenomena toward the principles, but that it includes the opposite direction from the principles to the phenomena as well. This interpretation is confirmed by a quote from Newton, which presumably was written in 1700 as a draft to an intended preface to his *Opticks*:

> As Mathematicians have two Methods of doing things which they call Composition & Resolution & in all difficulties have recourse to their method of resolution before they compound so in explaining Phaenomena of nature the like methods are to be used & he that expects success must resolve before he compounds. For the explications of Phaenomena are Problems much harder then those in Mathematics. The method of Resolution consists in trying experiments & considering all the Phaenomena of nature relating to the subject in hand & drawing conclusions from them & examining the truth of those conclusions by new experiments & drawing new conclusions (if it may be) from those experiments & so proceeding alternately from experiments to conclusions & from conclusions to experiments untill you come to the general properties of things. Then assuming those properties as Principles of Philosophy you may by them explain the causes of such Phaenomena as follow from them: wch is the method of Composition.[52]

Newton bases the possibility of treating the theory of colors mathematically on the correlation between the property of refrangibility and the rays' disposition to manifest a determinate color. Assigning the color scale to the indices on the scale of refraction creates the basis for introducing a metric into the theory of colors. It must be noted that the possibility of this assignation is based on the physical property of white light that it is composed of heterogeneous, differently refrangible rays.

After the analysis has come to a (provisional) conclusion, synthesis follows: "And the Synthesis consists in assuming the Causes discover'd, and establish'd as Principles, and by them explaining the Phaenomena proceeding from them, and proving the Explanations."[53] The interface of the transition from analysis to synthesis can be clearly determined in the lectures on optics. At the beginning of the seventh lesson, Newton observes that he had laid, in the previous lessons, the foundations of his own theory of colors by means of which the common color phenomena can be explained: "Thus far we have erected the foundation whereby the common appearances of colors produced by prisms can be most certainly ["certissime"] explained."[54] He answers the question whether a repeated treatment of the phenomena of the prismatic colors was not superfluous at this point by pointing out that this step was necessary because of the method he adhered to: "that consequently we may retain the proposed method, namely, to determine them scientifically from principles previously demonstrated."[55] This method apparently contains two steps: (1) of deriving the principles (causes) from the phenomena, and (2) of explaining the phenomena by means of these principles. It is not difficult to recognize that these two steps are analysis and synthesis. This demonstrates that Newton consciously adhered to a certain scientific method already in 1669/70, when he began to draft his lectures on optics, and that this consciousness did not emerge only as a result of his disputes about the *New Theory* in the years 1672/73.

What is new in synthesis, as opposed to analysis? Is this merely a case of a linear inversion of the direction of argument? It must be recalled here that the physical principles that emerged from analysis, within the scope of Newton's idea of mathematizing nature, had the function of serving as a foundation for a mathematical treatment of physical objects. This meant, in particular, that only such properties of natural objects were appropriate for these principles on which a metric could be founded. It was only after these properties had been ascertained that it became possible to attribute a physical meaning to the mathematical theorems and proportions. This is precisely what Newton does in pursuing the path of synthesis. The explanation of the elongated form of the color spectrum that follows analysis is characterized by the fact that attention is drawn specifically to the discrepancy between the spectrum image's ideal mathematical form and its real shape, and that the causes of this deviation are determined. This points to another interesting aspect of Newton's idea of mathematization. For Newton, mathematization precisely does *not* mean that the physical principles *exactly copy* the mathematical laws and determinations, but rather that they help to explain the *deviations* from such ideal laws. This explanatory function of synthesis is what is new, as compared to analysis. This is where the synthesis of the natural sciences is distinct from the synthesis of ancient geometry insofar as the latter did not have such a problem at all, being *exclusively* concerned with ideal mathematical objects.

The different character of *Opticks* and *Principia* is apparently due to the fact that, in optics, Newton did not manage to achieve the unification with regard to explaining phenomena that he achieved in mechanics. Essentially, it was the phenomena of refraction and reflection that he was able to unify under common principles. Newton himself probably saw quite clearly that his project of a mathematical theory of colors remained incomplete. But he also was clearly aware of the fact that science is not the isolated work of an individual, but must be conceived of as a social and historical process:

> But if without deriving the properties of things from Phaenomena you feign Hypothesis & think by them to explain all nature, you may make a plausible systeme of Philosophy for getting your self a name, but your systeme will be little better than a Romance. To explain all nature is too difficult a task for any one man or even for any one age. [...] Tis much better to do a little with certainty & leave the rest for others that come after you then to explain all things by conjecture without making sure of any thing.[56]

Abteilung Philosophie, Universität Bielefeld
Institut für Philosophie, Universität Dortmund

NOTES

[1] Newton, Mathematical Principles of Natural Philosophy. Vol. II, 547.

[2] Cf. Cohen (1956, 575-584); Cohen (1966); Hanson (1970); Shapiro (1989). This is only a small selection from the rich literature on Newton's concept of hypotheses.

[3] Kargon explains Newton's rejection of hypotheses from the latter's opposition to the ideal of a "hypothetical physics" propagated at the time by Descartes, Hobbes, and Gassendi that was unable to proceed beyond plausible (i. e., probable) hypotheses because of its own essence and because of its object's nature. Instead, Newton, continuing on from Francis Bacon and his own teacher Isaac Barrow, demanded a form of certainty going beyond mere hypothetical probability for the natural sciences as well. Barrow, in particular, had called for such a certainty, attempting to found it on the application of mathematics. Cf. Kargon (1965).

[4] Newton, Correspondence, Vol. I, 96 f.

[5] Two manuscripts of Newton's lectures on optics have been preserved. As date of authorship for the first, shorter manuscript, the time between the end of 1669 and the end of 1671 can be presumed. The second, longer manuscript, which obviously represents an improved version of the first, was probably finished subsequently in February 1672. Meanwhile, the excellent edition and translation of both manuscripts by Alan E. Shapiro is available: Shapiro, A. E.. The Optical Papers of Isaac Newton. Vol. I: The Optical Lectures 1670 – 1672, Cambridge 1984 (in the following quoted as: Newton, Optical Papers I).

[6] Newton, Optical Papers I, 87.

[7] Op. cit., 89.

[8] Shapiro comments at this point (Newton, Optical Papers I, 28): "Newton [...] makes a powerful plea for mathematical natural science, while offering his new, mathematical theory of color as an example of the value of mathematics in natural philosophy. Thus, at the beginning of his career he had already clearly formulated a program for the reform of natural science that would come to full fruition in his Philosophiae naturalis principia mathematica, that is, The Mathematical Principles of Natural Philosophy."

[9] Concerning the theories of light of the 17th century Shapiro has good reason to prefer, not to speak of a contrast between wave and corpuscular theory, but instead of one between "continuum theory" and "emission theory." Cf. Shapiro (1973, 136).

[10] Cf. for example Newton, Correspondence, Vol. I, 174. Here, Newton admits that it was possible to explain the properties of light he had found out not only by one, but "by many other Mechanicall Hypotheses." This is why he had preferred "to decline them all, & speake of light in generall termes, considering it abstractedly as something or other propagated every way in streight lines from luminous bodies, without determining what that thing is."

[11] Op. cit., 113.

[12] Op. cit., 202 (Letter by Hooke to Lord Brouncker of June 1672).

[13] Quoted according to Sabra (1967, 23).
[14] Descartes, Principles of Philosophy, Part IV, § 205, 287. For the clock-metaphor cf. op. cit., § 204, 286. Cf. for this also Laudan (1981, 27-58).
[15] Newton, Correspondence, Vol. I, 177 (Newton's answer to Hooke dated June 11th, 1672).
[16] Cohen (1958, 106).
[17] Newton, Correspondence, Vol. I, 363.
[18] Op. cit., 364.
[19] Newton, Optical Papers I, 86 f.
[20] Op. cit., 46-49; 80-85.
[21] Newton, Correspondence, Vol. I, 187 f. Cf. for this also Mamiani (1976, 107 f.).
[22] Newton, Correspondence, Vol. I, 209.
[23] Newton, Opticks, 404 f. A corresponding text already appeared in an abbreviated form in the first Latin edition of Opticks of 1706.
[24] Newton, Mathematical Principles of Natural Philosophy. Vol. I, Newton's Preface to the First Edition, XVII f.
[25] Op. cit., Cotes' Preface of the Second Edition, XX f.
[26] Newton, Opticks, 405: "In the two first Books of these Opticks, I proceeded by this Analysis to discover and prove the original Differences of the Rays of Light in respect of Refrangibility, Reflexibility, and Color, and their alternate Fits of easy Reflexion and easy Transmission, and the Properties of Bodies, both opake and pellucid, on which their Reflexions and Colors depend. And these Discoveries being proved, may be assumed in the Method of Composition for explaining the Phaenomena arising from them: An Instance of which Method I gave in the End of the first Book."
[27] Cf. Hall (1995, 33-83).
[28] Newton, Mathematical Papers, Vol. VIII, 449.
[29] Newton, Mathematical Papers, Vol. VII, 248-251; 304-311. According to Whiteside, Newton used the second edition of the Latin translation by Commandinus that was published in a first edition by Manolessi in 1610: Pappi Alexandrini Mathematicae Collectiones a Federico Commandino Urbinate in Latinum conversae, & Commentarijs illustratae, Bologna 1660. Cf. Newton, Mathematical Papers, Vol. VIII, 449, note 21. Commandinus translates the Greek terms of "analysis" and "synthesis" by "resolutio" und "compositio." This explains why Newton uses these expressions synonymously where he draws on Pappus' Latin translation. On the significance of the methods of analysis and synthesis in mathematics cf. Otte/Panza (1997).
[30] Cf. Newton, Mathematical Papers, Vol. IV, 218: "What the geometrical manuscripts now published do reveal is that in his middle-thirties Newton developed an acute interest in the Συναγωγη ([Mathematical] Collection) of the late Alexandrian mathematician Pappus, minutely studying its seventh and eight books in one or other of the available editions of Commandino's Latin translation."
[31] Cf. Newton, Mathematical Papers, Vol. II, 450-517.
[32] Op. cit., 493: "This construction can be demonstrated in the manner of problem 4 but for the sake of brevity and variety I prefer to employ the following analysis."
[33] Cf. Klein (1968, 154-178); Panza (1997, 401-405).
[34] Cf. Whiteside (1967, 73). Cf. also Westfall (1980, 106).
[35] Newton, Mathematical Papers, Vol. VII, 307; cf. also op. cit., 249. For the critical edition both of the Greek original and its English translation, cf. Pappus of Alexandria, Book 7 of the Collection, Part 1. Introduction, Text and Translation, ed. by A. Jones, New York et. al. 1986, 82 f.
[36] It should be noted that the Greek original raises grave problems of interpretation. Since the works of R. Robinson and F. M. Cornford from the 1930s, the discussion of these problems has focused mainly on the issue of the "direction" of analysis and synthesis. Does the analysis of a theorem to be proved mean to descend to further conclusions from this theorem, or vice versa to ascend to its conditions? Robinson pleaded in favor of the former possibility, whereas Cornford considered the second option plausible in connection with his own studies of Platon's dialectic. To this problem has to be added the further difficulty of how to link this general description of the methods at the beginning of the seventh book with Pappus' subsequent geometrical practice. The latter attempt has been made in particular in Hintikka/Remes (1974), Behboud (1994), and Mäenpää (1997). A brief summary of the discussion and additional references are to be found, for instance, in Behboud (1994, 53-57).

[37] Pappus of Alexandria, Book 7 of the Collection, Part 1, 82.
[38] Newton, Mathematical Papers, Vol. VII, 307. "Resolved locus" respectively "locus resolutus" is the translation of the Greek term "τόπος ἀναλυόμενος". This obviously meant a branch of the mathematics of antiquity concerned with geometrical analysis. Pappus enumerates a number of books belonging to this branch, like Euclid's "Data" and "Porismata," and further works by Apollonius, Aristaeus, and Eratosthenes. This branch obviously dealt with solving difficult geometrical problems and was intended only for those who were already sufficiently familiar with the elements and foundations of geometry.
[39] Op. cit., 309.
[40] Op. cit., 308, note 69.
[41] Newton, Mathematical Papers, Vol. VIII, 443/445. Cf. op. cit., 444 f., note 3: "The ancients in the resolution of problems used first to gather from the givens whatever might come to ensue, [...]. If by this method they were able to collect what was sought, the problem was resolved; but if not, they used to assume what was sought as though it were a given in order that they might thence gather some given as though it were sought, and so from the (connection) relationship between given and sought deduce the sought by going back."
[42] Op. cit., 451.
[43] Op. cit. 449: "Propositions in geometry, however, ought to be propounded in such a way that they may be appreciated by the great majority and thus most impress the mind with their clarity, and they need consequently to be synthetically demonstrated. Analysis is useful for finding out truths, but the certainty of a finding ought to be attested through the composition of a demonstration, and so made as transparent, clear and manifest to all as it is possible."
[44] The first authors, and to my knowledge the only ones, to call attention to the importance of the geometrical method of analysis and synthesis for Newton's natural science research, were Hintikka und Remes. They base this thesis on their interpretation of geometrical analysis as figural analysis, which shows, in its confinement to spatial constructions, some parallels to the analysis of experimental situations. Cf. Hintikka/Remes (1974, 106): "Newton, like any experienced mathematician, is thinking of the geometrical analysis as an analysis of figures, that is to say, as a systematic study of the interdependencies of the geometrical objects in a given configuration, including both the 'known' (controllable) and 'unknown' (uncontrollable) factors. [...] From this, it was but a short step to the idea that an experimental setup represented a kind of analytical situation, too, in that what is happening in a typical controlled experiment is a study of what depends on what in it - and hopefully also precisely what mathematical relationships these dependencies exemplify."
[45] Cf. Newton, Optical Papers I, 50. In what follows, I shall confine myself to Newton's presentation in his first lecture series of 1669 – 1670 that Shapiro designates as "Lectiones opticae." Cf. for this also Newton, Correspondence, Vol. I, 92.
[46] Newton gives the respective quote from Pappus in Newton, Mathematical Papers, Vol. VII, 309: "if we meet with what evidently cannot be done, then the problem will be likewise impossible."
[47] This position of the prism is derived in an exact mathematical way on the basis of comprehensive geometrical considerations in his "Lectiones opticae." This is followed by practical hints as to the simplest way of realizing this position experimentally. Cf. Newton, Optical Papers I, 52-63. In his New Theory, he intentionally left out this derivation, a fact that gave rise to the misunderstanding that he intended to claim the spectrum's image must appear circular in any position of the prism under the assumption of the homogeneity of the white sunlight.
[48] Op. cit., 61.
[49] Cf. for this op. cit., 62-65.
[50] Newton, Opticks, 69.
[51] Newton, Optical Papers I, 74 f.
[52] U. L. C. Ms. Add. 3970.5; quoted according to Guerlac (1977, 205f.). Cf. for this also U. L. C. Add. 3970.3. Folio 480v in McGuire (1970, 185).
[53] Newton, Opticks, 404f.
[54] Newton, Optical Papers I, 145.
[55] Op. cit., 145. The corresponding interface within the second series of lectures is found in part II, lecture 10. Cf. op. cit., 523.
[56] U. L. C. Ms. Add. 3970.5; quoted according to Guerlac (1977, 206).

REFERENCES

Behboud, A. (1994). Greek Geometrical Analysis. *Centaurus* 37, 52-86.

Blake, R. M., Ducasse, C. J., & Madden, E. H. (1966). *Theories of Scientific Method: The Renaissance through the Nineteenth Century*. Seattle/London: University of Washington Press.

Bricker, Ph., & Hughes, R. I. G. (Eds.) (1990). *Philosophical Perspectives on Newtonian Science*. Cambridge (Mass.)/London: M. I. T. Press.

Cohen, I. B. (1956). *Franklin and Newton*. Cambridge (Mass.): Harvard University Press.

Cohen, I. B. (Ed.) (1958). *Isaac Newton's Papers & Letters on Natural Philosophy and Related Documents*. Cambridge (Mass.): Harvard University Press.

Cohen, I. B. (1966). Hypotheses in Newton's Philosophy. *Physis* 8, 163-184.

Descartes, R. (1983). *Principles of Philosophy*. Transl. by V. R. Miller & R. P. Miller. Dordrecht/Boston/London: Reidel,

Engfer, H.-J. (1982). *Philosophie als Analysis. Studien zur Entwicklung philosophischer Analysiskonzeptionen unter dem Einfluß mathematischer Methodenmodelle im 17. und frühen 18. Jahrhundert*. Stuttgart/Bad Cannstatt: frommann-holzboog.

Guerlac, H. (1977). Newton and the Method of Analysis. In H. Guerlac, *Essays and Papers in the History of Modern Science*. Baltimore/London: The Johns Hopkins University Press, 193-216.

Hall, A. R. (1995). *All was Light: An Introduction to Newton's 'Opticks'*. Oxford: Clarendon Press.

Hall, A. R., & Hall, M. B. (Eds.) (1962). *Unpublished Scientific Papers of Isaac Newton*. Cambridge: Cambridge University Press.

Hanson, N. R. (1970). Hypotheses Fingo. In R. E. Butts, & J. W. Davis (Eds.), *The Methodological Heritage of Newton*. Oxford: Blackwell, 14-33.

Hintikka, J., & Remes, U. (1974). *The Method of Analysis*. Its Geometrical Origin and Its General Significance. Dordrecht/Boston: Reidel.

Kargon, R. (1965). Newton, Barrow, and the Hypothetical Physics. *Centaurus* 11, 46-56.

Klein, J. (1968). *Greek Mathematical Thought and the Origin of Algebra*. Cambridge (Mass.)/London: M. I. T. Press.

Laudan, L. (1981). *Science and Hypothesis*. Historical Essays on Scientific Methodology. Dordrecht/Boston/London: Reidel.

Mäenpää, P. (1997). From Backward Reduction to Configurational Analysis. In M. Otte, & M. Panza, (Eds.), *Analysis and Synthesis in Mathematics*. Dordrecht/Boston/London: Kluwer Academic Publishers, 201-226.

Mamiani, M. (1976). *Isaac Newton filosofo della natura*. Le lezioni giovanili di ottica e la genesi del metodo newtoniano. Firenze: La Nuova Italia Editrice.

McGuire, J. E. (1970). Newton's 'Principles of Philosophy': An Intended Preface for the 1704 Opticks and a Related Draft Fragment. *The British Journal for the History of Science* 5, 178-186.

McGuire, J E., & Tamny, M. (1983). *Certain Philosophical Questions: Newton's Trinity Notebook*. Cambridge: Cambridge University Press.

Newton, I. (1959-1977). *The Correspondence of Isaac Newton*. Ed. by H. W. Turnbull, J. F. Scott, A Rupert Hall, Laura Tilling. 7 Vols. Cambridge: Cambridge University Press.

Newton, I. (1967 – 1981). *The Mathematical Papers of Isaac Newton*. Ed. by D. T. Whiteside. 8 Vols. Cambridge: Cambridge University Press.

Newton, I. (1969). *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and his System of the World*. Ed. by F. Cajori. Vol. I + II. New York: Greenwood.

Newton, I. (1952). *Opticks or A Treatise of the Reflections, Refractions, Inflections & Colors of Light*. Based on the fourth edition London 1730. New York: Dover Publication.

Newton, I. (1984). *The Optical Papers of Isaac Newton*. Ed. by A. E. Shapiro. Vol. I: The Optical Lectures 1670 – 1672. Cambridge: Cambridge University Press.

Otte, M. (1994). *Das Formale, das Soziale und das Subjektive*. Eine Einführung in die Philosophie und Didaktik der Mathematik. Frankfurt a. M.: Suhrkamp.

Otte, M. (1997). Analysis and Synthesis in Mathematics from the Perspective of Charles S. Peirce's Philosophy. In M. Otte & M. Panza (Eds.), *Analysis and Synthesis in Mathematics*. Dordrecht/Boston/London: Kluwer Academic Publishers, 327-362.

Otte, M., & Panza, M. (1997). Mathematics as an Activity and the Analytic-Synthetic Distinction. In M. Otte & M. Panza (Eds.), *Analysis and Synthesis in Mathematics*. Dordrecht/Boston/London: Kluwer Academic Publishers, 261-271.

Otte, M., & Panza, M. (Eds.) (1997). *Analysis and Synthesis in Mathematics*. History and Philosophy. Dordrecht/Boston/London: Kluwer Academic Publishers.

Panza, M. (1997). Classical Sources for the Concepts of Analysis and Synthesis. In M. Otte, & M. Panza (Eds.), *Analysis and Synthesis in Mathematics*, Dordrecht/Boston/London: Kluwer Academic Publishers, 365-414.

Pappus of Alexandria (1986). *Book 7 of the Collection*. Part 1. Introduction, Text and Translation, Part 2. Commentary, Index, and Figures. Ed. by A. Jones, New York/Berlin/Heidelberg/Tokyo: Springer Verlag.

Sabra, A. I. (1967). *Theories of Light from Descartes to Newton*. London: Oldbourne.

Shapiro, A. E. (1973). Kinematic Optics: A Study of the Wave Theory of Light in the Seventeenth Century. *Archive for History of Exact Sciences* 11, 134-266.

Shapiro, A. E. (1980). The Evolving Structure of Newton's Theory of White Light and Color. *Isis* 71, 211-235.

Shapiro, A. E. (1984). Introduction. Newton I., In A. E. Shapiro (Ed.), *The Optical Papers of Isaac Newton*. Vol. I: The Optical Lectures 1670 – 1672. Cambridge: Cambridge University Press, 1-25.

Shapiro, A. E. (1984). Experiment and Mathematics in Newton's Theory of Color. *Physics Today* 37(9), 34-42.

Shapiro, A. E. (1989). Newton's 'Opticks' and Huygens' 'Traité de la Lumière': Persuing and Eschewing Hypotheses. *Notes and Records of the Royal Society* 43, 223-256.

Westfall, R. S. (1963). Newton's Reply to Hooke and the Theory of Colors. *Isis 54*, 82-96.

Westfall, R. S. (1980). *Never at Rest*. Cambridge (Mass.): Cambridge University Press.

Whiteside, D. T. (1967). Sources and Strengths of Newton's Early Mathematical Thought. In. R. Palter (Ed.), *The Annus Mirabilis of Sir Isaac Newton. 1666 – 1966*. Cambridge(Mass.)/London: M. I. T. Press, 69-85.

MIRCEA RADU

DID HERMANN AND ROBERT GRAßMANN CONTRIBUTE TO THE EMERGENCE OF FORMAL AXIOMATICS?

Abstract. This paper provides a critical assessment of the most influential ideas advanced by contemporary historians of mathematics in connection with the contribution of Hermann Graßmann to the of emergence of a new type of axiomatic approach during the 19th century. My analysis uncovers an wide variety of conflicting statements concerning the objectives and relevance of Graßmann's contributions. Based on this analysis, I argue that a renewed more careful examination of this contribution is needed.

Key words: axiomatics, foundations of arithmetic, foundations of linear algebra, history of mathematics.

In a relatively recent study dedicated to the impact of late 18th and 19th century mathematical developments on philosophy and on the foundations of mathematics, Donald Gillies writes:

> Since German philosophy of mathematics was driven by problems which arose outside philosophy in mathematics, we can classify the main developments by the mathematical results which gave rise to them. Following this principle, we can distinguish three main philosophical views (...). (...) I will argue that the discovery of non-Euclidean geometry gave rise to an empiricist philosophy of mathematics which was applied to geometry, even if not to arithmetic. (...) I will trace a path which led from the arithmetization of analysis by Cantor and Dedekind (...) to logicism in the philosophy of mathematics. Finally (...), I will argue that the development of a plurality of systems of geometry in the period following the discovery of non-Euclidean geometry was the main factor in the rise of formalism in the philosophy of mathematics (Gillies 1999, 173).

The emergence of the abstract concept of *n*-dimensional vector space in H. Graßmann's work is not included in Gillies' list. Hermann and Robert Graßmann's 1861 revolutionary treatment of arithmetic, or the foundational works published by R. Graßmann in 1872 and 1891 are also missing. In a footnote to the previously quoted passage, Gillies mentions that Volker Peckhaus and Georg Henrik von Wright did point this omission out to him (Gillies 1999, 173). This, however, did not have any impact on Gillies paper since, in the end, these contributions were left out of the classification scheme presented.

This seemingly minor detail indicates that Gillies' position was not really determined by the mathematical achievements that shaped 19th century foundational as well as philosophical thinking, but rather by those problems that contributed to the emergence of *intuitionism*, *logicism*, and *formalism* as basic directions of foundational research. Because the Graßmann brothers did not directly contribute to

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 263 – 274.

any of these directions, their ideas remain necessarily outside the horizon of Gillies' reflection.

Gillies' paper can certainly be read independently of the contribution of the Graßmann brothers and it does provide an insightful discussion of the topics considered. At the same time, however, Gillies' omission of the contributions of the Graßmanns is symptomatic for a discourse which despite isolated challenges (Webb 1980, Otte 1989) has established itself in present day research on the history of the foundations of mathematics. While acknowledging (as a rule in a few sentences) that the Graßmanns contributed to some extent to the development of the foundations of mathematics, this discourse attributes the emergence of the new axiomatic trend in mathematics during the 19th century to debates which grew around the discovery of the non-Euclidean geometries, and of the arithmetization of analysis or more generally of mathematics as a whole.[1]

The objective of my paper is to briefly review the diversity of positions forming this discourse. In my view, such an analysis represents a necessary step on the way towards a more appropriate appreciation of the contributions of the Graßmann brothers to the foundations of mathematics. At the same time, this sort of discussion may be helpful for facilitating a critical self-reflection of those working on the history of the foundations of mathematics, concerning the more general issue of understanding the variety of contributions and positions in respect to axiomatics before Pasch and Hilbert.

In his *The Axiomatization of Arithmetic*, Hao Wang writes: "The application of the axiomatic method in the development of numbers is not natural. Its rather late appearance is evidence" (Wang 1957, 146) and adds:

> In 1861, Hermann Graßmann published his *Lehrbuch der Arithmetik*. This was probably the first serious and rather successful attempt to put numbers on a more or less axiomatic basis. Instead of just the positive integers, Graßmann dealt with the totality of all integers, positive, negative and 0. Much of his method can be used to handle the smaller totality of all positive integers, too. He was probably the first to introduce recursive definitions for addition and multiplication, and prove on such a basis ordinary laws of arithmetic by mathematical induction. (Wang 1957, 147)

Wang continues by noting that "Graßmann did not present his development in an axiomatic form, although such a recasting is not difficult" (Wang 1957, 147).[2]

Even though, as it will be seen, H. Graßmann's arithmetic textbook is widely recognized as marking a decisive step towards the axiomatized arithmetic, at the same time there is considerable disagreement as to whether H. Graßmann's treatment as such should be seen as an axiomatic theory, and whether Hermann and Robert Graßmann themselves intended their treatment of arithmetic to be an axiomatization.[3] Among other things, it must be noted that H. Graßmann's textbook does not contain any terms that might be interpreted as synonymous with "axiom" or "postulate" (such as, for instance, "Grundsatz"). Instead, Hermann Graßmann speaks of "definitions [Erklärungen]"[4] only.

Despite these important remarks on the *Lehrbuch*, Wang's paper concentrates on Dedekind's contribution to the axiomatization of arithmetic as pursued in the 1888 *Was sind und was sollen die Zahlen,* and on the ideas contained in Dedekind's letter to H. Keferstein dated February 27th, 1890 (Wang 1957, 149 ff.). In this way

Dedekind rather than Graßmann is judged to have been the true promoter of the axiomatic treatment of arithmetic.

In his *Mathematics – the Music of Reason*, Dieudonné emphasizes what he sees as the discrepancy between Euclid's treatment of arithmetic and his axiomatic approach to geometry. He takes note of the gap between Euclid's treatment of arithmetic and the 19th century efforts, which eventually led to the axiomatization of arithmetic (Dieudonné 1992, 216).[5] Unfortunately, Dieudonné does not consider the problem in detail.

The *Abrégé d'Histoire des Mathématiques* – the most significant historical work edited by Dieudonné – contains a large section (written by M. Guillaume) examining the evolution of axiomatics. This chapter concentrates on the relationship between the evolution of axiomatics, geometry, and logic (Dieudonné 1978, Vol. 2, 315-418). Axiomatics is only briefly mentioned in the section *L'Axiomatisation durant les dix dernières années du dix-neuvième siècle.* Guillaume explains that Peano defined the notion of *real vector space* in the abstract axiomatic manner that is still in use today, and that he subsequently extended this approach from geometry to arithmetic (Guillaume in Dieudonné 1978, Vol. 2, 331). He then adds that: (a) Peano's approach to linear algebra was influenced by Hermann Graßmann's 1844 *Ausdehnungslehre*, and (b), that Peano's approach to arithmetic was influenced by Hermann Graßmann's 1861 *Lehrbuch.* Guillaume, however, choses to place Graßmann's treatment of arithmetic in the tradition that led to Gödel's incompleteness theorem rather than to the axiomatic tradition (ibid., 331 f.).

Even though Hermann Graßmann's work is described as the source of inspiration for Peano's axiomatization of linear algebra and arithmetic, neither Dieudonné nor Guillaume undertake a closer investigation of Hermann Graßmann's contributions to the foundations of mathematics or of his influence on Peano. In particular, the precise nature of Peano's debt to H. Graßmann remains unclear.

A similar situation can be found in the writings signed "Nicolas Bourbaki". Bourbaki does not attribute the axiomatization of arithmetic to H. Graßmann but to the already mentioned 1888 work of Dedekind. The 19th century shift towards the axiomatization of arithmetic is brought in connection with an increased tendency of looking for models for the various mathematical theories inside arithmetic rather than geometry, a trend which we are told began around 1880 (Bourbaki 1974 <1969>, 36 ff.). Bourbaki adds that prior to the 19th century hardly anybody seriously questioned the intuitive foundations of arithmetic already present in Euclid's *Elements.* He notes that even Weierstrass who is famous for his pursuit of rigor in analysis, did not see the need for a revision of the foundations of arithmetic itself. Graßmann is said to have been the first to pursue a "logical clarification" of arithmetic, but no details are given as to what this is supposed to mean (ibid., 38).

In his 1980 book *Mechanism, Mentalism, and Metamathematics*, Judson Webb also takes on the issue of the axiomatization of arithmetic:

> Returning to elementary arithmetic, we recall that it had always been traditionally conceived in terms of algorithms and calculation. Even Gauss' *Disquisitione Arithmeticae* of 1801, which shifted the focus of arithmetic to proof of theorems in his congruence theory, had as an essential goal the explanation of why certain algorithms work as well as proofs that they will work. In the middle ages, 'algorithm' was defined

> to be the arithmetic which calculates with Indo-Arabic numerals. Conversely the 1771 edition of the *Encyclopedia Britannica* defined 'Arithmetic' in terms of the notions of number and algorithm. Number was defined, following Euclid, as 'either a unit, or a multitudes of units', and once a person has 'the idea of number in his mind', he is ready for the 'science of arithmetic' (...). (Webb 1980, 43 f.)

Arithmetic is described as having been originally conceived as a collection of algorithms (Berkeley's work is an illustration), no attempts having been made to organize these algorithms into a coherent deductive system. As Webb points out, this position can be still encountered as late as 1801 in Gauß' work.

Webb distinguishes between two fundamentally different approaches to number: the *algorithmic* and the *deductive*. The *algorithmic* approach "regards the basic operations of arithmetic as algorithms rather than as functions in the modern sense, i. e., as rules rather than sets" (Webb 1980, 44). The deductive approach is more theoretical; it

> reduces the notion of number to concepts of pure logic and then concentrates on the proofs of arithmetical propositions. (...) The algorithmic conception tends to stress formalisms and concrete symbols while the deductive conception stresses concepts and abstract objects. (Webb 1980, 44)

Webb writes that Dedekind and Frege worked to reduce number to pure logic, and on this basis, he classifies them among those pursuing a *deductive* treatment of arithmetic. As for H. Graßmann's treatment of arithmetic, he writes:

> An early proponent of logicism was H. Graßmann, who held that while geometry depended to some extent on spatial intuition, number depended wholly on the law of thought. He was the first mathematician both to approach arithmetic axiomatically and to employ recursive definitions for the basic arithmetical operations. The two are related, for if one decides to use, say, the recursive 'definition' $a + (b + 1) = (a + b) + 1$ for addition, then the fact that it does not in general enable one to eliminate the plus sign, together with the difficulty of replacing it with anything more basic or obvious, obliges one to take it as an 'axiom'. (...) Recursive definitions for the basic arithmetic operations began to appear frequently in the literature after Graßmann and were seen as raising two kinds of problems, which we can classify roughly as logical and mathematical. (Webb 1980, 44)

Webb endorses the idea that the axiomatization of arithmetic was an achievement of 19^{th} century mathematics. Moreover, unlike Wang, Webb does not hesitate to call Graßmann's 1861 approach to arithmetic an axiomatization. He also provides a justification for assigning Graßmann's *Lehrbuch* an axiomatic nature. The introduction of formally expressed, recurrent definitions in arithmetic such as $a + (b + 1) = (a + b) + 1$ leads to the fact (already noticed and criticized by Frege, and emphasized by Wang) that the sign "+" (of *addition*) which is to be defined appears both in the expression of the *definiens* as well as in the expression of the *definiendum*. Therefore, formally expressed recurrent definitions are circular. The only way to avoid regarding this circularity as detrimental to the entire construction is to take them as *implicit definitions* of the signs involved, in Hilbert's fashion. For Webb this means that these definitions must therefore be interpreted as axioms.

According to Webb, Graßmann's axiomatization of arithmetic is a direct consequence of the shift towards explicit, formally stated, recurrent definitions of the basic operations of arithmetic. H. Graßmann's treatment of arithmetic is described as a combination of the algorithmic and the formal. No justification for the rather doubtful claim that H. Graßmann was a forerunner of logicism is given. This statement, however, needs further clarification, for neither Hermann nor Robert Graßmann ever maintained that logic should be taken as a foundation for arithmetic or for any other mathematical discipline, but on the contrary, R. Graßmann developed logic as a branch of mathematics. Moreover, logicism and axiomatics are not necessarily compatible (Compare Russell 1919; Otte 2002).

In 1982 D. Gillies published a book called *Frege, Dedekind, and Peano on the Foundations of Arithmetic*. In the introduction he explains that the broad reason for writing the book was to find an answer to the questions:

> Why did these authors get interested in the subject? Why did they feel that it would be desirable to provide a firm foundation for arithmetic? After all, the arithmetic of the natural numbers {0, 1, 2, ..., *n*, ...} had been widely employed by mathematicians in Western Europe since 1500. Why then was it only in the last quarter of the 19th century that serious attempts were made to examine the foundations of the theory of numbers? (Gillies 1982, 1)

On the whole, the thesis defended by Gillies is that the axiomatic turn in arithmetic (at least as pursued by the authors mentioned in the title of his book) was a late outcome of the arithmetization of analysis (Gillies 1982, 1 ff.). Hermann Graßmann's name is mentioned only once in a section of the book in which Gillies undertakes a comparison between the account of the foundations of arithmetic given by Dedekind, Frege, Peano, and Hilbert. Gillies quotes Wang's statement according to which "Historically, Peano borrowed his axioms from Dedekind" (Wang in Gillies 1982, 66) and adds:

> I think it is correct to speak of Peano's axioms rather than Dedekind's axioms; for Dedekind was not trying to axiomatize arithmetic, but rather to define arithmetical notions in terms of logical ones. Another way of putting it is to say that Peano is not a logicist, but a forerunner of Hilbert's later formalism. (...) Another difference between Dedekind and Peano is that Dedekind was a logicist, but did not use formal logic; while Peano was not a logicist, but did use formal logic. (Gillies 1982, 66)

Gottfried Martin attempted a detailed historical analysis of the development of axiomatics since Kant (Martin 1972). The starting point of Martin's work is the observation that Kant has had a number of close collaborators who wrote mathematical textbooks in which they attempted to give mathematical substance to Kant's epistemological ideas. The presentation of mathematics given in these books differs from that provided in the other books of the time through the fact that they all adopt an axiomatic path (Martin 1972, 20 f.). Martin speaks of the *axiomatic foundation of mathematics*, but his work focuses almost exclusively on the axiomatization of arithmetic and on combinatorics. Apart from Kant's writings, Martin's analysis focuses on the philosophical works of Johann Schultz and Jakob F. Fries (many other authors such as F. Murhard and M. Ohm are also considered).

Martin's work involves two related claims: a *weak* claim, and a *strong* one. According to the *weak claim*, Kant's treatment of the axiomatic nature of mathematics (in the *Critique* as well as in the Kant – Schultz correspondence) influenced Schultz, Fries etc. and it pushed them in the direction of providing an *explicit* axiomatic treatment of arithmetic. According to the *strong claim*, Kant's work provided not only the philosophical foundations and impetus for an axiomatization of arithmetic, but stronger: the axioms published in Schultz' first edition (1789) of his *Prüfung der Kantischen Critik der reinen Vernunft* (these are the *commutative* and *associative* laws for the addition of natural numbers), which, according to Martin, were taken over by Fries in his 1822 *Die mathematische Naturphilosophie*, belong to Kant. Martin claims that Kant was the first to recognize "the essentially axiomatic nature of arithmetic" and the first to have stated "two axioms of addition, namely the associative and the commutative law" (Martin 1972, 66).

Martin, however, was unable to find any conclusive historical evidence capable of sustaining his *strong claim*. As Friedman (Friedman 1992, 105 ff.) pointed out, Kant's own words go against Martin's *strong claim*. As far as the weak claim is concerned, Martin's case looks better. He, however, does not venture into any detailed investigation of this last thesis. Instead, he focuses unilaterally on establishing Kant's paternity of the axiomatization of arithmetic. In respect to Graßmann, Martin simply suggests a Kant-Schultz-Fries-Ohm-Hermann Graßmann link in the treatment of arithmetic, and in this way tries to present Graßmann's 1861 revolutionary approach to arithmetic as a late technical improvement of the axiomatic ideal put forward by Kant (Martin 1972, 50). This limits the scope of his analysis and narrows the interpretation of the various passages of the writings discussed by him.

Hans Wussing is best known for his important contributions to the history of group theory, and more generally, for his work on the evolution of the concept of algebraic structure. In his *Die Genesis des abstrakten Gruppenbegriffs* of 1969, as well as in other works, Wussing provides a detailed historical reconstruction of the various lines of thought involved in the emergence of the abstract group concept. Wussing writes that the decisive element in the development of the *abstract* group concept consisted in the shift from a bottom-up approach (in which the concept of group is only implicitly treated as embedded in the study of some particular mathematical topic – Gauss' introduction of the *congruence modulo a given natural number* in number theory, or Cauchy's study of the theory of permutations are examples of this), to a top-down approach in which a group appears as "System definierter Relationen zwischen abstrakten Elementen" (Wussing 1969, 171). The group structure emerged as an autonomous object of mathematical study as a result of such a shift. The decisive step towards the passage from the bottom-up to the top-down approach concerning the group concept is assigned to W. van Dyck's work (Wussing 1969, 182). This shift is important because as Wussing points out van Dyck was strongly influenced by Hermann Graßmann, Hankel, and Schröder (Wussing 1969, 180).

Even though Wussing mentions Graßmann several times in his book, he nowhere undertakes a detailed examination of the structural-algebraic ideas contained in Graßmann's work. This is perhaps due to the fact that Wussing places Graßmann not in the group-theoretic tradition, but rather in the tradition that led to the development of the concept of vector space which is not dealt with in his book. In any case, a particularly important factor in the development of van Dyck's abstract group concept was his discussion of the possibility of framing various "operations of multiplication which satisfy the associative but not the commutative law" (van Dyck in Wussing 1969, 180). As it is well known, Graßmann was one of the first to realize (as early as 1832) the possibility of introducing a non-commutative product concept in geometry, and this had an enormous impact on the development of his ideas. The idea of a possible H. Graßmann-van Dyck link on this issue is not considered by Wussing.

Wussing's work on the history of the abstract group concept, a concept whose emergence is strongly linked to the resurrection of axiomatics in 19th century mathematics, suggests that Hermann Graßmann's A_1 may have had a major influence. Yet the precise nature of this influence, the structural-algebraic ideas, and the place of axiomatics in Graßmann's work as such are not considered.[6]

Hermann Graßmann's 1844 *Ausdehnungslehre* contains a short section dedicated to the presentation of the *Allgemeine Formenlehre* (General Theory of Forms). Graßmann describes it as a "new mathematical discipline" (Graßmann 1894 – 1911, Bd. I.1, 33). To the modern reader this theory is bound to look like an axiomatic introduction of the basic structures of abstract algebra (semigroup, group, and field). In terms of the already mentioned distinction between the bottom-up and the top-down approach to algebra, H. Graßmann's approach was clearly and deliberately intended to be top-down. Moreover, Graßmann claims that the development of this theory must precede the treatment of all the other "particular [speciellen]" mathematical disciplines:

> Antecedent to the division of the theory of forms[7] into four branches is a more general subject that we may call the general theory of forms. In it are presented the general conjunctive law that apply to all branches alike.
> This preliminary subject is not intended simply to save repeating the same material in all four branches and thus to condense the treatment of the different parts, but also permits what naturally belongs together to appear together, and to act as the foundation of the whole. (H. Graßmann 1995 <1844>, 28; H. Graßmann 1894 – 1911, Bd. I.1., 28)

Commenting on this passage, Lewis writes:

> This description might imply to a modern reader that Graßmann is establishing a system of axioms. But what is presented is not a set of unproven statements from which succeeding statements are deduced; rather, principles of connection, expressed by means of the general concepts equality and difference, and connection and separation, are symbolized. To call the 'basis of the whole' may mean nothing more than that it precedes and is used in the succeeding presentation. (Lewis 1977, 140)

Even though Lewis recognized the possibility of interpreting the previous Graßmann quotes as an indication of an axiomatic intention, he believes that GTF was not so intended, and stronger, that it cannot be seen as an axiomatization of mathematics. Lewis' claims, however, are not easy to interpret. Ultimately, they seems to be based on the distinction between "unproved statements" and "deduction" on the one hand, and a "principle of connection" and establishing a "symbolic calculus" on the other. The term "axiomatics" is the one reserved for the former and denied of the latter. Unfortunately, Lewis does not explain his distinction any further so that its meaning remains somewhat unclear. In another short passage, Lewis attacks the same issue once again from another angle:

> There is another depiction of mathematical progress which has had a dominant role in historiography of mathematical progress and which emphasizes the rise of the axiomatic method as the fundamental characteristic of modern mathematics. It is possible to view Graßmann's general theory of forms also as a part of this line of progress, but only a misinterpretation could allow the A_1 as a whole to be viewed in this way.
> Such a misconception could come about from the association of the A_1 with Peano and Whitehead, who in turn are associated with the axiomatic development. But there is no evidence that the A_1 influenced these two in this way, and neither Peano nor Whitehead discussed the larger import of the A_1 in the development of mathematics. (Lewis 1977, 129)

Here again, Lewis seems to be explicitly acknowledging the possibility of consistently interpreting the GTF as providing a formal axiomatic foundation of the individual mathematical disciplines. In the end, however, this option is dismissed indirectly by claiming that Graßmann's A_1 did not participate to the axiomatic turn in algebra which is assigned to Peano and Whitehead.[8]

Lewis' paper contains a detailed discussion of many of the fundamental ideas of Hermann Graßmann's *Ausdehnungslehre*. Yet, Lewis is interested primarily in clarifying the relationship between Schleiermacher's philosophy and Hermann Graßmann's conception as outlined in the 1844 *Ausdehnungslehre*. Foundational issues are discussed only briefly. On the whole, Lewis concludes that Graßmann's GTF cannot be seen as an axiomatic foundation of mathematics and that it was not so intended.

Despite his renewed, perceptive discussion of the role of the *General Theory of Forms* in the A_1, Jean-Luc Dorier endorses Lewis' conclusion, and in addition to that, he claims that Graßmann's approach to the second edition of the *Ausdehnungslehre* which was published in 1862 was not axiomatic but merely "formalistic:"

> he bare formalism of the A_2 is the artificial result of Graßmann's attempt to satisfy the criticisms received after the edition of 1844; this resulted in the inaccessibility of most of the intuitive discussion, which is to be found in the A_1. (Dorier in Schubring 1996, 181)

The same position was defended in an earlier paper as well (Cf. Dorier 1995).

In his study *The Axiomatization of Linear Algebra: 1875 – 1940*, G. H. Moore attempts to settle a question raised by MacLane concerning the origin of "the definition of a vector space as a set of elements subject to suitably axiomatized operations of addition and multiplication by scalars, and not just as a set of n-tuples of scalars closed under these two operations?" (MacLane in Moore 1995, 263). Moore's answer is: "The relevant period for axiomatization and acceptance extends from 1875 to about 1940" (Moore 1995, 263). Both Dorier as well as Moore mention some of Hermann Graßmann's contributions, but they do not undertake a detailed examination of his ideas and they both refuse to credit Graßmann's contributions to have played a fundamental part in the development of axiomatics in general or in the axiomatization of the vector space structure.

The contributions of the Graßmann brothers to the development of algebra and arithmetic were also discussed by Volker Peckhaus (Peckhaus 1997, 243-250). Peckhaus' objective is to examine the interconnection between the development of logic and that of algebra since Leibniz. The presentation of the ideas of Hermann and Robert Graßmann does not touch upon the issue of the contribution of the Graßmanns to the emergence of axiomatics either in respect to abstract algebra or in respect to arithmetic.

Peckhaus' book contains just two more or less explicit references to the contributions of the Graßmann brothers to axiomatics. Firstly, we are told that the significance of H. Graßmann's *Allgemeine Formenlehre* for the development of "formalism [Formalismus]" was discussed by Jean Cavaillès in (Peckhaus 1997, 245; Cavaillès 1938). Secondly we are reminded that Peano's *Calcolo Geometrico* provides an axiomatization of vector algebra which builds on H. Graßmann's contributions (Peckhaus 1997, 246). Both claims are certainly correct. Both, however, are hardly helpful in establishing the merits of H. Graßmann's work. Even though axiomatics is indeed the central theme of Cavaillès remarkable book mentioned by Peckhaus, at the same time, the contribution of the Graßmann brothers is only briefly mentioned in it. Cavaillès does not provide a detailed examination either of the complex foundational ideas that can be found in the work of the Graßmanns nor of any other aspect of these contributions (Cavaillès 1981 <1938>, 48 ff.). Moreover, Cavaillès writes that Graßmann's theory was developed based on extensive use of intuitive evidence taken from the peculiar use made of pictorial representations, and that the formalisms used were just an attempt to dissimulate this intuitive input (Cavaillès 1981 <1938>).

Peckhaus does not undertake a detailed examination of the contributions of the Graßmann brothers. He is only interested in the contributions of the Graßmanns in as far as they prepared the ground for Schröder's work on algebra. Schröder's contribution is assigned a greater historical significance. This option is reflected in Peckhaus' comparison between the Graßmann brothers and Schröder concerning the use of recurrent definitions. Speaking about recurrent definitions such as $a + (b + e) = (a + b) + e$, Peckhaus writes:

> This type of definition which plays only a marginal role in [Hermann – M.R.] Graßmann's work was taken over by Schröder who used it systematically in his approach to the introduction of the arithmetical operations. (Translation M. R.)
> Diese bei [Hermann – M. R.] Graßmann nur beiläufig verwendeten Definitionsarten werden von Schröder übernommen und zu einem durchgängigen Gestaltungsprinzip für die Einführung der arithmetischen Rechnungsarten ausgebaut (Peckhaus 1997, 247).

Peckhaus' analysis certainly has the merit of uncovering various important aspects of Schröder's seminal contributions. At the same time, it must be pointed out that recurrent definitions were by no means used "beiläufig" by the Graßmann brothers but were central both to the 1861 treatment of arithmetic as well as to the 1872 treatment of the General Theory of Forms.

As far as the contribution of the Graßmann brothers to the emergence of formal-axiomatics in general, and to the axiomatization of arithmetic and algebra in particular is concerned, the discussion outlined above leaves us with an unsatisfactory situation. While we are told that this contribution represented a turning point in the treatment of the foundations of arithmetic, and that it had a significant impact on the development of the abstract concepts of group and vector space, we are not given any significant information about the mathematical and epistemological origin of this shift, of the nature of the shift, or of its relevance for the evolution of axiomatics in general.

Many important questions remain open and the answers provided controversial. Some of them are: (1) when did the shift from the pre-axiomatic to the axiomatic stage in the treatment of arithmetic really occur (in H. Graßmann's work, or earlier in the writings of Kant and Schultz, or rather subsequently in the work of Peirce, Dedekind, etc.)? (2) Although it is true that events such as the discovery of the non-Euclidean geometries, the arithmetization of analysis and of geometry, and the paradoxes of set theory have played an important role in the resurrection of axiomatics during the 19^{th} and the first half of the 20^{th} century, these topics did not play any role whatsoever in the works of the Graßmann brothers. If this is true, then what were the mathematical and epistemological sources of Hermann and Robert Graßmann's axiomatic turn? (3) Indeed, why did an outstanding mathematician such as H. Graßmann bother taking up the foundations of arithmetic? (4) Did Hermann and/or Robert Graßmann regard their treatment of arithmetic as an axiomatization? (5) Does the work of the Graßmann brothers contain other ideas that are relevant for a history of axiomatics? (6) Should the foundational ideas of the Graßmann brothers be seen as a source of logicism or rather of formalism, or perhaps as belonging to some other tradition? (7) Over and above all that, the previous discussion also hinges upon another delicate general distinction between an axiomatic mathematical theory and a deductive but non-axiomatic mathematical theory. As far as I can see, there is no consensus on this issue between the various historians dealing with the foundations of mathematics. As long as one is interested only in the technical outcomes, this may not have a great significance. As soon as we turn to historical and to the philosophical matters related to them the situation changes and a much deeper investigation of these matters becomes necessary. It seems to me that the time has come to follow Judson Webb's and Michael Otte's example, abandoning the rather prudent position in respect to the contributions of Hermann Graßmann to the development of the foundations of mathematics, and replace old prejudice by new substantial research.

Institut für Didaktik der Mathematik, Universität Bielefeld

NOTES

[1] This attitude is expressed in even stronger terms by Herbert Stachowiak: "As far as mathematics is concerned, the only relevant contributions to axiomatics made between Euclid and Hilbert come down basically to the debates over Euclid's postulate." (Stachowiak 1971, 311).

[2] Wang does not explain why he thinks that an axiomatic treatment of arithmetic is "unnatural." He also does not explain why he thinks H. Graßmann's treatment of arithmetic is not cast in axiomatic form.

[3] Compare Radu 2000 and Radu 2003.

[4] The German word "Erklärung" means explanation, but Hermann Graßmann as well as his brother use this term in the sense of "definition."

[5] A similar view is advocated by Medvedev (Medvedev 1981, 223).

[6] Demidov also considers the relation between group theory and the axiomatic method. Among the mathematicians which contributed to the axiomatization of the "new algebra" during the 19th century he

mentions Hankel, Weber, and Dedekind. Although H. Graßmann's work is also mentioned, no connection is made between Graßmann and the axiomatization of algebra (Demidov 1970).

[7] Here the expression "theory of forms" is used as another name for "mathematics".

[8] This rejection of placing the *Allgemeine Formenlehre* in the axiomatic tradition is not based on Lewis analysis of its content, but rather on the historical point according to which there is no clear evidence of the fact that (as an axiomatic theory) the GTF has had an impact on Peano or Whitehead. Lewis' historical argument is rather weak. Indeed, Graßmann's treatment of the *Calculus of Extension* and the account of arithmetic developed in the *Lehrbuch* are closely related to the axiomatic spirit which permeates the A_1 at several levels, and this work was explicitly connected to axiomatics. Lewis' thesis is called into question by Wussing's discussion of van Dyck's development of the abstract group concept.

REFERENCES

Bourbaki, N. (1974 <1969>). *Éléments d'histoire des mathématique.* Paris: Hermann.

Cavaillès, J. (1981 <1938>). *Méthode axiomatique et Formalisme.* Paris: Hermann.

Demidov, S. (1970). Évolution, extension et limites de la méthode axiomatique dans les sciences modernes sur l'example de la géométrie. *Archives Internationales d'Histoire des Science,* 3-30. Paris.

Dieudonné, J. (Ed.) (1978). *Abrégé d'histoire des mathématiques 1700 – 1900.* 2 Vols. Braunschweig: Vieweg.

Dieudonné, J. (1992). *Mathematics – the Music of Reason.* Berlin: Springer.

Dorier, J.-L. (1995). A General Outline of the Genesis of Vector Space Theory. In *Historia Mathematica* 22, 227-261.

Dorier, J.-L. (1996). Basis and Dimension – from Graßmann to van der Waerden. In G. Schubring (Ed.), *Hermann Günther Graßmann (1809 – 1877) Visionary Mathematician, Scientist and Neohumanist Scholar* (Boston Studies in the Philosophy of Science). Dordrecht: Kluwer, 175-196.

Friedman, M. (1992). *Kant and the Exact Sciences.* Cambridge, Mass.: Harvard Univ. Press.

Gillies, D. A. (1982). *Frege, Dedekind, and Peano on the Foundations of Arithmetic.* Assen: van Gorcum.

Gillies, D. A. (1999). German philosophy of mathematics from Gauss to Hilbert, In A. O'hear (Ed.), German Philosophy Since Kant. Cambridge Mass.: Cambridge Univ. , 167-192.

Graßmann, H. (1894 – 1911). *Gesammelte mathematische und physikalische Werke.* Friedrich Engel (Ed.) Vols. I.1 to III.2. Leipzig: Teubner (Quoted in respect to the volume).

Graßmann, H. (1861). *Lehrbuch der Arithmetik.* Berlin: R. Graßmann.

Graßmann, H. (1995 <1844>). A new branch of mathematics: The Ausdehnungslehre of 1844 and other works. Translated by L. C. Kannenberg. Chicago a. o.: Open Court.

Graßmann, R. (1966 <1872>). *Die Formenlehre oder Mathematik.* Hildesheim: Georg Olms Verlagsbuchhandlung.

Graßmann, R. (1891). *Die Zahlenlehre oder Arithmetik – streng wissenschaftlich in strenger Formelentwicklung.* Stettin: Verlag von R. Graßmann.

Grattan-Guinness, I. (1996). Where does Graßmann fit in the History of Logic? In G. Schubring (Ed.). *Hermann Günther Graßmann (1809 – 1877) Visionary Mathematician, Scientist and Neohumanist Scholar* (Boston Studies in the Philosophy of Science). Dordrecht: Kluwer, 211-216.

Lewis, A. C. (1977). H. Graßmann's 1844 Ausdehnungslehre and Schleiermacher's Dialektik. *Annals of Science* 34, 103-162.

Martin, G. (1972). *Arithmetik und Kombinatorik bei Kant.* Berlin: Walter de Gruyter.

Medvedev, F. A. (1981). On the Role of the Axiomatic Method in the Development of Ancient Mathematics. In J. Hintikka, D. Gruender, E. Agazzi (Eds.), *Theory Change, Ancient Axiomatics, and Galileo's Methodology.* Proceedings of the 1978 Pisa Conference on the History and Philosophy of Science. Vol. 1 (Synthese Library 145). Dordrecht: Reidel.

Moore, G. H. (1995). The Axiomatization of Linear Algebra: 1875 – 1940. *Historia Mathematica* 22, 262-303.

Otte, M. (1989). The Ideas of Hermann Graßmann in the Context of the Mathematical and Philosophical Tradition since Leibniz. *Historia Mathematica* 16, 1-35.

Otte, M. (2002). Einleitung. In B. Russell, *Einführung in die mathematische Philosophie.* Hamburg: Meiner.

Peckhaus, V. (1997). *Logik, Mathesis universalis und allgemeine Wissenschaft. Leibniz und die Wiederentdeckung der formalen Logik im 19. Jahrhundert*. Berlin: Akademie Verlag.

Radu, M. (2000). Justus Grassmann's Contributions to the Foundations of Mathematics: Mathematical and Philosophical Aspects. In: *Historia Mathematica* 27.1: 4-35.

Radu, M. (2003). A debate over the axiomization of arithmetic: Otto Holder against Robert Grassman, in *Historia Mathematica*,30 (2003):341-377.

Russell, B. (2002 <1919>). *Einführung in die mathematische Philosophie*. Hamburg: Meiner.

Stachowiak, H. (1971). *Rationalismus im Ursprung – Die Genesis des axiomatischen Denkens*. Wien: Springer Verlag.

Wang, H. (1957). The Axiomatisation of Arithmetic. In: *The Journal of Symbolic Logic* 22.2: 145-158.

Webb, J. C. (1980). *Mechanism, Mentalism, and Metamathematics*. Dordrecht: D. Reidel Pub. Co.

Wussing, H. (1969). *Die Genesis des abstrakten Gruppenbegriffes*. Berlin: DVW.

GERT SCHUBRING

A CASE STUDY IN GENERALISATION

The Notion of Multiplication

Abstract. While the operation of multiplication is presented by historiography as an unproblematic notion which did not see substantial evolution, a polemic by Ampère against Bézout's arithmetic textbook, highly popular in France over many decades, is used here to unravel decisive restrictions imposed on multiplication in various mathematical cultures since Old Babylonian times. The analysis illustrates not only the controversy about the existence of a Greek "geometric algebra," but makes accessible also reflections on non-commutativity of operations in the 18th century already. The concept of multiplication thus underwent characteristic changes in the process of generalising mathematics.

Key words: arithmetic operations, de-contextualization, geometric algebra, non-commutativity, qunatities vs. numbers, restrictedness of multiplication.

Multiplication seems to be a simple and conceptually and epistemologically unproblematic, innocent notion. This widespread assumption is reinforced when one consults the chapter on multiplication in the famous *Tropfke* for algebra, i. e. the major classical historical source for the development of concepts and notions in elementary mathematics. Of its twenty-five pages, the chapter devotes barely more than half a page to the conceptual history of multiplication *stricto sensu* – the rest discussing various uses of signs for the operation of multiplication, the terms used, and representations in different cultures. The main content of the short conceptual paragraph is given by a reference to Euclid's definition of multiplication as repeated addition, in Book VII, definition 15:

> A number is said to multiply a number when that which is multiplied is added to itself as many times as there are units in the other, and thus some number is produced (Heath 1956 II, 278; Tropfke 1980, 207-208).

Multiplication thus seems to present the case of a stable notion with a meaning remaining identical over millennia.

I began to doubt this harmonic view of an unchangeable identity when I stumbled on a harsh critique, by Ampère, of how Bézout presented multiplication in one of his textbooks. As Etienne Bézout (1730 – 1783) was one of the most influential French textbook writers – on an international level, too –, meanings of multiplication not conforming to the traditional view would imply that such "non-conformist" views were rather broadly disseminated, and even widely accepted. Moreover, Ampère proved to be an analyst highly sensitive to conceptual problems in elementary mathematics. This may sound surprising, since André-Marie Ampère (1775 – 1836) is remembered as an outstanding researcher in physics and chemistry, and also in

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 275 – 285.

philosophy; his profound research into the foundations of mathematics are not well known, however. Even in a recent authoritative biography of Ampère, it is claimed that he would have had the ability to contribute to rigorous foundations of calculus, but it is also claimed that his interests never focused on this topic (Hofmann 1996, 59)

As his *Nachlaß* reveals, however, Ampère strove for clarity and rigor in the foundations of mathematics – while teaching calculus and mechanics at the *École Polytechnique*. Numerous manuscripts in his *Nachlaß* prove that he worked intensively over many years to prepare a comprehensive treatise on pure mathematics, from arithmetic to the calculus. He never succeeded in completing this text: in particular, the repeated attempts to come to terms with the first chapters in arithmetic show on the one hand his effective concern for foundations, but on the other hand that there were still unsolved problems inhering in the number concept.

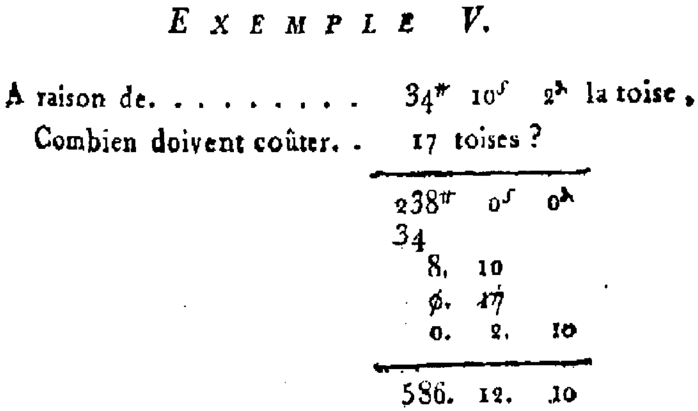
In one of these numerous fragments of the arithmetical part within his projected treatise, Ampère discussed the concept of multiplication; and there he made the following critical remark:

> Avant de finir cette petite digression, je crois devoir dire un mot d'un passage de Bezout sur la nature de l'opération dont je traite, et qui tendrait renverser le premier principe, qui consiste dans l'invariabilité du produit quelque soit celui des deux facteurs qui servent de multiplicateur: On trouve à l'article 117 de l'artillerie deux multiplications de 17 toises par $34\#.10^{s}.2^{d}$. Dans l'une, le produit est exprimé en livres et dans l'autre en toises, et il dit qu'il n'a donné ces deux exemples de multiplication que pour prouver qu'en changeant le multiplicande avec le multiplicateur on peut changer le produit.

As Ampère points out here, Bézout's claim in his series of textbooks for artillery officers, that multiplying two quantities in a different order will result in different products would undermine the basic property of multiplication: that the product is independent of the choice of one of the two factors to serve as multiplicator. In fact, commutativity had always constituted a fundamental element of the notion of multiplication. Euclid had proved it, as proposition 16 of Book VII, and Antoine Arnauld had even postulated the independence of order as an axiom, in his famous innovatory textbook *Nouveaux Élémens de Géométrie* of 1667:

> c'est la mesme chose dans la multiplication de commencer par lequel on veut des deux nombres que l'on multiplie (Arnauld 1667, 2).

Let us look therefore at Bézout's exposition in his arithmetic textbook for the artillery[1] There, Bézout had in fact multiplied the two factors mentioned in a different order. First, the compound factor by 17 *toises* (fathoms):



EXEMPLE V.

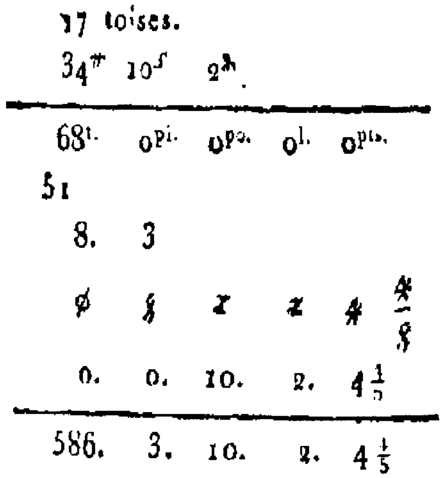
A raison de. . . . . 34# 10s 2d la toise,
Combien doivent coûter. . 17 toises?

238# 0s 0d
34
8, 10
6, 17
0. 2, 10

586. 12. 10

Figure 1. Example from Bézout(1800).

And, thereafter, in inverse order:



17 toises.
34# 10s 2d.

68t. 0pi. 0po. 0l. 0pts.
51
8. 3
0. 0. 10. 2. 4 1/5

586. 3. 10. 2. 4 1/5

Figure 2. From Bézout (1800, 90).

For Bézout, this was not a marginal observation; rather, he emphasized the importance of distinguishing between multiplicator and multiplicand, since – despite the equality of the factors in both cases – "the two products are different" (ibid., 91).

One observes two differences between both products. First, there is the **numerical** difference to which Bézout is apparently alluding. That is due to the complex system of non-metrical sub-units of *toises* (fathoms) and of *livres* (pounds) respectively:

Table 1. List of sub-units used by Bézout

sign	*lengths*	*sign*	*prices*
t	*1 toise = 6 pieds*	*#*	*1 livre = 20 sous*
pi	*1 pied = 12 pouces*	*s*	*1 sous = 12 deniers*
po	*1 pouce = 12 lignes*	*d*	*1 denier*
l	*1 ligne = 12 points (pts)*		

The second difference was a difference in the **dimension** of the products, and this constituted the most problematic feature for Ampère, since we saw that he objected to the one product being expressed in *livres* and the other in *toises*.

These differences reveal the hitherto hidden problem in the history of the notion of multiplication. Bézout's intention had been to show "the importance" of distinguishing between multiplicand and multiplicator when both of them are concrete ("concrets") (Bézout 1800, 91). What was hence at stake was the multiplication not of abstract entities, i. e. numbers, but of "concrete" or "named" quantities, i. e. magnitudes. Euclid's definition quoted above only applied to numbers (ἀριθμός): actually, only to integer numbers; likewise, Arnauld had formulated his axiom of commutativity for numbers, too. Operations on numbers are not necessarily defined for quantities, too.

The historical problem of multiplying general quantities, and its evolution, has not been recognized and studied in traditional historiography. Only relatively recently has research on the nature and meaning of operations in arithmetic and in algebra begun, in particular for some key periods and particular key cultural settings. I will briefly recapitulate here some important results for Babylonian mathematics, Greek mathematics, and early modern European mathematics and show some common patterns which highlight the relation between magnitudes and numbers as a key problem which went beyond the illusion of a stable notion of multiplication which always remained identical. At the same time, it will become evident that the problem was well known and controversially discussed in particular in the pre-modern period.

BABYLONIAN MATHEMATICS

Since the pioneering work of Otto Neugebauer in the 1920s and 1930s, who was the first to seriously study Babylonian cuneiform texts and unravel their mathematical meaning, his interpretation of these texts as constituting a Babylonian **algebra** has been generally shared and even been widely disseminated by B. L. van der Waerden. Recent critical research has reassessed this pioneering work and the underlying assumptions of interpretation. One of the major outcomes of this research consists in showing how deeply contextualised the mathematical notions were, and how far Babylonians really were from dealing with "pure" concepts.

A first perspective for these new approaches was provided by research into the generalisation of the number concept from an utterly complex contextualisation of magnitudes. Due to the strenuous efforts of an international group analysing Sumerian and Babylonian cuneiform texts, which began mainly as book-keeping records for the needs of state administration, it has been possible to reconstruct the process of transformation of signs for measuring quantities into number signs, and the establishment of number systems. The conceptual and technical problem to be solved for this reconstruction was determined by the fact that the first sign systems were all closely tied to the concrete objects they were to measure, so that the same sign may have a different meaning when applied to a different class of objects. In fact, the research group identified about 6,500 different number signs; the researchers were able to trace their transformation into more and more "abstract" and general numbers, loosening their ties to the object classes and quantities they were meant to measure (see Nissen, Damerow, Englund 1993).

And only recently has it become clear that arithmetical operations had to undergo an analogous process of generalisation, of de-contextualisation. Hitherto, the postulated Babylonian algebra had – as Høyrup has put it – "looked astonishingly modern and similar to ours" (Høyrup 2002, 7-8). The research result which entirely overthrew this received algebraic interpretation was to show that clear-cut arithmetical operations did not exist which would be the necessary basis for the supposed algebra. There were no uniform, "numerical" operations; rather, there existed several operations molded according to different – mainly geometrical – contexts, instead of just one of the four classical operations.

In fact, there is not just the one canonical type of addition which is familiar to us; instead there are two additive operations. The first can be understood as "appending" and is non-commutative, since one entity is always "appended to" another so that it becomes absorbed into the other which "conserves its identity while increasing in magnitude". This operation is only used for "concretely meaningful 'additions'" (Høyrup 2002, 19). The other additive type can be understood as "accumulating"; it is commutative. "It adds the measuring numbers of two or more" quantities and hence resembles the later standard notion of addition (ibid.).

As for subtraction, Høyrup identified two distinct forms for this, too: "removal" and "comparison". He sees *removal* as the inverse of *appending*: "it can be used only when the subtrahend is really *part* of the entity from which it is subtracted". The other form, *comparison*, is not a reversal of any kind of addition; it is a concrete operation, and not on numbers, "used to say how much one magnitude *A* exceeds another magnitude *B which it does not contain*" (ibid., 21).

The most complicated case is presented by multiplication, however! Høyrup distinguished four different operations – even with several numbers of terms expressing them – for what had traditionally been understood as multiplication. The first comes near to the repeated addition of numbers: it applies to the multiplication of two numbers *a* and *b* as "*a* steps *b*" (ibid., 22). The next is called "raising" or "lifting" and applies to the determination of a concrete magnitude by means of multiplication; it is used for multiplying magnitudes with scalar factors and for the calculation of volumes from base and height (ibid.). The third type, literally named "the double", applies to repeating concrete magnitudes and is translated by Høyrup as "to repeat" or "to repeat until *n*, *n* between 2 and 9" (ibid., 23).

In fact, this is an exclusively geometrical operation, since it does not represent an operation between numbers and numbers, nor between numbers and quantities. It is the geometric operation of producing a rectangle. Høyrup translates its Babylonian term as "to make [two segments *a* and *b*] hold each other".[2] This is sometimes expressed in the texts as "I have built a surface" (ibid.).

And division as an arithmetical operation in its own right and as the reverse of multiplication did even not exist in Babylonian mathematics. There were instead several procedures which exerted for certain types of magnitudes some analogous functions (ibid., 27 ff.).

GREECE AND EUCLID

The controversy of the 1970s and the 1980s as to whether a "geometric algebra" existed in ancient Greece, and in particular in Euclid's *Elements*, has revealed somewhat analogous limitations and the contextual nature of the allegedly arithmetical operations within this mathematics. Since 1975, Sabetai Unguru has attacked the largely accepted view similar to Neugebauer's interpretation for Babylon that Greek – and especially Euclid's – geometry was really nothing but algebra dressed up geometrically. Whereas Unguru first explored historical, philosophical, and linguistic arguments against this received view, he questioned its mathematical justification in two seminal papers in 1981 and 1982, jointly with David Rowe (Unguru/Rowe 1981 and 1982). The starting point for this endeavour was the consideration that

> the existence of a coherent system of arithmetical operations is a necessary (though not sufficient) precondition for the existence of any *system* of algebra (ibid. 1981, 4).

Investigating whether there was a coherent system of arithmetical operations in Greek mathematics, one notes a striking generalisation with regard to Babylonian times: addition and subtraction now prove to be well-defined, universal and unrestricted operations. Both operations are applicable to all kinds of quantities, numbers as well as magnitudes – provided that these are – within one operation – of the same kind (ibid., 14 ff.). The multiplication operation is not general, however, and more restrictive. In fact, as outlined above, multiplication is explicitly defined for numbers only and can at best be extended to multiplying quantities by numbers. The reason for this restriction is that only thus can the homogeneity between the concerned terms be maintained. And **homogeneity** between the magnitudes is the basic precondition for all operations in Greek geometry. Thus, one will never find, for instance, a line added to a rectangle since these are not homogeneous figures.

The consequence of having homogeneity as the fundamental property which any operation has to possess is that it destroys the likewise basic assumption of the supposed "geometric algebra" – namely, that the operation of rectangle formation can be understood as the missing multiplication of a line segment by a line segment. Since the product would be two-dimensional, it would not be homogeneous with the one-dimensional factors. Rectangle formation is hence an entirely geometrical procedure and does not constitute the missing generalised multiplication. (ibid., 30 – 31).Within the four arithmetical operations addition, subtraction, multiplication, and ratio formation (as a form of division) which form a system when applied to homogeneous quantities (i. e. of the same dimension), multiplication plays an exceptional rôle:

> Whereas addition, subtraction, and ratio-formation are defined for arbitrary pairs of homogeneous magnitudes, multiplication requires that [at least] one of these magnitudes be a number (ibid., 28).

Grattan-Guinness has recently systematised these results with respect to the restrictedness of the various operations; he also printed a table showing for which of the three types of Greek quantities – numbers, magnitudes, and ratios – which operation is defined and which different meaning it entails according to the respective type of quantity. It can thus be nicely demonstrated that it is mainly due to the missing generality of multiplication that there was no algebra in Greek mathematics (Grattan-Guinness 1996, 371).

EARLY MODERN EUROPE

As far as early modern Europe is concerned, recent research into the history of mathematics has shown that the problem of establishing a general operation of multiplication remained unsolved: due to the fact that it was mainly geometric quantities which were the subject of arithmetical operations, the homogeneity of the product was still violated.

The seminal tendency for mathematics in the early modern period was to strive for a fusion of arithmetic, algebra, and geometry, and to remove or at least to diminish the fundamental differences between numbers and arithmetical calculation, and geometrical magnitudes and construction respectively. One was prepared to suspend the strict methodological prescriptions of Greek geometry, and to relax the restrictions of multiplication.

Whereas the introduction of a unit length had been excluded in Greek geometry (Unguru/Rowe 1981, 21-22), one now employed numbers in practical geometry and units of lengths and areas so that, at least in practical contexts, multiplication of geometrical magnitudes was interpreted numerically (Bos 2001, 126). On the other hand, the dimensional interpretation of the operations in geometry remained an obstacle to the merging of arithmetical and geometrical methods; the use of numbers continued to be judged as inappropriate in geometry, from the viewpoint of theoretical mathematics (ibid., 131).

By the end of the sixteenth and the first half of the seventeenth centuries two important yet decisively different approaches to overcome these obstacles had been developed – one by François Viète (1540 – 1603) and the other by René Descartes (1596 – 1650). A major feature of both approaches was – as will come as no surprise now – a proposal for a solution to the multiplication problem.

Viète's intention was to provide an algebra for abstract magnitudes, thus uniting arithmetic and geometry. As Henk Bos has shown, his approach was inspired by geometry: he not only allowed the multiplication of two line segments, he even allowed an – in principle – unlimited scale of successive higher-dimensional species of magnitudes. Since in geometry the highest dimension for magnitudes was three – the dimension of space – Viète refrained, according to Bos, from an interpretation of products of more than three line segments (ibid., 148).

Viète expressed "scalar" quantities, resulting from successive multiplications of a length, a "side", by itself thus essentially powers, in dimensional terms:

- side, square, cube, square-square, square-cube, cube-cube, etc.,
  and "comparative" quantities, resulting from multiplying lengths and widths such as:
- length or width, plane, solid, plane-plane, plane-solid, solid-solid, etc.

The product of magnitudes increased in dimension, its dimension being the sum of the dimensions of the two factors. In Viète's approach, multiplication thus did not constitute a closed operation (Bos/Reich 1990, 188).

Descartes, for his part, chose a different route which preserved dimensional homogeneity. As is well known, he introduced – for the first time explicitly in non-practical mathematics – a **unit** length segment and thus succeeded in obtaining line segments as results, thus achieving closed operations. Descartes did this by reinterpreting Euclid's construction of the fourth proportional in book VI, 12, using the first element as his unit length. In the same way, he was able to interpret Euclid's construction of the mean proportional (VI, 13) as square root extraction. Descartes thus proceeded to solutions for quadratic equations and even to higher-order roots by means of mean proportionals. It is important to stress that Descartes – while operating with line segments and introducing a unit line segment – did not identify line segments with their numerically expressed lengths, apparently due to the problematical conceptual status of irrational numbers (Bos 2001, 293-296).

CONVOLUTIONS AND ALTERNATIVE SOLUTIONS

Notwithstanding Descartes' generalisation of the notion of multiplication, the conceptual problems in establishing a truly general multiplication persisted. Descartes' solution was still a special case: multiplying line segments by line segments. Practical problems, in particular in the considerably extending physics, implied multiplication between different kinds of magnitudes, however – like mass and velocity, or in commercial contexts quantities and prices. Practitioners would not be impressed by theoretical prohibitions to multiply arbitrary magnitudes by magnitudes.

In fact, the case of Bézout's arithmetic, criticised by Ampère, shows that the major problem in practice was to multiply different magnitudes – a problem which has not been studied at all by historiographers. Since Bézout was a well-known and typical author of textbooks one can reasonably expect that his approach to multiplying magnitudes was not an isolated one in his time. Alerted by the Bézout case, I found several authors exhibiting analogous practices. A particularly telling example is presented by the Italian mathematician Leonardo Salimbeni (1752 – 1823), since he attempted, in a *mémoire* of 1794, to clarify the notion of multiplication. His thinking started from the observation that Euclid's definition only concerned numbers and that a rigorous definition for magnitudes was still missing.

Salimbeni's solution was to propose a modification of Descartes' definition for the multiplication between line segments:

> The true and general definition of the algebraic multiplication is the following: One says that one magnitude multiplies another magnitude when one postulates: as the concrete unit of the multiplying magnitude is to this magnitude, so is the multiplied magnitude to the other magnitude which becomes produced (Salimbeni 1794, 484; my transl., G. S.).[3]

The modification seems to be a minor one: the unit is restricted just to the quantity of the multiplicator, which can be symbolised thus:

$$1_A : A :: B : P\,,$$

yet this shows at the same time the incoherence within this definition. The unit is only able to measure the magnitude A, not however magnitude B, which is of another kind. Salimbeni tried to circumvent this contradiction by giving the rank of a "theorem" to a *petitio principii*:

> Theorem: When one magnitude multiplies a magnitude, the product will be homogeneous with the multiplied magnitude (ibid., 487).[4]

The product should thus always have the same dimension as the multiplicand – or, as the first factor. Multiplication should maintain the dimension of the multiplicand, and should not be affected by the dimension of the multiplicator. After "proving" this theorem – basing himself on his "true" notion of multiplication – Salimbeni vehemently attacked traditional approaches, and in particular that of Viète, dismissing them as "falsissimos":

> Simple and obvious as this theorem is, it is nevertheless contrary to the commonly received idea. Who has not read again and again that a line multiplying a line produces a surface, and that a line multiplying a surface produces a solid? This all is entirely false [...] since we have proved that the magnitude produced is of the same kind as the multiplied one (ibid., 487).[5]

One wonders how Salimbeni would have reacted to reverting the order of A and B in multiplying, and will note that he was not only fully aware that, in his conception, multiplication was no longer commutative, but he seemed even quite proud of this achievement. In fact, he criticised the fact that a postulate that the product of two magnitudes should not depend on the order was accepted in Algebra:

> *when two magnitudes are multiplied by each other in a different order, the products are mutually equal.* [...] But, taken generally, this theorem is false; being true only in the case when both magnitudes A and B are of the same kind (ibid.).[6]

Salimbeni confidently explained his notion by an example analogous to the one by Bézout: having to multiply a length, 3 *piedi*, and a price, 70 *libbre*, the product will be either 210 *piedi* or 210 *libbre* (ibid.).

It is remarkable to see the notion of non-commutativity emerging through attempts to generalise the notion of multiplication. On the other hand, these efforts did not lead to the desired generalization; rather, they aggravated the problem. The eventual solution was achieved by a radical algebraisation, by separating numbers from geometric and other magnitudes. For calculation, magnitudes became definitely represented by numbers (now based on the concept of real numbers). Moreover, the introduction of metrical units made questions of conversion less salient. And the triumph of pure mathematics relegated dealing with magnitudes to fields of application, of a no longer foundational nature.

It often goes unnoticed that – this strict algebraisation notwithstanding – an alternative approach to generalising multiplication has been implemented in the history of mathematics: in the realm of **geometry**. In fact, it was a major motivation for Hermann G. Graßmann (1809 – 1877) – when elaborating his *Ausdehnungslehre* since the 1830s – to achieve a general notion expressly of multiplication for geometric magnitudes. The notions and theories introduced and developed by him – including the first coherent notion of non-commutativity – led eventually to the successes of linear algebra and multilinear algebra. There are effectively alternative paths in the history of mathematics – even for such apparently straightforward processes as generalisation.

Institut für Didaktik der Mathematik, Universität Bielefeld

NOTES

[1] The second, parallel series of his Cours de Mathématiques was intended for the training of engineers.

[2] This comes close to the formation of rectangles as defined in Euclid's book II where rectangles are formed by two segments which "contain" them (in German somewhat clearer: "umfassen") (Heath 1956 I, 370).

[3] "La vera e generale difinizione della moltiplicazione algebricha è questa: Una grandezza dicesi moltiplicare una grandezza, quando facciasi come l'unità concreta della grandezza moltiplicante alla stessa, così la grandezza moltiplicata ad un'altra grandezza che si produce."

[4] "Teorema: Se una grandezza moltiplichi una grandezza; il prodotto sarà omogeneo alla grandezza moltiplicata."

[5] "Quantunque semplice e manifesto sia questo Teorema, egli è però contrario ad una idea comunemente ricevuta. chi è, che non abbia molte e molte volte letto: che una linea moltiplicando una linea produce una superficie, e che una linea moltiplicando una superficie produce un solido? tutte cose falsissime. [...] poichè abbiamo dimostrato che la grandezza prodotta è dello stesso genere della moltiplicata."

[6] "che se due grandezze con vario ordine moltiplichinsi insieme, i prodotti sono uguali fra loro. [...] Ma questo Teorema preso in senso generale [...] è falso; non essendo vero che nel caso in cui le le grandezze A [e] B sieno dello stesso genere."

REFERENCES

Archives de l'Académie des Sciences, Paris – Papiers de A.-M. Ampère. -chem. Cart. 1 Chap. 1, chem. 3 : Précis élémentaire D'arithmétique. [Antoine Arnauld], *Nouveaux Elémens de Géométrie* (Paris: Ch. Savreux, 1667).

Bézout, Etienne (1800). *Cours de Mathématiques, à l'usage de corps de l'Artillerie. Tome premier: Contenant l'Arithmétique, la Géométrie et la Trigonometrie rectiligne.* Nouvelle édition par Guillard (Paris: Richard, Caille et Ravier, an VIII).

Bos, Henk, & Reich, Karin (1990). Der doppelte Auftakt zur frühneuzeitlichen Algebra: Viète und Descartes. In Erhard Scholz, *Geschichte der Algebra.* Mannheim: Bibliographisches Institut.

Bos, Henk (2001). *Redefining geometrical exactness: Descartes' transformation of the early modern concept of construction.* New York: Springer.

Grattan-Guinness, Ivor (1996). Numbers, Magnitudes, Ratios and Proportions in Euclid's *Elements.* How did he Handle Them? *Historia Mathematica 23*, 355-375.

Heath, Thomas L. (1956 <1926>). *The thirteen books of Euclid's "Elements".* Transl. from the text of Heiberg with introd. and commentary. Vol 1 Introduction and books I, II. Vol 2 Books III – IX . – Unabridged and unaltered republ. of the 2nd ed. New York: Dover.

Hofmann, James R. (1996). *André-Marie Ampère.* Cambridge: Cambridge University Press.

Høyrup, Jens (2002). *Lengths, Widths, Surfaces. A Portrait of Old Babylonian Algebra and Its Kin.* New York: Springer.

Nissen, Hans J.; Damerow, Peter; & Englund, Robert K. (1993). *Archaic bookkeeping: early writing and techniques of economic administration in the ancient Near East.* Chicago: University of Chicago Press.

Salimbeni, Leonardo (1794). Intorno alla Moltiplicazione ed alla Divisione Algebraiche. In *Memorie di Matematica e Fisica della Società Italiana*, Tomo VII. Verona: D. Ramanzini, 482-507.

Tropfke, Johannes (1980). *Geschichte der Elementarmathematik.* 4. ed. Vol 1, Arithmetik und Algebra. Vollständig neu bearbeitet von Kurt Vogel, Karin Reich, Helmuth Gericke. Berlin: de Gruyter.

Unguru, Sabetai, & Rowe, David (1981/1982). Does the Quadratic Equation have Greek Roots? A Study of "Geometric Algebra", "Application of Areas", and Related Problems. *libertas mathematica* 1, 1-49; 2, 1-62.

CIRCE MARY SILVA DA SILVA DYNNIKOV

SOME GERMAN CONTRIBUTIONS TO MATHEMATICS RESEARCH IN BRAZIL

Abstract. This text discusses the institutionalization of mathematics research in Brazil with emphasis on the cooperation networks involving Brazil and Germany. These cooperation networks began systematically at the Instituto de Matemática Pura e Aplicada [IMPA] (Institute of Pure and Applied Mathematics) in Rio de Janeiro, Brazil. An example is the work of the Brazilian mathematician Paulo Ribenboim at Bonn University and his contact with the German mathematician Wolfgang Krull. Another example is the contribution from the German mathematician Otto Endler in Brazil. Another collaboration arose during the 1990s: Michael Otte and his work preparing Brazilian researchers for the investigation of the history of mathematics.

Key words: history of mathematics, mathematics institutions, mathematics in Brazil.

Cooperation networks involving Brazil and Germany began in the 1950s. On one side, there is Brazil, a country without any tradition in mathematics research, that, nonetheless, set upon governmental initiative a specialized institute of mathematics research at the beginning of the 1950s, well in advance of such moves in several European countries. On the other side, there is Germany, a country with a long tradition in mathematics research, that nonetheless did not set up its specialized institute in Bonn until the 1980s.

In the 1930s and 1940s, mathematicians from Italy, France, and the United States, among other countries, had already cooperated with the Brazilian mathematicians for short and long periods. Such cooperation favoured the formation of university professors and researchers in the faculties of philosophy created in the 1930s. Experienced mathematicians such as Luigi Fantappiè, Giacomo Albanese, Jean Dieudonné, Andre Weil, Oscar Zariski, Alexander Grothendieck, Marshall Stone, Adrien Albert, and Antonio Monteiro, among several others, had been in Brazil, mainly in São Paulo and Rio de Janeiro. However, similar relations were not established with German mathematicians. It was only after the World War 2, when IMPA was founded, that cooperation links began between the two countries. One of the first German mathematicians to visit IMPA was the applied mathematician Lothar Collatz from the University of Hamburg. Collatz's stay in Brazil was short, but it already disclosed a certain trend toward establishing cooperation acts with the German-speaking countries.

In this chapter, I shall start by reporting on the creation of some specialized mathematics institutes on the international level. Afterwards, I shall present the conditions that had favoured the foundation of IMPA in Rio de Janeiro. I shall then go

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 287 – 298.

on to address the following questions: How were the first contacts established with German mathematicians? How did the scientific cooperation network between the two countries develop? Who were the first German mathematicians to come to Brazil, and, among them, who were the ones with whom a long-time cooperation was established? Answers to these questions are only explored here because the goal is not to document the entire cooperation, but only to point out how the first steps in this process occurred.

CREATION OF SPECIALIZED MATHEMATICS INSTITUTES AROUND THE WORLD

Pyenson and Sheets-Pyenson (1999) relate the word discipline, which possesses several meanings, to the word authority. In our context, the subject matter of mathematics is what we considered as the discipline of mathematics. *Disciplines* function according to principles; to general and abstract rules. In contrast, *institutions* operate according to corporate structures and private convenience. In a contraposition between both, it can be said, like Pyenson and Sheet-Pyenson, that disciplines display an abstract solidarity whereas institutions show an earthly and organic solidarity. "*Exploring the authority of disciplines and institutions to elaborate the counterpoint of tradition and innovation, in Kuhn's word, is the project that has animated historians of science since* the *1960*s" (Pyenson & Sheet-Pyenson, 1999, 20-21). The present chapter takes the same perspective to analyze the discipline of mathematics in a scientific institution directed mainly toward research.

Institutes devoted primarily to mathematics research only appeared in the 20th century. Some of the main pioneers were: the Steklov Institute (1919, Saint Petersburg), the Institute for Advanced Study (1930, Princeton), the Institute of Pure and Applied Mathematics (IMPA, 1952, Rio de Janeiro); the Institut des Hautes Études Scientifiques (IHES, 1958, Bures-sur-Yvette), and also the Max-Planck-Institut für Mathematik (1981, Bonn). Earlier initiatives had not been very successful. For example, in Sweden, in 1916, Gösta Mittag-Leffler and his wife Signe established the Mittag-Leffler Institute. Although incorporated to the Royal Swedish Academy of Sciences in 1919, financial difficulties and Mittag-Leffler's death in 1927 led to an almost total lack of activity apart from its library. It was only in 1969 that Lennart Carleson managed to turn Mittag-Leffler's dream into a concrete one. The next section sketches two specialized institutes that continued to function without interruptions.

STEKLOV INSTITUTE (SAINT PETERSBURG)

In 1919, the Mathematical Cabinet of the Academy of Sciences of Russia was set up at Saint Petersburg. This was due to the initiative of V. Steklov, a mathematician and theoretical mechanic and vice-president of the Academy of Sciences of Russia. Despite the state of civil war, resources for the creation of this cabinet had been provided by a decree of Lenin. Later, in 1921, the Institute of Mathematics and Physics was created. This contained not only the Mathematical Cabinet but also the Laboratory of Physics and a Network of Cosmic Stations at the Academy. Steklov was its first director. In 1926, after Steklov's death, the institute started to be called the Steklov Institute.

THE CREATION OF IMPA

In Brazil, in the 1950, university professors and researchers were aware that entrepreneurs had little interest in producing scientific knowledge and paid little attention to science. Thus, the possibility of promoting science and technology was restricted to the action of the state. Therefore, the rise of agencies to promote science and technology would require the political initiative of the government itself. In countries like England (1916), the United States (1918), and Canada (1916), National Councils of Research had appeared during World War 1 as agencies that could guide technological and scientific production.

In the middle of the 19^{th} century, science acquired supremacy in the western thought, and it became socially recognized for its instrumental potentials and its capacity to develop technology. In Brazil, after the World War 2, with the high *status* of science and technology, discussions intensified over the creation of an agency for promoting research. The pivot was the atomic bomb. For the historian Ana Maria Ribeiro de Andrade, the "*Second World War modified mentalities, revealed ideologies and intervened directly with the scientific work*" (Andrade 1999, 15). The Brazilian army was pledged to produce nuclear energy in Brazil. Several societal groups joined together to produce nuclear energy and develop science in Brazil. These included scientists, the army, government and scientific societies. Special attention must be given to admiral Alvaro Alberto da Motta e Silva. After negotiations with other countries, possessing the monopoly on atomic energy technology, he tried to find the path that would ensure Brazil's attainment of its scientific and technological potential.

In Brazil, the National Council of Research (CNPq, *Conselho Nacional de Pesquisa*) was created in 1951, directly linked to the Presidency of the Republic, as an autarchy, with corporate entity and administrative autonomy. The creation of the CNPq forms a watershed in Brazil; it is possible to speak of science in the country before and after the rise of this Council. Before its foundation, only a few states developed some scientific research. Within 10 years, CNPq had already created several research institutes and been able to promote scientific investigations outside the axis Rio de Janeiro-São Paulo, thus expanding scientific research throughout the country.

One of the first institutes that the CNPq created was IMPA – Institute of Pure and Applied Mathematics. It was established in 1952, and had a well-defined goal: "Teaching and scientific inquiry in the field of pure and applied mathematics, as well as the diffusion and rise of the mathematical culture in the country" (Arquivo CNPq, t.6.3.002). However, its existence was only acknowledged by Decree 39,687 of August 7, 1956. The Institute would be led by a director, nominated in commission by the president of the CNPq, and would also have a Supervising Council, composed of six members, whose function would be to guide the institute scientifically, technically, and administratively. The IMPA started its activities in 1953 without its own building and as a guest of the Brazilian Centre of Physics Research (CBPF).

Even though, according to its statute, the institute's goal was to develop research in both pure and applied mathematics, its first 20 years of existence focussed predominantly on pure mathematics. Since 1960, a special prominence has been given to research on dynamic systems. The IMPA started its activities with Lélio Gamma as director, Maurício Peixoto and Leopoldo Nachbin as titular researchers, and Paulo Ribenboim and Carlos Benjamin Lyra as assistant researchers. Carlos Lyra was a professor at the University of São Paulo (USP) and did not take an active role as a researcher in IMPA.

The two first decades of the IMPA were an attempt to empower institutions. IMPA sought to promote several institutions with the main goal of promoting research and international interchange. CNPq provided financial support to researchers, and IMPA's main concern was to encourage new mathematicians. CNPq provided the financial resources to pay the salary from of foreign professors who where guests at IMPA for shorter or longer periods.

In 1960s, the BNDE (National Bank of Economic Development) was a major funder of IMPA. Later, according to Lindolpho Dias, a former director of this institute, the FUNTEC (Technological National Foundation) was established and covered 70 % of the personnel payroll of IMPA. Furthermore, another institution for promoting research, the [FAPESP] (São Paulo State Foundation of Support to Research) granted funds. In this same decade, beyond the OEA, the Ford Foundation, the Foundation SLOEN, and the National Science Foundation granted resources to IMPA. The Ford Foundation was created in 1936, but its international expansion occurred only in 1950. The National Science Foundation was founded in Washington in 1950.

One can observe from these facts that IMPA received not only strong support from Brazilian governmental institutions such as the CNPq but also from foreign foundations that funded guest residences of prominent international mathematicians at the institute. In sum, IMPA received a lot of support in promoting the goal of the development of mathematics research. Without these resources, IMPA would have hardly survived and grown. After this brief overview, I shall now introduce the people who played a central role in establishing the cooperation network between Brazil and Germany, and consider the implications of such cooperation for mathematics research in Brazil.

PAULO RIBENBOIM AND WOLFGANG KRULL (1899 - 1971)

The first person to deserve special attention is Paulo Ribenboim, who was born in 1928 in Recife, Pernambuco. He got his Bachelor degree in mathematics from the National Faculty of Philosophy at the University of Brazil in Rio de Janeiro in 1948. In this same year, he attended lessons on integral equations and grids given by the Portuguese mathematician Antonio Aniceto Ribeiro Monteiro. With Leopoldo Nachbin he attended lessons on topological vector spaces; with the American mathematician Adrian Albert, he attended a course on Galois theory; and with the American mathematician Francis Murnaghan, he studied vector geometry.

Since 1950, different funding agencies had granted scholarships to Paulo Ribenboim. The first scholarship was for studying with Jean Dieudonné at Nancy, in France. There he met Laurent Schwartz and Alexander Grothendieck. In 1951, he received a 6-month scholarship from UNESCO. He enrolled in a course of Jean Delsarte, studying Lie groups; with Jean Dieudonné, he studied algebraic numbers and valuations; and with Laurent Schwartz, he studied distributions theory. In 1952, he received another scholarship, this time from CNPq. In 1953, he got a scholarship from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior – Coordination of Enhancement of Personnel from Higher Education), another funding agency from Brazil created in 1951. At that time, the scholarship was to carry on studies at the University of Bonn, in Germany. In 1953, Leopoldo Nachbin recommended him and affirmed that he was the best young mathematician to graduate from the National Faculty of Philosophy at the University of Brazil in Rio de Janeiro:

> He is a very young competent mathematician with a solid culture who is already engaged in research in the field of modern algebra. It is my opinion that he is the best student in mathematics produced by the National Faculty of Philosophy, from Rio de Janeiro (Arquivo Leopoldo Nachbin).

In the University of Bonn, he started to work under the supervision of the algebraist Wolfgang Krull, in the summer semester of 1953/1954. Krull was very impressed by the scholarship holder as a letter to the National Council of Research shows:

> My impression of Mr. Ribenboim is very positive, and I would really welcome to have him as a co-worker for a longer period. Especially, I like his profound mathematical knowledge, the logical präcision of his thinking and the attentiveness of his proofs, his personal initiative in attacking problems and making them his own, and his great diligence. (Mein Eindruck von Herr Ribenboim ist sehr günstig, und ich würde es sehr begrüssen, wenn ich ihn noch längere Zeit als Mitarbeiter haben könnte. Besonders schätze ich an Herrn Ribenboim seine gründliche mathematische Bildung, die logische Schärfe seines Denkens und die Sorgfältigkeit seiner Beweisführung, seine persönliche Initiative in selbständigen Angreifen von Problemen und seinen grossen Arbeitseifer.)
> (Krull, 12.5.1954, translated into English by the editors) Source: Paulo Ribenboim: Career up to 1995, Queen's University, Ontário, 1995.

Krull assisted Ribenboim personally after his arrival in Bonn. On the very first day, Krull advised him to study Jaffard's articles on rings of Dedekind. The first important results came from the study of Krull "Allgemeine Bewertungstheorie." This work was so fertile in terms of further results and methods that it functioned as a permanent source in which one could seek inspiration for further studies in the area and also for dealing with problems in similar areas. Although Paulo Ribenboim worked with Krull from 1953 until 1956, he did not produce his doctoral dissertation until he returned to the University of São Paulo. Some of the publications resulting from these first studies in cooperation with Krull are:

- Sobre a teoria das valorizações de Krull. Boletim da Sociedade de São Paulo 11, (1956a).

- Un thèoréme sur les anneaux primaires et complètement integralement clos. Mathematische Annalen 130, 399-404, (1956b).
- Sur la théorie du prolongement des valuations de Krull. Mathematische Zeitschrift 75, 449-466, (1961).
- An Existence Theorem for Fields with Krull Valuations. *Trans. Amer. Math. Soc. 105*, 278-294, (1962).
- On the Existence of Fields with few discrete valuations. *Journal für reine und angewandte Mathematik 216*, 43-49, (1964).

Ribenboim's article published in the Mathematische Zeitschrift in 1957, is based on Krull's work of 1932. One of the main results of the evaluations of Krull concerns the theorem of the approach. The objective of Ribenboim's article was to establish the theorem of the approach for finite sets of valuations, not necessarily two by two-independent ones. The bibliographical references used were Bourbaki, Krull, Jaffard, and Nagata.

In 1956, Ribenboim returned to Brazil, and made the exceptional move of CNPq granting him another scholarship for a further year. In this same year, he entered the Academia Brasileira de Ciências [Brazilian Academy of Sciences] as an associated member. In August of 1957, under the formal supervision of Cândido da Silva Dias, he presented his doctoral work to the Faculty of Philosophy, Sciences and Languages at the University of São Paulo. His doctoral thesis was entitled "About the theory of valuations of Krull." Clearly, the real supervisor of Paulo Ribenboim's doctoral work was Wolfgang Krull, but as the work was presented in Brazil, it was necessary to comply with Brazilian laws and formalities requiring a Brazilian professor to be the doctoral supervisor. In this same year, IMPA appointed the new doctor of philosophy to head research.

In 1959, he received a scholarship from the Fulbright Commission to study at the University of Illinois. He remained in the United States, teaching in several universities until 1962, when his Visa could not be renewed. He accepted a post at Queen's University, Ontario, Canada where he remained until his retirement. In 1988, the Springer Verlag published Ribenboim's collected works in a book entitled "The book of the prime numbers record." The book originated from a course he gave at Queen's University in 1984.

It can be concluded that with the research of Paulo Ribenboim, IMPA entered a phase of contact with the modern algebra produced in Germany by one of its most brilliant researchers Wolfgang Krull, who was a disciple of the algebraist Emmy Noether. This established the connection with German mathematics. This connection would gain continuity through the German Otto Endler, who produced his doctorate under Krull's supervision. Later, this link would continue with Karl Otto Stöhr, also a disciple of Krull, who is currently at IMPA, doing research and supervising graduate students. Paulo Ribenboim was responsible for Otto Endler coming to Brazil.

OTTO ENDLER (1929-1988), A GERMAN IN BRAZIL

Otto Endler was born in 1929 in the city of Mikulasovice, Czechoslovakia,. In 1950, he entered the University of Bonn where he received his doctorate in 1955, with the thesis *"Differentiation in algebraischen Funktionenkörpern von n Variabeln,"* under the orientation of W. Krull. This is when he met Paulo Ribenboim, and the partnership with Brazil started. He came to IMPA in 1957, as an invited researcher. In the following year, Otto Endler gave a seminar on Riemann surfaces at the ITA (Instituto de Tecnologia da Aeronáutica – Institute of Aeronautics Technology) and another one on holomorfic functions at the IMPA. In 1959, Otto Endler published his first work in Brazilian periodicals: "On rings of fractions" in the Summa Brasiliensis Mathematicae, Vol. 4. In the following year, Otto Endler met the physicist Ana Maria Freire at ITA. They got married in this same year and had a son.

Otto Endler did not cut his links with the University of Bonn. He continued to make annual trips to Germany, and in 1962, he gained his postdoctoral dissertation with the thesis "Bewertungstheorie unter Benutzung einer Vorlesung von W. Krull." This thesis was published in Bonn, in the *Bonner Mathematische Schriften, n. 15*, eds.: Hirzebruch, F.; Krull, L.; Peschl, E.; Unger, H., in 1963. This dissertation entitled him to become a full professor at the College of Mathematics and Natural Sciences at the University of Bonn, in 1962. The work was made up of two parts: foundations of the theory of valuation and Galois theory. Among the 53 works cited by Otto Endler in his thesis, one finds 11 works from Krull and 7 from Ribenboim, indicating the important role of the work of these two mathematicians. In Brazil, he also published the following work in *Anais da Academia Brasileira de Ciências* [Annals of the Brazilian Academy of Sciences], in *Notas de Matemática*, and in *Atas dos Colóquios de Matemática*:

- Modules and Rings of Fractions. *Summa Brasiliensis Mathematicae, 4* (1959a), 149-182;
- On the inverse problem of Galois theory. Anais da Academia Brasileira de Ciências, 31 (1959b); 331-332;
- On pseudovalued complete Rings. Atas do Segundo Colóquio Brasileiro de Matemática, (1960), 9-10;
- The resolution of algebraic equations and the inverse problem of Galois theory. Notas de Matemática, 24 (1961), Rio de Janeiro.

In parallel, he published works on mathematics in international periodicals and publishing houses. More specifically, his book entitled Valuation Theory was published in 1972 by Springer Verlag. The collected work dedicated to the memory of Wolfgang Krull was written from his notes for a course given in 1969/70 at the University of Rochester (New York State). But, according to the author, the most of this book was worked out during his stay as a visiting professor in IMPA at Rio de Janeiro. He called attention in the introduction to the fact that many advanced topics in the theory of the valuations, such as the theory of the fields of maximal values and Ribenboim's generalization of the theorem of the approach were not explored in the book.

In a Brazilian book publication, entitled Theory of the Algebraic Numbers, Otto Endler offered an important contribution to training mathematics professors and mathematics researchers in Brazil. The "Brazilian Society of Mathematics" published this work in 1985 as a part of the Euclides project. This book is still important, and continues to be assigned as a compulsory reference in IMPA. In 2001, it was compulsory reading in the discipline of Algebra II for the master's mathematics course. For the discipline of Algebraic Theory of Numbers, from the doctoral program of mathematics from IMPA, the books from Otto Endler and Paulo Ribenboim are also given as basic references for students. In IMPA, Otto Endler supervised the following doctoral dissertations:

- Gervásio Gurgel Bastos – "A problem of existence of the valuations", in 1974. This work was published in the Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität. Vol. 41 (1974), 154-157.
- Antonio Jose Engler – "A study on dependence and valuation ring composition", in 1976. It was published in the Manuscripta Mathematica, Vol. 24 (1978), 83-95.

Endler also supervised in IMPA the following work for a master's degree: Gonzalo Bueno Angulo – "Ideals versus valuations in the introduction to the theory of the algebraic numbers," in 1982.

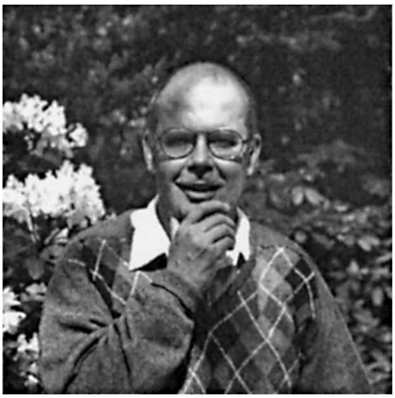
It can be seen that all the works supervised by Otto Endler are tied closely with the research on the theory of the valuations started by Krull. Endler invited Krull to come to Brazil in 1969. During this visit, Krull delivered lectures in IMPA. Otto Endler was responsible for the first collaboration between Brazil and Germany in the area of mathematics, with the support of the GMD [Gesellschaft für Mathematik und Datenverarbeitung, Society for Mathematics and Date Processing], in Germany. This German institution was founded in 1968 in Bonn. This bilateral accord was initiated in 1969, just one year after the foundation of GMD. The accord permitted several mathematicians to visit Brazil. Before this official agreement IMPA had received German mathematicians very sporadically. For example, in 1963, beyond Otto Endler, IMPA also received Wilhelm Klingenberg, a mathematician from the University of Göttingen, who spoke about "On closed geodesics on manifolds." German mathematicians who came to IMPA as a result of the accord with GMD include Gunter Bengel from the University of Münster, Ernst Ruh from the University of Bonn, Jürgen Herzog from the University of Essen, Herbert Heyer from the University of Tübingen, Lothar Collatz from the University of Hamburg, Heinz Jürgen from the University of Bielefeld, Klaus Foret from the University of Oldenburg, Henning Stichtenoth from the University of Essen, Jürgen Neukirsch from the University of Bonn, and Peter Roquete from the University of Heidelberg, Friedrich Hirzebruch from the University of Bonn, Heinz Helling from the University of Bielefeld, Alexander Prestel from the University of Konstanz, and Karl-Otto Stöhr from the University of Bonn. These mathematicians investigated not only in the area of algebra but also in the areas of algebraic topology, logic, theory of numbers, differential geometry, among others. In 1969, on the recommendations of Otto Endler, Jürgen Syman signed a professional contract with IMPA, and in 1972, he also recommended Karl-Otto Stöhr to sign a contract with IMPA. Krull also supervized Karl-Otto Stöhr in Bonn. Karl-Otto Stöhr is still in Brazil at the present

Karl-Otto Stöhr in Bonn. Karl-Otto Stöhr is still in Brazil at the present time. He has distinguished himself mainly through training a large number of mathematicians.

The German mathematicians referred to in this text were certainly not the only ones who contributed effectively to the development of mathematics research in Brazil. Several others need to be investigated and should have their stories told. But as mentioned at the beginning, this chapter does not claim to exhaust the theme of the German-Brazilian network in the field of mathematics. The basic aim is to show how this cooperative network became established and to report its first important steps.

A NEW COLLABORATION WITH GERMANY

In the last decade of 20th century, we observed the rise of another German collaboration with Brazil, but this time, in the area of mathematics education, a recent field compared with other disciplines and scientific fields. Mathematics education has only established its identity as a scientific field in the 20th century. It involves knowledge and research from several fields including mathematics, education, history of sciences, psychology, sociology, and anthropology. In the last decades, professionals trained specifically in mathematics education and in the history of mathematics have contributed a lot to the establishment of the field of mathematics education around the world. In the English-speaking countries, it is normal to use the terms mathematics education or mathematical education, whereas in some European countries, it is more common to use the term didactics of mathematics.



Michael Otte, professor at the University of Bielefeld, (see the photograph above) is one of the most active members of the Institut für Didaktik der Mathematik (IDM, Institute for Didactics of Mathematics) at this university. In 1990, he began collaboration with the Universidade Estadual Paulista (UNESP, Paulist State University), in Rio Claro, São Paulo. This collaboration has involved delivering lectures, seminars, and courses; supervising master's students in the graduate program on mathematics education; and also publishing articles and books. Michael Otte published a book in

Portuguese with the title *"O formal, o social e o subjetivo: Uma introdução à Filosofia e à Didática da Matemátic"* [The formal, the social, and the subjective: An introduction to philosophy and didactics of mathematics]. His interest, great involvement, and dedication to training of students have extended to other Brazilian universities. He has also been actively engaged at Universidade Federal do Mato Grosso [Federal University of Mato Grosso], where he has supervised more than 10 master's students working on topics in the history of mathematics. Michael Otte often comes to Brazil and has received financial support from Brazilian and German agencies. From Brazil, he has received funds from CNPq and CAPES; from Germany, the DAAD (Deutscher Akademischer Austausch Dienst). This sort of financial cooperation has been made possible by programs to stimulate and enhance graduate programs and research.

More recently through the program of cooperation between Brazil and Germany (PROBAL – Programa de Cooperação Brasil Alemanha), Michael Otte has extended his activities to two universities in the state of São Paulo. These are the Universidade Estadual de Campinas (UNICAMP, State University of Campinas) and the PUC-SP (Pontifícia Universidade Católica de São Paulo, Pontifical Catholic University of São Paulo). He has collaborated in graduate programs supervising master's and doctor's theses from several students.

Michael Otte's initial contact with Brazil arose from his supervision of two Brazilian students enrolled in graduate programs in Germany funded by CAPES and CNPq. These students are Circe Mary Silva da Silva and Fernando Raul de Assis Neto. Both have written their doctoral theses on topics in the history of mathematics in 1991 and 1992. Circe Mary Silva da Silva's thesis (1991) is – "Positivismus und Mathematikunterricht: Portugiesische und Französische Einflüsse in Brasilien im 19. Jahrhundert" (Silva, 1999) and Fernando Raul de Assis Neto's thesis (1992) is – "Géométrie de Position – eine Studie zum Werk von Lazare Carnot" (1753 – 1823).

The return of these professors to Brazil with their PhDs had an important impact in the universities where they worked. Circe Mary Silva da Silva has worked in the last years at Federal University of Espírito Santo and Raul Fernando Assis Neto has worked at Federal University of Pernambuco. They returned with experience in conducting research in the history of mathematics acquired in Germany. These Brazilian professors have helped to promote and establish the field of mathematics education in their universities and in Brazil. They are contributing particularly to research on the history of mathematics in Brazil and helping to prepare master's students in graduate programs in different states (Espírito Santo and Pernambuco). They are also contributing to the advance of other research fields in the country. They maintain contact with German investigators and have organized symposiums and conferences to bring some of them to Brazil, as well as returning to Germany themselves for short periods.

In the last years of the 1990s, new members have joined the group of German researchers who maintain contact with Brazilian professors working mainly in history of mathematics. These include Gert Schubring, also from University of Bielefeld. Schubring is the German coordinator of the collaboration project between Brazil and Germany (PROBAL) in the area of history of mathematics. The Brazilian coordinator is João Bosco Pitombeira de Carvalho.

FINAL CONSIDERATIONS

In the decade 1957 – 1967 IMPA consolidated itself as a specialized research institute in Brazil. It was also during this time that several international partnerships were established. Furthermore, it was in 1962, that IMPA established officially its graduate program in mathematics.

Paulo Ribenboim was responsible for the initiating the productive knowledge exchange with German mathematicians since way back in 1953. Nonetheless, it is impossible to measure properly the scientific production of German supervisors and their corresponding Brazilian students or disciples at this stage. It is still too early to evaluate the multiplicative effects, results, and power of such scientific networks between Brazilian and German mathematicians. And more research in other Brazilian institutions and in the other states needs to be implemented with respect to the influential persons involved in this mathematics research network, as well as a more through investigation at the network dealing with the history of mathematics.

The contribution of international mathematicians was without doubt an important aspect in establishing IMPA as a well-known research institution. And those international researchers played a crucial role for the development of mathematics research in Brazil. Nonetheless, without the presence of local scientific leaders taking the first steps to the form the initial nucleus of mathematics research, the international contribution would not have had such a strong impact. Among the principal Brazilian mathematicians who performed this vital role in mathematics research we cite, for example: Leopoldo Nachbin, Maurício Peixoto, Elon Lages Lima, and Manfredo Perdigão do Carmo. These important leaders not only served to make crucial links and implement cooperative networks between Brazil and abroad; they were also the support basis for the development of mathematics research in IMPA. The research work developed by these Brazilian mathematicians has been acknowledged of their peers on both national and international levels as an important contribution to the field of mathematics. Another relevant point to consider refers to the political negotiation to get support for IMPA by Brazilian personalities, such as Lélio Gama, Lindolpho Dias, Leopoldo Nachbin, and Jacob Palis. These were the first ones to firm accords to get financial support to IMPA that could assure the attainment of the main goals of this institute. Naturally, without the money granted through these agreement, IMPA would not have gained its reputation as an important specialized research institute in mathematics.

While the history of mathematics research is relevant and important being able to develop the skills to write and implement such research requires a solid theoretical and philosophical position. The theoretical foundation on the history and philosophy of mathematics developed by Michael Otte provided Brazilian researchers with a framework for implementing further research in the field.

ACKNOWLEDGEMENT

The work reported in this text received financial support from CNPq. All opinions and information presented here are the sole responsibility of the author. Special thanks go to Vânia Maria Santos-Wagner for her editing comments and translating the work into the English language.

Universidade Federal do Espírito Santo (Brazil)

REFERENCES

Andrade, A. M. R. (1999). *Físicos, mésons e política: A dinâmica da ciência na sociedade*. Rio de Janeiro: Hucitec/MAST/CNPq.

Arquivo Conselho Nacional de Pesquisa (Museu de Astronomia e Ciências Afins).

Arquivo Leopoldo Nachbin. (Museu de Astronomia e Ciências Afins).

Endler, O. (1959a). Modules and Rings of Fractions. *Summa Brasiliensis Mathematicae 4*, 149-182.

Endler, O. (1959b). On the inverse problem of Galois theory. *Anais da Academia Brasileira de Ciências* 31; 331-332.

Endler, O. (1960). On pseudovalued complete Rings. *Atas do Segundo Colóquio Brasileiro de Matemática*, 9-10.

Endler, O. (1961). The resolution of algebraic equations and the inverse problem of Galois theory. *Notas de Matemática* 24, Rio de Janeiro.

Endler, O. (1963). Bewertungstheorie unter Benutzung einer Vorlesung von W. Krull. Bonn, *Bonner Mathematische Schriften*. 15, Org.: Hirzenbruch, F.; Krull, L.; Peschl, E.; Unger, H.

Endler, O. (1972). *Valuation Theory*. Berlin: Springer Verlag.

Endler, O. (1985). *Teoria dos Números Algébricos*. Rio de Janeiro: Projeto Euclides, IMPA.

Favero, M. L.; Peixoto, M. C. & Silva, A. E. G. (1991). Professores estrangeiros na Faculdade Nacional de Filosofia, RJ (1939 – 1951). In *Cadernos de Pesquisa*. São Paulo (*78*), agosto, 59-71.

Gomes, J.M. & Santos, G.T. (1988). Otto Endler in Memoriam. In *Matemática Universitária 7*, junho, 1-6.

Krull, W. (1999). *Gesammelte Abhandlungen*. Berlin: Walter de Gruyter.

Otte, M. (1993). *O formal, o social e o subjetivo*: Uma introdução à filosofia e à didática da Matemática.(Transl. *Das Formale, das Soziale und das Subjektive. Eine Einführung in die Philosophie und Didaktik der Mathematik*, 1994). São Paulo: Editora da UNESP.

Pyenson, L. & Sheets-Pyenson, S. (1999). *Servants of nature: A history of scientific institutions, enterprises and sensibilities*. Londres: Fontana Press.

Ribenboim, P. (1956a). Sobre a teoria das valorizações de Krull. *Boletim da Sociedade de São Paulo 11*.

Ribenboim, P. (1956b). Un théoréme sur les anneaux primaires et complètement integralement clos. *Mathematische Annalen 130*, 399-404.

Ribenboim, P. (1961). Sur la théorie du prolongement des valuations de Krull. *Mathematische. Zeitschrift 75*, 449-466.

Ribenboim, P. (1962). An Existence Theorem for Fields with Krull Valuations. *Trans. Amer. Math. Soc. 105*, 278-294.

Ribenboim, P. (1964). On the Existence of Fields with few discrete valuations. *Journal für reine und angewandte Mathematik 216*, 43-49.

Ribenboim, P. (1988). *The Book of the prime numbers record*. Berlin: Springer Verlag,

Ribenboim, P. (1995). Paulo Ribenboim Career up to 1995. Ontario: Queen's University.

Sarita, Albagli. (1987). Marcos institucionais do Conselho Nacional de Pesquisas. In *Perspicillum, Rio de Janeiro* 1(1), 1-166.

Silva, C. M. S. (1999). *Matemática Positivista e sua difusão no Brasil* (Translation from the doctoral thesis of 1991). Vitória: Edufes.

Silva, C. M. S. (2000). A Faculdade de Ciências e Letras da USP e formação de professores de Matemática. In CD-Rom. *Anais da 23ª Reunião Anual da ANPED*, Caxambu.

ANDREAS DRESS

DATA STRUCTURES AND VIRTUAL WORLDS

On the Inventiveness of Mathematics

Abstract. It is argued that the real impact of mathematics is its power to shape concepts that allow us to understand the real world by manipulating virtual worlds.

Key words: conceptualisation, early hominids, Fourier analyses, virtual worlds.

A current notion is that it is the task of mathematics to find proofs for difficult conjectures and answers to tricky questions – presuming that the question or conjecture has already been phrased in the language of mathematics and that it should therefore be possible to find the requested answer or proof provided one is gifted with a sufficiently well-developed and well-trained power of combination.

According to this point of view, success or failure in mathematics depends exclusively on such achievements: who comes up (first) with the correct answer or finds a convincing proof merits recognition and fame, nothing else ever counts.

Doubtlessly, this view covers essential aspects of present-day mathematics. And just as doubtlessly, it falls short in many ways: Not alone that there may be more than one correct answer to some questions, and none at all to others, that some conjectures are false, and several, like e. g. the continuum hypothesis, may even be undecidable – more important is that the real progress achieved by finding an answer to an open question does often not so much consist in that very answer, but rather in the new methods developed to find it and the new concepts on which such methods are based, not to mention that neither question nor conjecture could even have been stated before the fundamental concepts and definitions to which they refer had been worked out properly.

A good case in point is Jean Baptiste Joseph Fourier's (1768 – 1830) theory of *Trigonometric Series* developed to solve the heat equation, published first in 1807. This theory, today called *Harmonic or Fourier Analysis,* has not only become one of the cornerstones of modern mathematics, yielding innumerable applications in physics and engineering as well as in pure and applied mathematics itself – from number theory to stochastics. Fourier's ideas presented also the crucial challenge for clarifying all of the most fundamental concepts of analysis during the 19th century like *function, continuity, convergence, integrability, differentiability* etc. And they led Georg Cantor (1845 – 1918) to establish his *Mannigfaltigkeitslehre* (theory of multifariousness) called *Set Theory* today which – after some initial hesitation – is now considered to provide the most efficient foundation for mathematical reasoning so-far.

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 299 – 304.

Moreover, Fourier's theory is a paradigm for the use of mathematical methods in *data analysis:* It permits (i) to completely decompose a complicated signal into simple elementary components in a purely formal way that can be fully automatized, (ii) to reconstitute that signal from those components, and (iii) to identify its most important components. In consequence, Fourier theory allows to separate the *essential message* contained in a signal from random noise. Furthermore, the definition of what may and should be considered to represent a simple elementary component can be adapted flexibly to whatever operational symmetry constraints appear to be pertinent. This has been giving rise to further, not even presently fully exhausted possibilities of applying Fourier's ideas inside and outside mathematics (see Figure 1).

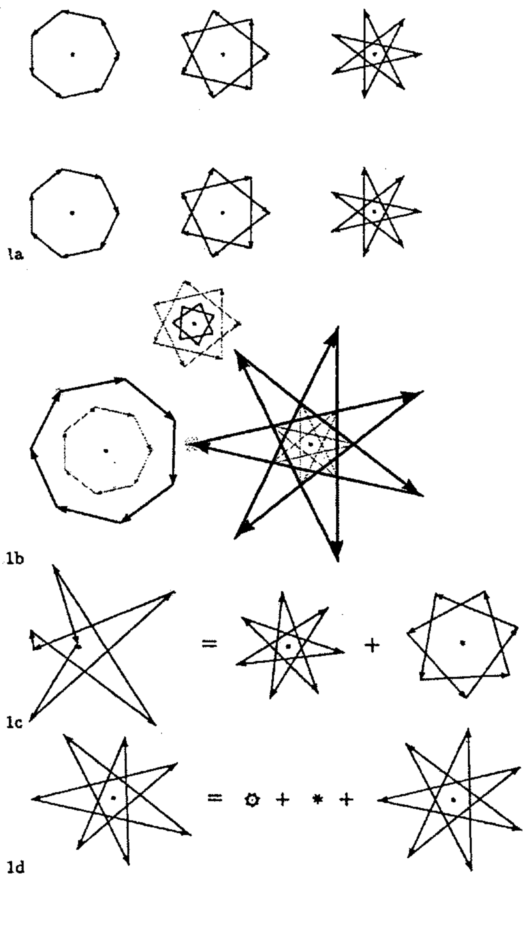


Figure 1

Explanation for figure 1: We are considering signals that consist of seven successive points $Q_1, Q_2, \ldots, Q_7$ in the plane, repeated in this order, having a fixed point of gravity $P = (Q_1 + Q_2 \ldots + Q_7)7$. Such a signal may be considered to be elementary if it (a) transforms into itself under rotation by 51.4285714 ... = 360/7 degrees around their point of gravity. Figure 1a shows six distinct elementary signals of this form from which all further elementary signals can be obtained by rotations around P and dilatations (Figure 1b). Fourier analysis now permits us to represent any arbitrary signal $Q_1, Q_2, \ldots, Q_7$ of the form described above as a superposition of those six elementary signals, appropriately rotated and dilated (Figure 1c). The (relative) seize of the 6 dilatations may not only permit to recognize a given signal as being – except for minor random noise – essentially just one of those six elementary signals, it also allows to quantify exactly the extent of perturbation due to random noise (Figure 1d).

In summary, the relevance of the *conceptual tools* introduced by Fourier and the progress of the mathematical sciences initiated by his *ideas* and *methods* outweighs by far his contributions to what was his primary concern, his – rather elegant – solution of the heat equation.

In this context, it is also whorthwhile to discuss the – at first glance rather banal – observation that there may be more than one correct answer to a given question, and that a problem may have more than one correct solution. It implies that instead of looking for the *one* correct solution, it is often feasible to concern oneself first with the collection of *all* possibly correct solutions, to list and to examine all of them, and to relate their properties to one another. It is one of the fundamental achievements of modern mathematics – if not the most decisive one – that it offers tools that permit to form and handle such collections. I. e., it permits to conceptually grasp and to explicate the *space* formed by those (potential) solution, and to produce, to analyze, and to axiomatically describe such spaces in rather concrete and constructive ways. In yet other words, it is the capacity of modern mathematics to design *virtual worlds* and to successfully use their structure for interpreting *real-world* data that is probably its most eminent achievement.

From an evolutionary perspective, such ability to design adequate and stringent virtual representations of our environmental reality presumably began to develop among our ancestors 4 to 5 million years ago: Separated by the East African Rift Valley from their own kind in the Central and West African rain forests, the predecessors of the present-day chimpanzees, our closest relatives today, they began to be exposed to a climate turning increasingly arid, transforming the eastern part of this rain forest into grass lands. Thus, the early hominids were compelled to descend from the gradually disappearing trees and to survive in new habitats which had already been occupied and very successfully exploited by many other mammals. The early hominids were weaker than the carnivores living there, and much slower than the latter's prey.

One feature, however, that they brought along into those savannahs as former climbers was a highly developed optomotoric system supported by a large visual cortex. This had enabled them to safely grasp a branch that – having already left their field of vision directed towards the next and second-to-next support for their treetop swinging – they could no longer envision but with their "inner eye". The hominids might have died out under these constraints. Instead, we systematically used and elaborated, in the course of evolution, this ability to constitute and to manipulate a *virtual world* before our inner eyes – a virtual world that was based on *those* traits of our environment that were most crucial to our existence and could be manipulated in such a way that successful strategies of coping with the real world resulted.

Certainly, there are a number of other frequently named factors which contributed towards our becoming humans: For instance, developing the ability to communicate verbally among ourselves about our respective virtual inner worlds (albeit only within certain limits, as we sadly experience again and again). Further, to plan actions as a group using such modes of communication basis was a crucial step accompanying this process. Other decisive boosts for the hominids' development were that the erect posture liberated hands for purposes of actually implementing what had been imagined, and that our "instinct of play" continued to grow ever stronger, allowing and training us to playfully generate and investigate virtual worlds. The foundation of the hominid success story, however, is in my opinion the human ability – culminating in mathematics and in the exact sciences it supports – of deliberately and progressively extending, if not overcoming, the boundaries and constraints of the real world by confronting it with suitably constructed virtual worlds.

Surely still within prehistoric times, this development led towards forming a tentative concept of *number.* Numbers constitute *the* data structure that is probably the most important one today for the virtual reconstruction of contexts and the shaping of virtual worlds which represent basically all of those traits of the real world that we really need to know about.

Once begun, the elaboration of the concept of numbers required many additional millenniums, and it can perhaps not even today be considered accomplished, as recent contributions by John Conway demonstrate. It is undisputed, however, that solving specific problems (as e. g. the Delian problem of doubling the cube) and computing specific numbers (the infinite alternating sum $1 - 1/2 + 1/3 - 1/4 + \ldots$ or like $\exp^{\pi i}$) makes sense only within the context of the number system, the space of all possible solutions of all numerically posed problems.

At the same time, our present number system forms the building material for the overwhelming majority of spaces (of solutions) designed by present-day mathematics. One case in point are, e. g., the separable Hilbert spaces describing the collection of *all* signals accessible by means of classical Fourier analysis.

Beyond that, mathematical disciplines like topology, algebra, combinatorics, or category theory offer a wealth of further possibilities for designing solution spaces for the most diverse classes of problems. Their focus often extends far beyond the usual tasks of mathematics dealing with basic quantitative relationships. Instead, they allow to grasp and to treat as well purely qualitative aspects:

The concepts of *groups* and of *group actions* for instance was developed in the 19th century. It provides a data structure that permits to virtually record, analyze, and classify practically all phenomena linked to *symmetry.* It can be viewed as the culmination point of a development that had started most certainly already in the neolithicum more than 4000 years ago – as stone artefacts shaped like octahedrons and dodecahedrons in the Ashmolian Museum in Oxford amply testify.

In the same vein, the concept of a *topological space* developed between 1850 and 1930 provides a data structure that permits to state and virtually handle concepts like shape without inadequate recourse to purely quantitative levels of description. Classifying all possible shapes under certain appropriate limiting conditions – say all knots, or all closed two-dimensional manifolds, or even all simply connected and closed three-dimensional manifolds – is one of the central tasks in this area. How difficult this is can be seen from the fact that methods – often just as profound as subtle and astucious like e. g. *algebraic topology* – which only serve the subordinate goal of enabling one to orient oneself within the virtual world of all imaginable shapes and to systematically distinguish different shapes from one another, already count among the secular achievements of mathematics in the last century.

In contrast, *combinatorial graph theory* provides an almost disappointingly elementary data structure that permits, however, to model and to construct the most complex networks and, thus, to virtually examine all sorts of interdependence between different parts of highly interconnected worlds.

And last, not least there is *category theory*. It encompasses all these three theories and simultaneously extends their analytic power, providing the most elaborate, powerful, and flexible tool for designing mathematically structured virtual worlds that we presently know.

It must be noted, however, that these abstract concepts designed for grasping and handling qualitative aspects often unfold their full power only in combination with numerically defined aspects:

Many achievements of group theory are based on representation theory including the theory of *group characters* – i. e. systems of numerical invariants that encode decisive properties of any given collection of symmetry operators forming a group.

Likewise, the continuous real-valued *functions* that can be defined on a topological space form the presumably most important invariant of this space. The existence of sufficiently many of such functions can be shown by rather abstract means for a surprisingly large class of spaces. And perhaps even more importantly: Topological spaces supplied with an appropriate *measure* that allows *integration* of continuous real-valued functions offer the opportunity to systematically simulate and investigate the phenomenon of probability, truely fundamental for so many aspects in present-day live.

And finally, the networks studied in combinatorics are of particular theoretical and practical interest – e. g. within the context of optimization theory – as soon as their elements (the nodes of the networks as well as the connecting elements or *edges* linking these nodes) are numerically quantified by suitably defined *weights* or *capacities* etc.

In summary, the thesis advanced here asserts that the crucial achievement of mathematics that has shaped our present culture in a fundamental way is based first of all on its ability – present already in the early evolution of the hominids – to design phenomenologically adequate and stringent conceptual or *virtual* worlds and data structures, and on their amazing penetrating power.

This corresponds to the fact observed by G.-C. Rota that much, if not the majority of mathematical work is devoted to the task of obtaining new perspectives on things already known, of exploring virtual worlds already discovered, and of detecting new and surprising connections between them – work that is concerned with making virtual worlds habitable and allowing people to feel at home in them.

It also corresponds to the fact that, in mathematics, one quite frequently does not attempt (as the example of Fourier analysis shows indeed) to find the correct proof for a theorem, but rather the correct conceptual framework for a proof, that is, finding out what has really been established by that proof.

And it corresponds finally to the fact that successful tricks are, as a rule, embedded sooner or later into a conceptual framework in such a way that one is almost routinely led towards these tricks once one accepts working within this framework as exemplified, e. g., by the ground-breaking work of N. Bourbaki.

Above all, however, this perspective offers a way to understand the *unreasonable effectiveness of mathematics in the natural sciences* discussed by Eugene Wigner (1902 – 1995) in a lecture given in 1959: the real strength of mathematics is its inventiveness with which it designs and investigates new worlds and data structures corresponding to them. Evolution would probably have eliminated us several million years ago without further ado if the virtual worlds designed by us had proved inadequate, that is, to keep within the metaphor presented above, if the branch imagined in front our inner eye had not actually been where we imagined it, thus saving us just in time from a disastrous fall. This metaphor also explains why the persistent experience of the might and power of our imagination has led to the misunderstanding – so popular in present-day social sciences – according to which the branch we grasp and which saves us from falling was only at this precise locality because we ourselves had imagined it to be there.

The question remains, however, whether our ability of confronting the real world successfully with virtual worlds has not meanwhile set a dynamic in progress that targets a perhaps disastrous fall of quite another kind, making our planet uninhabitable at a time at which we are – due to our steadily increasing power of manipulation – just on the point of subjugating it completely . Yet, this is another story better to be discussed in quite a different context.

Forschungsschwerpunkt Mathematisierung, Universität Bielefeld

HERMANN DINGES

VARIABLES, IN PARTICULAR RANDOM VARIABLES

Abstract. We argue that the concept of a variable as classically understood in mathematics and physics is didactically and scientifically inadequate for the needs of probability and statistics and we propose an alternative concept. A revision should have political implications for the relation between mathematical and statistical teaching.

Key words: axiomatics, foundations of probability, random variables, statistical teaching, variables.

It can hardly be doubted that the concept of a variable is fundamental and ubiquitous in analysis. The textbooks are very reserved with explanations or definitions, however. On the other hand, they take great pain in explaining and defining the concept of a function, a concept which is clearly secondary to the concept of a variable. The concept of a function (or a mapping) builds on the concept of variables; a function associates the value of the dependent variable with the value of the independent variable.

I find it remarkable that the textbooks do not say or suggest, that the concept of a variable may remain undefined just as the concept of a point in euclidean geometry. They rather seem to suggest that it must be left to the users of analysis to build up that concept of a variable which suits their needs. I am not aware of any didactical or philosophical discussion of this kind of an approach to mathematics, however.

At first sight the didactical strategy seems to work in present day elementary analysis. Different contexts go along with different conceptions of variables. Nobody is really bothered by the observation that variables in thermodynamics are not exactly the same as variables in mechanics or economics. I will argue here, that the disconcern with a mathematical concept of a variable has shown its Achilles-heal in the case of random variables. It seems to me, that present day education in elementary analysis is not open for an intuitive and workable conception of a random variable. This has lead to a very unfortunate situation which cries for action.

A very basic innovation is needed to pave the way for a broader and deeper understanding of stochastics. Mathematicians must understand that the common reductionist approach to the concept of a random variable is inadequate and that this must have consequences for the teaching of elementary analysis already at school level.

There are of course historical reasons for the divergence between the needs of stochastics in the practical sense and present day mathematics in the narrow sense. When probability around 1930 (after a dormant period of a century) became again a respectable field of mathematics, the teachers of mathematics wanted a definition: what is a random variable? Bourbaki did not offer a definition, he left stochastics aside. But for the mathematicians who believed in Bourbaki at that time it was clear,

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 305 – 311.

that a random variable is a measurable function (or a measurable mapping) from a measurable space $(\Omega, \mathcal{A})$ into another measurable space $(S, \mathcal{B})$. And they did not try to please traditional stochastic thinking. P. Halmos argued in these times that probability theory is legitimate mathematics, in as far that it is the theory of normed boolean algebras.

The mathematics departments have put probability theory into their curricula, because they were told in the sixties that mathematicians need competency in stochastics, when they were to leave the realm of genuine mathematics, in other words, when they leave the mathematics department. By this move the advanced students of mathematics towards the end of their studies after an intensive training in traditional analysis got the opportunity to learn the reductionist approach to stochastics. The majority of university mathematicians considered this as a concession which was not justified by the strive for mathematical progress.

On the other hand, the concession turned out to be a fairly small contribution to the needs of a better stochastical education. Only very few students got acquainted with the rich world of stochastic intuitions. Not much was achieved to assist students of the sciences working with unprecise data or students of economics dealing with decisions under uncertainty. A mathematics based philosophy of stochastical thinking didn't come to the attention of any kind of students. A gap was opened: the non-mathematicians defied measure theory, and the mathematicians defied "cookbook"-statistics, as they called stochastics without measure-theoretic foundations. In particular, mathematicians would not accept the "definition" of a random variable, which is commonly presented in books for non-mathematicians: *"A random variable is a quantity whose value depends on chance".*

More and more people have stressed in recent years that action has to be taken in order to provide a common ground for mathematical stochastics and stochastical modelling in the sciences. There is a remarkable article of two prominent statisticians: David S. Moore and George W. Cobb (2000). The introduction says:

> It has become a truism, at least among statisticians, that while statistics is a mathematical science, it is not a subfield of mathematics ... In mathematics, context obscures structure. In data analysis, context provides meaning. ...

They conclude their article:

> Mathematics, a core discipline, looks inward and risks being seen as increasingly irrelevant. Statistics, a methodological discipline, looks outward but risks being swallowed by information technology. Both professions have a stake in the survival of statistics as a subject informed and structured by mathematics.

The article also points out political consequences of a reconciliation:

> To mathematics statistics offers not only the example of an outward looking culture, but also entree to new problems ripe for mathematical study. To statistics, mathematics offers not only the safe harbor of organizational strength, but intellectual anchorage as well: mathematical understanding is an essential part of what distinguishes statistical thinking from most of the rest of information technology.

Mathematics is challenged both didactically-philosophically and politically in these days; and stochastics will undoubtedly be a focus of renewal. There are pleas in this direction also by prominent mathematicians.

David Mumford (former IMU-president, originally coming from the very pure end of mathematics) contributed to the volume "Mathematical Frontiers and Perspectives 2000" an article "The Dawning of the Age of Stochasticity." There he points out:

> The basic object of study in probability is the random variable, and I will argue that it should be treated as a basic construct, like spaces, groups and functions, and it is artificial and unnatural to define it in terms of measure theory.

Mumford challenges the teachers of probability:

> Put the concept of 'random variable' on center stage and work with manipulations of random variables wherever possible.

The challenge has been felt by many teachers of stochastics for years. In our courses at the Department of Mathematics at Frankfurt University, Goetz Kersting and myself have been working hard over 20 years to deal with random variables independently of measure theory. In more recent times we have got encouraging results both in technical aspects and in practical teaching. We feel that the successes of teaching have considerably improved since the curriculum advises the students of mathematics to take up stochastics early. In former times we had to deal with the problem that students with an extensive training in traditional analysis were disappointed or confused, when the course in stochastics didn't conform to the familiar scheme. These students insisted that they would understand much more easily on the basis of measure theory; there was little curiosity for stochastic intuition. Our younger students now find it less difficult to build up intuitions around the concept of random variables. In our elementary courses this concept is introduced in a way, which is open for enlargement fitting the needs of professional mathematical stochastics.

In the search for a flexible conception of a variable it is appropriate to recall the historical origins of analysis. It was Leibniz who invented functions. Newton on the other hand had curves in mind when he invented his calculus. Euler (1768) considered it fruitless to unite the conceptions in Newton's tradition with the ideas of the continental school. In early 19th century a variable was commonly understood as something which runs through a domain. It was in opposition to this dynamical conception when Bolzano gave a statical definition of continuity. This lead to the concept of a function "in the sense of Dirichlet" and finally to Cantor's set theory, forbidding any dynamical association.

The dynamical conception of a variable has survived in school mathematics. Unfortunately there it is so intimately connected with special functions on an interval that every extension of the intuition about variables seems to be difficult. (Already Riemann's idea of an abstract "multiply extended manifold" seems to be beyond the intuition of most of our high-school teachers). Thus there is very little hope that an intuitive conception of random variables might profit from the traditional concept of a variable in analysis. As far as I know, nobody has ever tried to teach an intuitive concept of a random variable using analogies with the algebraic concept of an indeterminate. Thus random variables have to find their own intuitive basis.

Modern analysis in the spirit of Cantor's set theory has variables of course; but nothing is varying here. The mathematician may insert a value of the independent "variable" into a function $f(\cdot)$ (which is seen as a black box) and he retains the value of the dependent "variable"; the evaluation is a singular act at every instance; the mathematician is choosing a value in an arbitrary, completely discontinuous act. For every singular x in a specified set the black box $f(\cdot)$ produces a singular $y = f(x)$.

In the case of a random variable we have an actual variability and this variability is due to chance. Chance inserts a "typical" value into the black box $f(\cdot)$ and the output is a "typical" value of the dependent variable.

$$X \xrightarrow{f} Y = f(X).$$

The technical term is not "typical", but rather "random". "Typical value" might be confused with "generic value" which is a technical term in geometry. So we say: The blackbox associates the random variable $Y = f(X)$ with the random variable X.

In stochastics the calculus of functions is more or less the same as in traditional analysis. The arguments of the functions have a different logical status, however. There are instances where the difference has technical consequences. For example: assume $Y = f(X)$. We may ask for conditions which guarantee the existence of a $g(\cdot)$ such that $X = g(Y)$. (In our "polish" calculus to be sketched below there exists an easy and intuitive criterion).

There exist teaching traditions for elementary statistics (or probability) in various applied fields. They all deal with random variables. And most of them seem to get along without any elaborate mathematics. One may object that all of them are fairly incoherent. But this is not the most serious objection. Looking closer one finds that the non-mathematical traditions tend to produce misconceptions which hamper a deeper understanding of stochastic modelling. For example: in certain traditions of teaching randomized algorithms, the concept of randomness (implicitly) includes stochastic independence or uniform distribution or both.

Almost all traditions of teaching elementary probability theory tie the concept of a random variable to a fixed "true" distribution. Many textbooks stick to the idea that the distribution is the only thing which is worthwhile considering when we speak about random variables; there, probability theory is the theory of probability measures. This has its price. We find that very often students who have been brought up in a narrow teaching tradition, have serious problems, when they have to think in wider terms. For example, the concept of an equivalent martingal measure, which is crucial in mathematical finance, presents enormous conceptual difficulties to most of these students, as they are used to think in terms of the familiar "true" distribution.

Let me now briefly sketch our proposal, how random variables might be conceived as objects within an axiomatic context. We start with four axioms which can be supplemented at various levels all compatible with professional mathematical stochastics.

- Every random variable takes values in a measurable space ("domain"). When X_1 and X_2 are random variables, then (X_1, X_2) is a random variable with values in the product space.
- When X is an S_1-valued random variable and $\varphi: S_1 \to S_2$ is a measurable mapping, then $\varphi(X)$ is an S_2-valued random variable.
- For every measurable B in the domain of X there exists a $\{0, 1\}$-valued random variable $1_{\{X \in B\}}$ which represents the event $\{X \in B\}$. The S-valued random variables X and Y are equal if and only if the event $\{X \in B\}$ is equal to the event $\{Y \in B\}$ for every measurable B (or a family of subsets which generates the σ-algebra).
- The set $\mathscr{A}$ of observable events is a σ-complete boolean algebra (Equality of events is given axiomatically and cannot be reduced to equality of more concrete mathematical objects).

It is true that the events can be identified with the $\{0, 1\}$-valued random variables. However the universe of all random variables (in any particular stochastic model) is much richer than the set of those random variables which can be constructed in an obvious way from the $\{0, 1\}$-valued random variables. Even in school stochastics, we need nondiscrete random variables, in particular random variables with values in $\mathbb{R}$, and random vectors, i. e. random variables with values in $\mathbb{R}^n$. In elementary university teaching we need random functions and random measures. Ultimately our axiomatic system admits random variables with values in all measurable spaces $(S, \mathscr{B})$ of a certain type $\mathcal{S}$. This type $\mathcal{S}$ has to be specified and I may say at this point, that my favorite class $\mathcal{S}$ is the class of polish spaces.

Beyond the four axioms stated above, one has to make a couple of non-obvious technical and didactical decisions. The teacher has to decide how far he wants to extend the perspective at a certain point – but he should be aware of what he is doing and he should be careful not to brain-wash his students at a level which cannot be extended in an intuitive way.

We believe that an essential point of stochastic reasoning would be missed, if one would teach a toy version which would require all the admissible domains to be denumerable. Such a toy version would be counterintuitive for most (high-school) students, since the students at the same time might be confronted with random variables in the sense of unprecise measurements in some applied discipline.

Philosophically speaking, such a toy version would miss a "principle of continuity" (in a sense which I have learned in long discussions with Michael Otte). We insist, that it is basic for a thinking in stochastical terms, that the events (in general) cannot be broken down to any kind of "elementary events" (which in the measure-theoretic reduction would correspond to the values of some independent variable $\omega \in \Omega$). There are school teachers who have tried to tell their students what elementary events are in the case of an infinite sequence of coin tosses. They all have failed; and we argue that there are fundamental philosophical reasons for this failure. Events are not like sets, which are equal if and only if they have the same elements;

the equality of events is more abstract and flexible. The partial analogy between sets and events is the philosophical trap which every teacher of stochastics should be aware of. (Of course the advanced analyst will be happy to learn that there exists a technical although highly non-constructive bridge to set theory. 1948 Loomis has proved that every σ-complete Boolean lattice can be represented by a set of equivalence classes of measurable sets.)

Here is an example of a nontrivial didactical decision. At some instance the system of axioms has to say under what conditions an infinite sequence of random variables X_1, X_2, ... may itself be considered as a random variable in the universe. In particular it has to be fixed axiomatically under what conditions an infinite product of admissible spaces (S_i, $\mathscr{B}_i$) can be considered as an admissible space. We think that an infinite sequence of coin tosses or a random path through a rooted tree ought to qualify as a random variable at a fairly early stage of stochastical education.

For a second course at university level I strongly recommend the full class of polish spaces. This opens a wide range of interesting questions and fruitful considerations which fortunately have nice intuitive answers. It turns out, that the theory of "sure" convergence of random variables is more or less a well reflected version of the material of convergence and continuity, which is commonly treated in elementary analysis. The main ingredients of sophistication are the following: we experience the distinction between sure events and mathematically true assertions. Secondly, all manipulations of the objects are bound to be of denumerable type. Thirdly the considerations about nullsets in measure theory get a clear profile. Finally the floor is set for the theory of weak convergence (which is a subject for advanced courses). It gains a nice intuitive background by the concept of sure convergence; stochastic convergence appears as the topological weakening of sure convergence.

This is not the place to say more about mathematical technique. A thorough scrutiny of the technical aspects of our axiomatization is in preparation. On Kersting's homepage you can find a manuscript documenting a first university course which puts random variables on center stage. More about tensions between stochastics and mathematics proper can be found in Dinges (2001).

As a conclusion I may summarize the main thrust of this article in honour of my dear friend Michael: Students on all levels ought to be encouraged to trust in the intuitive concept of a random variable. The variability of an observable may be due to chance; it is not necessarily caused by the specification of some "independent" variable.

ISMI, Institut für Stochastik & Mathematische Informatik, Fachbereich Mathematik, Universität Frankfurt/Main

REFERENCES

Dinges, H. (2001). Stochastisches und deterministisches Denken. *Allgemeines Statistisches Archiv* 85(2), 173 ff.

Moore, D. S. & Cobb, G. W. (2000). Statistics and Mathematics: Tension and Cooperation. *Mathematical Association of America Monthly* 107, 615-629.

Mumford, D. (2000). The Dawning of the Age of Stochasticity. In V. I. Arnold, M. Atiyah, P. D. Lax, & B. Mazur (Eds.), *Mathematics: Frontiers and Perspectives.* Providence, RI: AMS.

JOHANNES LENHARD

DEDUCTION, PERCEPTION, AND MODELING:

The Two Peirces on the Essence of Mathematics

Abstract. Charles Sanders Peirce, the celebrated philosopher of pragmatics and semiotics, viewed mathematics as the basic science. But, according to him – what is it?

In providing an answer, he gave reference to his father Benjamin Peirce, a leading Harvard mathematician. Charles quoted him with: *Mathematics is the science which draws necessary conclusions.* However, he went further than his father's position by asking what is necessary reasoning. His analysis led him from the clean world of pure reasoning to the more down-to-earth circumstances of perception and experimentation. Even deductive reasoning proceeds by using signs and their iconic qualities and is based on the perception and experimental manipulation of diagrams. Moreover, Peirce accomplished a pragmatic shift that was oriented toward mathematical practice and especially included the process of modeling as a mathematical key activity.

This Peircean standpoint will be explored in more detail, and it will be shown (so I hope) that it offers a perspective for a genetic philosophy with an impact on the didactics of mathematics.

Key words: deduction, diagrammatic reasoning, modeling, Peirce, perception.

1. INTRODUCTION

If one is reading, as currently, in a philosophical context, about "two Peirces", or even a greater number, one will probably expect to read about the different standpoints taken by Charles Sanders Peirce (1839 – 1914) in different articles or at different times. C. S. Peirce is prominent as the "most original and versatile of American philosophers and America's greatest logician" (Weiss 1934), and as the founder of philosophical pragmatism and semiotics. But, in equal measure, his works appear as complex and heterogeneous.

The title of the present article, however, refers to two different persons – Charles and his father, Benjamin Peirce. Whereas for today's philosophy, and especially the philosophy of mathematics, Charles is by far the more prominent of them, the situation was completely reversed a hundred years ago. Benjamin was the most famous scientist of the 19th century in the USA. He was a mathematician and he was the admired idol of his son Charles. It is in contest with his father's standpoints that Charles formed an important part of his philosophy of mathematics. At least this is my thesis and the topic of the present article.

C. S. Peirce was a passionate system-builder. He conceived of mathematics as the fundamental science in the system of the sciences (cf. his lectures on pragmatism). It is worth noting that Peirce did not choose logic, his home discipline, as the fundamental science. In the literature, this position is controversial. Grattan-

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 313 – 323.

Guinness (1997) or Hull (1994) vote in favour, whereas Fann (1970) or Murphey (1993) propose that Peirce has seen logic as foundational after all. I myself shall not contribute to this debate that traditionally dominates the view on Peirce's conception of mathematics. I shall ignore the relation between logic and mathematics and hope to gain a clear view on Peirce's perspective on mathematics.

By the way, this coincides with my personal career. I have become acquainted with pragmatism, or pragmaticism in Peirce's terminology, as a philosophical standpoint through Karl-Otto Apel in Frankfurt. Alas, he consciously ignored the mathematical texts of Peirce. Later, I appreciated a philosophical position playing a kind of *basso continuo* in Michael Otte's works: that the core of mathematics itself, closely examined, shows deep connections to philosophy.

Back to the issue in my article: What is the essence of mathematics according to Peirce? The common view is that he took over the definition from his father Benjamin. Indeed, Benjamin plays a crucial role as admired scientist. Born in 1809, he graduated at Harvard where he began an equally successful and long-lasting career as professor for mathematics and astronomy, lasting 47 years until his death in 1880. His leading positions as consultant for the government, in the Coast Survey, or the National Academy of Sciences is aptly summarized by Murphey: "His position was thus a commanding one in almost every field of physical science in America" (Murphey 1993, 11).

The young Charles Peirce seemed to be predestined for a bright career at Harvard and delivered his Harvard lectures on the philosophy of science in 1864 – 65 and the Lowell lectures 1866 – 67. But he always stood in the shadow of his all-powerful father. "At Harvard and on the Coast Survey, he was still Benjamin Peirce's son" (ibid., 19). Joseph Brent, in his highly readable biographical study, identifies the relation between Charles and his father as the key to the former's life and consequently to the failure of Charles' academic efforts. "Charles spent his life trying to surpass his father at his own subtle and demanding calling, the exploration of the abstract" (Brent 1998, 340). Whatever the biographical motivations may be, it is certain that, concerning the reflection about mathematics, the standpoint of Benjamin served as the starting point for Charles both personally and as regards content.

Charles repeatedly and affirmatively quoted the famous definition of Benjamin published in his late main work "Linear Associative Algebra" from 1870 and surely in familiar dialog for a longer time: *Mathematics is the science which draws necessary conclusions.*

This definition expresses two aspects: First, mathematics cannot be limited by referring only to certain objects. Stan Ulam expressed this in a similar vein about 100 years later:

> The mathematical *method*, as presently used, probably would not appear strange to the Greeks. However, the *objects* to which mathematical thought is devoted today have been vastly diversified and generalized. It is their proliferation that would perhaps appear so striking not only to the ancients but even to mathematicians of the last century. (Ulam 1986, 2)

The second aspect of Benjamin's definition, again in accordance with Ulam, is that the method is made the criterion of demarcation, namely, *necessary reasoning*.

It is uncontroversial that this definition played a central role for Charles (cf. Engel-Tiercelin 1993, 30/31; Grattan-Guinness 1997, 34; or Hull 1994, 274). But contrary to the suggestions of the mentioned literature, I shall claim that this is the beginning and not the end of Peirce's considerations about an adequate definition of mathematics. Admittedly, Benjamin's definition serves as an accepted basis for Charles, and therefore it can justly be called the family doctrine of the Peirces. Later on, Charles had modified, or better, transformed this statement in a twofold way. This transformation can be read as a critical discussion of his father's position.

In the following I shall interpret a text of Charles Peirce and quote extensively from it. This is worthwile, because this text is a self-contained statement published in the *Educational Review* of 1898 as "The Logic of Mathematics in Relation to Education". The determination of mathematics is the main theme, and the modifications of Benjamin's definition are clearly expressed. The first transformation is a *pragmatic turn*. It enlarges the conceptual framework and looks at mathematics as an activity oriented toward applications, containing necessary reasoning only as a part. The main question shifts from: *what is the essence of mathematics*, to *what is the business of the mathematician?* The second trans-formation is a semiotic specification. Charles asks further *what is necessary reason-ing like?* And his answer consists in the semiotic proposal that each deduction contains elements of perception in an essential way, or proceeds diagrammatically, as he baptized it.

2. OF MATHEMATICS IN GENERAL – C. S. PEIRCE'S PRAGMATIC AND SEMIOTIC VIEWPOINT

The subject of this section will be the interpretation of Peirce's 1898 article "The Logic of Mathematics in Relation to Education," bearing the subtitle "Of Mathematics in General." It is contained in the *Collected Papers* (CP), covering paragraphs 553 to 562 of the third volume, according to the usual way of citation: CP 3.552 – 3.562. In this short text, Peirce reasons about the essence of mathematics from his philosophical viewpoint. One can find, therefore, characteristic features of the Peircean philosophy in a nutshell.

Commonly, one finds cited only the lines in which Peirce quotes his father's definition. The citation should support the identity of the two viewpoints. But this is not the whole truth. Therefore, I shall quote the Peircean text extensively to show how Charles develops his own position, modifying Benjamin's definition in subtle but essential aspects.

While searching for a definition of mathematics, Peirce is concerned at first with "the definition of mathematics as the science of quantity." He judges a Greek origin as rather implausible, because Aristotle reasoned "that mathematics ought not to be defined by the things which it studies but by its peculiar mode and degree of abstractness." (3.554) This aspect, that mathematics is independent from the objects it investigates, will build the backbone of the later definition by Benjamin Peirce. The origin of the definition as the science of quantity may be somewhere, this does not affect that "the definition of mathematics as the science of quantity suited well enough such mathematics as existed in the seventeenth and eighteenth centuries." (3.555) But it did not suit any longer, because obviously mathematics began to cover a broader field.

In the next paragraph, Peirce discusses Kant and his conception of mathematics given in the *Critique of Pure Reason*. For Kant, as for Peirce, mathematics played a fundamental role for epistemology.

> Kant, in the Critique of Pure Reason (Methodology, chapter I, section 1), distinctly rejects the definition of mathematics as the science of quantity. What really distinguishes mathematics, according to him, is not the subject of which it treats, but its method, which consists in studying constructions, or diagrams. That such is its method is unquestionably correct; for, even in algebra, the great purpose which the symbolism subserves is to bring a skeleton representation of the relations concerned in the problem before the mind's eye in a schematic shape, which can be studied much as a geometrical figure is studied. (3.556)

Even this short passage shows clearly how Peirce embeds his semiotic thesis about diagrammatic reasoning in the tradition of Kant. What Kant described as construction was interpreted by Peirce as constructing an observable picture or diagram that allows for further empirical analysis.

Peirce defends Kant against a cursory reading, as he ascribes to Hamilton and De Morgan, who had characterized the Kantian definition of mathematics as "science of pure time and space."

> Not only do mathematicians study hypotheses which, both in truth and according to the Kantian epistemology, no otherwise relate to time and space than do all hypotheses whatsoever, but we now all clearly see, since the non-Euclidean geometry has become familiar to us, that there is a real science of space and a real science of time, and that these sciences are positive and experiential – branches of physics, and so not mathematical except in the sense in which thermotics and electricity are mathematical; that is, as calling in the aid of mathematics. (3.557)

What is the relation between mathematics and sciences that call in the aid of mathematics? The answer to this question is pivotal for Peirce's philosophy of mathematics: Defining mathematics requires it to be embedded into the process of scientific research. Only by looking at the whole process does it become possible to acknowledge those aspects of genesis and evolution that were so close to Peirce's heart.

Now come the decisive paragraphs, introduced by the position of his father who formulated very clearly that mathematics cannot be defined through its objects, but has to be defined "subjectively."

> Of late decades philosophical mathematicians have come to a pretty just understanding of the nature of their own pursuit. I do not know that anybody struck the true note before Benjamin Peirce, who, in 1870, declared mathematics to be "the science which draws necessary conclusions," adding that it must be defined "subjectively" and not "objectively." (3.558)

Charles quotes his father affirmatively, and therefore one can speak of this definition as the family doctrine. Nevertheless, he modifies it in the following paragraphs and thereby works out the original Peircean pragmatic and semiotic shifts.

The definition in the *Encyclopedia Brittannica*, according to which "the essence of mathematics lies in its making pure hypotheses, and in the character of the hypotheses which it makes" (3.558) is in line with this family doctrine.

> What the mathematicians mean by a "hypothesis" is a proposition imagined to be strictly true of an ideal state of things. In this sense, it is only about hypotheses that necessary reasoning has any application; ... (3.558)

This unspectacular statement contains the fundamental insight that the applicability of mathematics depends on the construction of *ideal states* – or expressed in a more modern fashion: it depends on the construction of model worlds. Mathematics makes its assertions only by the use of models and only about models.

> Hence to say that mathematics busies itself in drawing necessary conclusions, and to say that it busies itself with hypotheses, are two statements which the logician perceives come to the same thing. (3.558)

One could exchange Peirces "hypotheses" for "models." What is the subject matter of "busies itself?" At first sight, it seems as if mathematicians would act solely in the world of models, because it is only there that they can do what characterizes them, namely, to reason necessarily.

This self-sufficient attitude of mathematics is closely related to so-called "if-thenism," and seems to result from the Peircean family doctrine. Exactly at this point, Peirce takes a step that shapes his philosophy of mathematics. He embeds the necessary reasoning, as a part of mathematical activity, into a more global practice of mathematics. The question is not about *any* hypotheses, but about ones that are judged as adequate for certain reasons, that is, for criteria of application. In other words, according to Peirce, the steps of modeling (that Peirce called "abductive"), of necessary reasoning within a model, and of testing the results inductively by application to the world of phenomena form one and the same process. This process, in turn, refers not only to a theoretical and methodological practice, but also to a social one. Thus, Charles Peirce performs a pragmatic shift in the definition of mathematics. The question of identifying the "essence of mathematics" is transformed into the question of the "business of the mathematician." In Peirce's own words:

> A simple way of arriving at a true conception of the mathematician's business is to consider what service it is which he is called in to render in the course of any scientific or other inquiry. Mathematics has always been more or less a trade. An engineer, or a business company (say, an insurance company), or a buyer (say, of land), or a physicist, finds it suits his purpose to ascertain what the necessary consequences of possible facts would be; but the facts are so complicated that he cannot deal with them in his usual way. He calls upon a mathematician and states the question. Now the mathematician does not conceive it to be any part of his duty to verify the facts stated. He accepts them absolutely without question. He does not in the least care whether they are correct or not. (3.559)

Here, Peirce poses slightly too much, because the process of modeling is a highly dialogic undertaking and it affords a lot of mediation. Modeling is dependent on aspects of correctness, or better, of adequacy.

> He (the mathematician) finds, however, in almost every case that the statement has one inconvenience, and in many cases that it has a second. The first inconvenience is that, though the statement may not at first sound very complicated, yet, when it is accurately

> analyzed, it is found to imply so intricate a condition of things that it far surpasses the power of the mathematician to say with exactitude what its consequence would be. At the same time, it frequently happens that the facts, as stated, are insufficient to answer the question that is put. (3.559)

That is the lesson of the naive model builder: In the process of application, it is neither obvious which facts are relevant, nor how these facts could be molded into a model that is able to allow for a mathematical treatment. There is no unique relation between the field of concrete problems of application and mathematical models.

> Accordingly, the first business of the mathematician, often a most difficult task, is to frame another simpler but quite fictitious problem (supplemented, perhaps, by some supposition), which shall be within his powers, while at the same time it is sufficiently like the problem set before him to answer, well or ill, as a substitute for it. (3.559)

At this point, the process of modeling is getting off the ground. It is worth noting that "to frame another problem" has nothing to do with necessary reasoning, rather it is a question of judgment. Nevertheless, it is "the first business of the mathematician." Peirce is assigning the problems of adequate modeling to mathematics itself! Thus he is enlarging the family doctrine considerably. Although mathematics is still not defined by certain objects, the difficulties in the handling of objects are still very much involved. In the quotation, Peirce has named two criteria:

1. The question to the model should be practically treatable, "shall be within his powers," and
2. the model should be sufficiently adequate, in the sense of applicability, "sufficiently like the problem set before him."

The characteristic tension in applied mathematics is caused by the condition to fulfill (i) and (ii) "at the same time." Both aspects stand in a complementary relation – the more one makes it easier with the one aspect, the more it gets difficult with the other. Normally, neither very simple nor very complex models lead to a solution, because they either have little significance or are hardly possible to analyze.

The (always only preliminarily accepted) outcome of the process of modeling is preparing the ground for necessary reasoning, one can say. And the latter cannot claim the rank of a definition of mathematics.

> This substituted problem differs also from that which was first set before the mathematician in another respect: namely, that it is highly abstract. All features that have no bearing upon the relations of the premisses to the conclusion are effaced and obliterated. The skeletonization or diagrammatization of the problem serves more purposes than one; but its principal purpose is to strip the significant relations of all disguise. (3.559)

After the pragmatic transformation from *essence* to *business* described above, the second philosophically meaningful step beyond the family doctrine consists in a specifying request. What is necessary reasoning like?

> Kant is entirely right in saying that, in drawing those consequences, the mathematician uses what, in geometry, is called a "construction," or in general a diagram, or visual array of characters or lines. Such a construction is formed according to a precept furnished by the hypothesis. Being formed, the construction is submitted to the scrutiny of observation, and new relations are discovered among its parts, not stated in the

> precept by which it was formed, and are found, by a little mental experimentation, to be such that they will always be present in such a construction. Thus, the necessary reasoning of mathematics is performed by means of observation and experiment, and its necessary character is due simply to the circumstance that the subject of this observation and experiment is a diagram of our own creation, the conditions of whose being we know all about. (3.560)

This quote gives the principal elements of diagrammatic reasoning, always including "observation and experiment." Peirce's semiotic answer consists in analyzing necessary reasoning as diagrammatic reasoning. This Peircean conception is enjoying growing attention in different areas as semiotics, didactics of mathematics, or artificial intelligence (cf. Stjernfelt 2000; Hoffmann 2002; Chandrasekaran 1995). The pictorial approach seems to me to be especially promising, because the philosophy of mathematics usually suffers from a strong linguistic orientation, which it has inherited from the common analytical philosophy of science (cf. for an exception, e. g., Giere 1999).

At the end of the foregoing quotation, Peirce is overshooting slightly in claiming a complete knowledge about the self-created diagrams. In this passage, he comes very near to the *verum est factum* theory of Cusanus. Therefore, he seems to miss his own point: Particularly the observed relations in a diagram can give rise to further questions. On, for example, the assumptions that underlie the construction, and whether some of them can be weakened without destroying the observed relations. This results in an analysis of the further attributes of a created diagram or a built model – attributes that were not known from the first and which are not under the command of the diagram's creator. That is, as I see it, the point of experimenting with diagrams. Michael Otte (2002) gives a nice example of a diagrammatic proof analysis, which starts from certain geometrical points in a triangle, observes their invariance under certain transformations (based on dynamical geometry software), and leads eventually to the theorem of Desargues. Peirce refers once more to Kant – my last quotation from the text:

> But Kant, owing to the slight development which formal logic had received in his time, and especially owing to his total ignorance of the logic of relatives, which throws a brilliant light upon the whole of logic, fell into error in supposing that mathematical and philosophical necessary reasoning are distinguished by the circumstance that the former uses constructions. This is not true. All necessary reasoning whatsoever proceeds by constructions; and the only difference between mathematical and philosophical necessary deductions is that the latter are so excessively simple that the construction attracts no attention and is overlooked. The construction exists in the simplest syllogism in Barbara. Why do the logicians like to state a syllogism by writing the major premiss on one line and the minor below it, with letters substituted for the subject and predicates? It is merely because the reasoner has to notice that relation between the parts of those premisses which such a diagram brings into prominence. If the reasoner makes use of syllogistic in drawing his conclusion, he has such a diagram or construction in his mind's eye, and observes the result of eliminating the middle term. (3.560)

Peirce is arguing explicitly in favour of the essential similarity of all necessary reasoning, that is, for the thesis that all such reasoning proceeds diagrammatically. In this context, it provides a good argument to interpret the compelling force of a syllogism as an effect of observation.

By the way, this passage indicates a point about the Kantian philosophy of mathematics. Peirce criticizes Kant indulgently for his erroneous claim that philosophy and mathematics would reason in essentially different manners (namely conceptually vs. intuitively). The missing acquaintance with the logic of relations, so Peirce, has caused Kant's mistake. For Peirce, in contrast, it is just the logic of relations that provides a sufficiently general framework to perceive the structural similarity of all necessary reasoning. All such reasoning involves elements of observation. As one of the founders of the logic of relations, Peirce surely has highlighted this instance proudly.

Interestingly, one can find a very similar critique of Kant that points to the very same mistake, but reaches the opposite conclusion. Bertrand Russell (1956), or, in the same vein, Michael Friedman (1992), blame Kant for his out-dated conception of logic, that was based on the Aritotelian subject-predicate logic, which, in turn, was fundamentally broadened by relational logic. So far, their critique agrees with Peirce's. But they draw the opposite conclusion: All elements of construction, intuition, and perception are judged to be superfluous, because they would be introduced only as a crook for Kant's insufficient conception of logic (cf., for a critical discussion, Lenhard 2004). Maybe this illustrates that Peirce saw, as mentioned already at the beginning, mathematics as a fundamental science and did not argue for logicism in the sense of Russell, whose goal was to deduce mathematics from logic.

Let us leave the interpretation of the Peircean text here and sum up the results of the reading.

3. CONCLUSION

The Peircean family doctrine, that is, the statement of Benjamin Peirce that mathematics is the science that draws necessary conclusions, was adopted also by Charles Peirce. But, as we have seen, Charles modified this doctrine in a twofold manner. First, he broadened it pragmatically and second he made it more specific semiotically. Let us review the two modifications briefly.

1. Peirce has broadened the task of determining mathematics from the "essence of mathematics" to the "business of the mathematician." This presents a *pragmatic shift*, insofar as necessary reasoning is embedded in a process of research, that is, a social practice, which especially includes the systematically important stage of modeling.

   This transformation is in good agreement with the general feature of Peirce's pragmatic philosophy, as expressed by Karl-Otto Apel, namely to see the *logic of science* as a kind of "logical Socialism (the ideal of a practising community of inquirers)." And to act on the assumption "that the world cannot be known or explained merely by its previously fixed, lawful structure, but must rather continue to be developed as a historical, social world of institutions and habits for which we must assume responsibility ... (Apel 1981, 193)." That this pragmatic turn affects Peirce's philosophy of mathematics is broadly underestimated in the literature (cf. the above-mentioned references to Engel-Tiercelin, Grattan-Guiness and Hull).

One could presume that C. S. Peirce's activity as an applied mathematician for the *Coast Survey* contributed to this picture. This is oriented toward the practice of mathematical research and not the particular subspecies of so-called "pure" mathematics. Finally, Peirce takes up a position, very likeable for me, according to which it is a philosophical challenge to investigate mathematics on the basis of a close examination and acquaintance of its course of action.

2. The second modification consisted in a *semiotic shift*, which takes necessary reasoning itself as the subject of investigation and localizes elements of observation and experiment in any such reasoning.

   The principal direction of the family doctrine is maintained; that is, no specific domain of objects is ascribed to mathematics. But this should not hide the fact that diagrams do consist in concrete individual objects. And, furthermore, mathematical deduction explores, using the framework of a model, the properties of those objects. Thus, the semiotic shift goes hand in hand with a strong relation to objects. It can count as one of Peirce's main points that mathematics proceeds by transforming abstract concepts into concrete objects of investigation, a process that he has called "hypostatic abstraction."

Peirce strived toward a pragmatic, or pragmatistic, and semiotic viewpoint on mathematics. Insofar, the interpreted passages are "typically Peirce." Another typical feature is that Peirce was seeking for a methodological answer to the question on the "essence" of mathematics.

Is this approach getting rid of all philosophical problems? Admittedly, this is a rhetoric question – it does not. Here I should like to point to the already mentioned problems concerning the adequacy of mathematical models that turn up in the context of application. Modeling is an activity in two worlds, so to speak, and Peirce has seen it as a fundamental problem to give a philosophically valid argument for a trustworthy relation between the world of real phenomena and that of an "ideal state of things." The difficult task is not to mix them up and, at the same time, to mediate them. Following Leibniz, Peirce has connected this problem with a "principle of continuity," and he has worked out a metaphysical hypothesis, the so-called *synechism*, a kind of doctrine of connection.

For me, the main achievement of Peirce's philosophical account of mathematics does not lie in the classification of mathematics in the architectonic system of sciences (that was not mentioned at all in our reading). Rather it consists in the convincing proof that questions concerning mathematics inevitably lead to questions of the broadest philosophical significance. This seems to me to be also the essential viewpoint in Michael Otte's considerations about the Peircean principle of continuity that are part of his "Mathematik und Verallgemeinerung – Peirces semiotisch-pragmatische Sicht" (1997).

Finally, I shall hint at the significance of these philosophical reflections for the didactics of mathematics. First, Peirce's article contains the "relation to education" in its title, and one can ask what this relation is. Actually, Peirce does not aim at didactical considerations, at least not in the examined article (cf., about Peirce's didactics, Radu 2003).

Nevertheless, the text contains a didactic message. I read it as the appeal to *ground didactics philosophically*, that is as the demand to base didactics of mathematics on a philosophically adequate picture of mathematics. (Perhaps the collaboration with Michael Otte has made me hypersensitive in this respect). Today, the dominant picture views the essence of mathematics as being realized in so-called "pure" mathematics, which is thought of as marking the application-far pole of all sciences. This picture is in obvious conflict with Peirce's view. And, in addition, this culturally established picture of mathematics is, at least partly, to be blamed for the recently diagnosed dreadful state of mathematics education. Thus there lies a common didactical *and* philosophical challenge. Currently, this challenge seems to be taken up by different approaches as exemplified in Corfield (2003), Lenhard and Otte (2003), or van Kerkhove (2003) – and some contributions to the present volume.

Institut für Didaktik der Mathematik und Institut für Wissenschafts- und Technikforschung, Universität Bielefeld

REFERENCES

Apel, K.-O. (1981). *Charles S. Peirce: From Pragmatism to Pragmaticism.* Amherst: The University of Massachusetts Press.

Brent, J. (1998). *Charles Sanders Peirce. A Life.* Bloomington and Indianapolis: Indiana University Press.

Chandrasekaran, B. e. a. (Ed.) (1995). *Diagrammatic reasoning: cognitive and computational perspectives.* Menlo Park, Calif.: AAAI Press.

Corfield, D. (2003). *Towards a Philosophy of Real Mathematics.* Cambridge: Cambridge University Press.

Engel-Tiercelin, C. (1993). Peirce's Realistic Approach to Mathematics: Or, Can One Be a Realist without Being a Platonist? In E. C. Moore. (Ed.), *Charles S. Peirce and the Philosophy of Science. Papers from the Harvard Sesquicentennial Congress.* Tuscaloosa and London: The University of Alabama Press, 30-48.

Fann, K. T. (1970). *Peirce's Theory of Abduction.* The Hague: Nijhoff.

Friedman, M. (1992). *Kant and the Exact Sciences.* Cambridge, Mass.: Harvard Univ. Press.

Giere, R. N. (1999). *Science without laws.* Chicago: The Univ. of Chicago Pr.

Grattan-Guinness, I. (1997). Peirce between Logic and Mathematics. In N. Houser, D. D. Roberts, & J. van Evra (Eds.), *Studies in the Logic of Charles Sanders Peirce.* Bloomington and Indianapolis, Indiana University Press, 23-42.

Hoffmann, M. H. G. (2002). How to get it. Diagrammatic reasoning as a tool of knowledge development and its pragmatic dimension. *Foundations of Science,* in print.

Hull, K. (1994). Why Hanker After Logic? Mathematical Imagination, Creativity and Perception in Peirce's Systematic Philosophy. *Transactions of the Charles S. Peirce Society* 30, 271-295.

Lenhard, J. (2004). Kants Philosophie der Mathematik und die umstrittene Rolle der Anschauung. [Kant's Philosophy of Mathematics and the controversial role of intuition.] *Kantstudien,* to appear.

Lenhard, J. and Otte, M. (2003). Grenzen der Mathematisierung – von der grundlegenden Bedeutung der Anwendungen. [Limits of Mathematization – on the Fundamental Role of Applications.] Manuscript.

Murphey, M. G. (1993). *The Development of Peirce's Philosophy.* Indianapolis, Cambridge: Hackett Publish Comp.

Otte, M. (1997). Mathematik und Verallgemeinerung – Peirces semiotisch-pragmatische Sicht. [Mathematics and Generalization – Peirce's Semiotic-Pragmatic View.] *Philosophia naturalis* 34, 175-222.

Otte, M. (2002). Proof-Analysis and the Development of Geometrical Thought. *Representations and Mathematics Visualization.* F. Hitt. Mexico: Cinvestav-IPN.

Peirce, C. S. (CP). *Collected Papers of Charles Sanders Peirce*. Cambridge, Mass.: Harvard UP.
Radu, M. (2003). Peirces Didaktik der Arithmetik: Möglichkeiten ihrer semiotischen Grundlegung. [Peirce's Didactic of Arithmetic: Possible Foundations in Semiotic.] *Mathematik verstehen – Semiotische Perspektiven*. M. H. G. Hoffmann. Hildesheim, Franzbecker: 160-194.
Stjernfelt, F. (2000). Diagrams as Centerpiece of a Peircean Epistemology. *Transaction C.S.P.Society* 36(3): 357-392.
Ulam, S. (1986). The Applicability of Mathematics. *Science, computers and people*. G.-C. R. Mark C. Reynolds. Boston, Birkhäuser: 1-8.
Van Kerkhove, B. (2003). Mathematical Naturalism: Origins, Guises, and Prospects. *Foundations of Science*. To appear.
Weiss, P. (1934). C. S. Peirce. *Dictionary of American Biography*, 14 , 398-403.

C. ULISES MOULINES

MODELS OF DATA, THEORETICAL MODELS, AND ONTOLOGY

A Structuralist Perspective

Abstract. A general scheme for dealing with ontological issues form a "Scientistic point of view" is proposed. "Ontological commitments" (in Quine's sense) should always be examined with respect to a well-established scientific theory. By means of a schematic example of "reconstruction" of a piece of experience within a theoretical frame, it is shown what the essential steps in the process of constructing of a scientific ontology are. This steps involve, first, the construction of a "model of data" for a given "experiential situation", second, the selection of a model of a mathematized theory, an third, the subsumption of the data model under the selected mathematical model. A further schematic example provides the clues for answering the question of ontological reduction between different experiential domains. A final word is said about what it would mean to have a really unified universal ontology.

Key words: ontological commitment, data model, theoretical model, subsumption, echelon set, ontological reduction.

In the traditional wording, ontology is the discipline of Being in general. In a less bombastic, but eventually equivalent way of speaking, Quine has characterized it as the discipline of what there is (Quine 1953). And all "what there is," also following Quine, is the values of bound variables. Though Quine's slogan is in need of some minor qualifications and revisions, I still think it is a good starting point for doing ontology.[1] The question, however, is to determine where these bound variables are to be found. My answer (which I think also follows the spirit though perhaps not the letter of Quine's slogan) is that they are to be found in *scientific* texts, or, more precisely, in the formulations of scientific theories in standard scientific texts. I feel no inhibition in pleading for a "scientistic" ontology: If you want to find out what there is in the world around you, ask science! I shall not argue at length in favor of this kind of "scientism," I would just like to point out that it responds to an "economic" strategy founded on historical induction: Whenever there has been a dissent on "what there is" between science on the one hand and common sense, metaphysics, religion, or whatever on the other, in the long run it has always been science that has won the battle.[2] Though I think there are very good reasons why this is so, I cannot dwell upon this point here.

Therefore, having taken this "scientistic" decision in matters ontological, our next task is to investigate the way scientists handle their "ontological commitments" (to employ another of Quine's famous phrases) within the theoretical frame of their respective disciplines. Now, the first thing we notice is that the starting point for any

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 325 – 333.

scientific investigation is a more or less loosely determined piece of (ultimately sensorial) experience.[3] Nonetheless, unless we are hard-nosed ontological phenomenalists (which I am not), we shall not assume that the elements of a particular piece of experience are actually *what there is*. Let us rather say that scientific research starts with a particular *experiential situation (ES),* and the final goal is to find out what there is behind *ES* – to find out the "Hidden Being Behind Appearances." Scientists (or whoever) may describe *ES* by means of particular expressions of ordinary language, like "wet," "hot," "sweet," "brown," and so forth. But a "scientific ontology" will not be one in which it is assumed that these predicates apply to what really is. Our starting point is necessarily an *ES* we describe in ordinary language, but our goal is to get "behind" *ES* and describe it within the conceptual frame of a scientific theory.

A first obstacle to attain this goal, however, comes from the fact that the identity criteria for any *ES* considered are too fuzzy and uncertain for our purposes. The conceptual demarcations of experiential objects characteristic of ordinary life have to be further refined and modified for scientific purposes. To do this, scientists perform a series of actions and interactions, mainly consisting in linguistic communication with their peers plus the systematic observation and/or manipulation of medium-sized objects. This leads, in a first step, to the constitution of what we may call an *"operational base," OB*, for *ES* ; in this way, the original *ES* becomes transformed and codified into an *intersubjectively controlled experiential situation (ICES)*. The same or another *OB* may serve to determine other *ICES*s.

It is important to be aware of the fact that a successful codification process leading from an original *ES* into an *ICES* is never the outcome of the heroic action of a single individual but rather that of intersubjective communication. Or, to put it more cautiously, the only *ICESs* that science takes seriously for ontological purposes are those constructed intersubjectively. The construction of *ICES*s always takes place within a collective entity, a "group of partners" (*GP*) standing in regular interaction. At least since Thomas S. Kuhn, we know that these entities are the real subjects of science. In this sense, scientific ontology depends on pragmatics.

Of course, this is only the first step toward a full-fledged scientific ontology. What comes next? To discuss this question, I propose to lay out a couple of schematic examples instead of a general argument.

FIRST EXAMPLE

While lying on his terrace, Johnny (or Nebuchadnezzar) faces an experiential situation, let us call it "ES_p." This intrigues him, and may be described (without any ontological commitment) as follows: "Experience of the sky on a clear night; slowly moving sparkling points at various positions." In order to interpret ES_p correctly, Johnny-Nebuchadnezzar meets a group of partners, $GP°$, all of them equally intrigued by the night sky, and they agree to proceed in the following way:

First, they agree to undertake systematic observations on many following nights.

Second, they agree to determine lapses of time by means of a device (a medium-sized object) called a "clock." (We need not imagine here a quite sophisticated, i. e., "theoretized" apparatus; it would be enough to have a sandglass, on which GP° makes some marks: The time elapsed is determined by the number of marks covered by the falling sand.)

Third, on a long series of following nights, always beginning at the same time (i. e., starting from the same mark on the sandglass) and after equal lapses (i. e., after the same number of covered marks), GP° focuses its special attention on those sparkling points that move irregularly.

Fourth, after some hesitations, GP° decides to devote itself *only* to the irregularly moving points and to give them a generic name: "planets." They also give proper names to the individual points: "Mercury," "Venus," and so forth.

At this point, we can say that GP° has transformed the original ES_p into a clearly demarcated $ICES_p$.

We now go on to the next phase of the scientific enterprise. GP° notices that a thorough investigation of the issue of the night sky requires more than mere observations and denominations. One has to "fix them on paper." GP° decides to **represent** $ICES_p$ in the following way: They take squared paper, agree on marking a "center of coordinates" on the paper representing a particular sparkling point (e. g., the so-called "North star"), and they determine the relative positions of the planets with respect to the North star by means of successive marks on the paper within regular lapses of time. On a great number of sheets, they obtain in this manner a great number of marked points.

We may call this representation an "**$ICES_p$-corresponding data model**," or just DM_p. The representation has been established according to a set of conventions or "axioms:"

(A_1) "Point p_1 on the paper, when the sand covers mark m, represents Mercury."
(A_2) "Point p_2 on the paper, when the sand covers mark m, represents Venus."

.
.
.

(A_m) "Point p_m on the paper, when the sand covers mark m', represents Mercury."

.
.
.

(A_n) "The distance from p_1 to the center, when the sand covers mark m, is r_1."
(A_{n+1}) "The distance from p_2 to the center, when the sand covers mark m, is r_2."

.
.
.

After some quarrels and reconciliations within GP° (what sociologists of science call "negotiations"), GP° agrees to **accept** "axioms" (A_1), (A_2), ..., (A_m),..., (A_n), (A_{n+1}),.... This process of acceptance is followed partly by convention (e. g., for (A_1), (A_2),...) and partly through procedures normalized and admitted by GP° (related, e. g., to the way the number of boxes on the squared paper separating each p_i from the center of coordinates for establishing (A_n), (A_{n+1}), ..., is to be counted). As soon as the process of acceptance comes to an end, GP° declares the statements (A_1), (A_2),...,(A_{n+1}),... to be **true**.

Consider now the following structure:

$$DM_p =: < \{p_1,\ldots,p_5\}, \{\ldots m_i,\ldots\}, d >,$$

in which d is a diadic function consisting of the triples

$$< p_i, m_i, r_i > , (\text{with } r_i \in \mathbf{Q}),$$

that are the values of those parameters correlated in each statement (A_{n+j}), with $j \geq 0$.

The construction process just illustrated leads to the conclusion that DM_p is a kind of structure that is a "model" of the axioms (A_1), ..., (A_n), ... in a sense similar, though *not identical*, to the standard Tarskian notion of "model of a formalized theory." It is similar to the standard model notion in the sense that the triples $<p_i$, m_i, $r_i>$ satisfy (in the strong Tarskian sense of "satisfying)" the formulae (A_{n+j}) for $j \geq 0$. However, it is not quite a Tarskian model, because we cannot really speak of "satisfaction" with respect to the formulae (A_k) for $k < n$: What we have here instead are operational presuppositions to settle the "universe of discourse" of DM_p. Taking into account both aspects (the analogy as well as the nonidentity with the formal model notion), it seems to be justified to apply to the structure DM_p the special denomination "data model".

Up to this point, we still have not talked about ontology. We still have not said what there is. Neither the sparkling points in the night sky, nor the marks on the sandglass, nor the marked points on the squared paper; none of these elements of $GP°$'s experiential space are what really is – at least from the point of view of a scientific ontology. For, up to this point, $GP°$ has not made any **serious ontological commitment.** According to the "scientism" in matters ontological we have adopted from the beginning, in order to come to such a commitment, $GP°$ has got to work within the frame of at least one well-established scientific **theory**. This brings us to a qualitatively new phase in our enterprise.

Let us assume somebody ($GP°$ itself or some other group of partners) has initially elaborated a specific theory, call it "$\boldsymbol{T}_p$," for the time being only as a purely mathematical formalism. According to the structuralist standpoint I adopt here, the identity criterion for it is its proper model class, $M[\boldsymbol{T}_p]$. Let us suppose the elements of this class have the form:

$$x \in M[\boldsymbol{T}_p] \rightarrow x = <D, I, \boldsymbol{IR}, R_1, \ldots, R_m, f_1, \ldots, f_n>,$$

so that the components of any x are characterized as follows :

(*1) D is a finite, nonempty set (whose elements $\boldsymbol{T}_p$'s inventor just calls "dots").

(*2) I is isomorphic to an interval of $\boldsymbol{IR}$ (whose elements are called "instants").

(*3) R_i and f_j are relations and functions defined over D and/or I and/or $\boldsymbol{IR}$.

Further, any $x \in M[\boldsymbol{T}_p]$ satisfies some "proper axioms" or "laws," that is, formulae relating D, I, R_i, and f_j with each other. (For the purposes of illustration, we can imagine they are the formulae known as "the laws of Kepler," "the laws of Newton," or something of the sort.)

Up to this point, we are still outside ontology. All this is pure mathematics (and I have already said that we just make the simplifying assumption that mathematics has nothing to do with what **there really is**). For the time being, to be a "dot" just means being an element of any set D that is, in turn, the first component of a given structure x that is, again in turn, an element of the class $M[T_p]$ characterized in purely formal terms. The same goes for the term "instant."

Now, suppose $GP°$ is able to give a mathematical proof for the following claim:

There is at least a concrete $u° \in M[T_p]$ such that:

DM_p* is a substructure of *$u°$ (abbreviated as "$DM_p \S u°$").

The concept of *substructure* relevant here may be explicated in following terms:

$$\text{Let } u° = <D°, I°, \boldsymbol{IR}, R°_1,\ldots,R°_m, \ldots, f°_1,\ldots f°_n>,$$

in which, for some $f°_i$, we change this symbol into $s°$ and characterize it as follows:

$$s°: D° \text{ X } I° \rightarrow \boldsymbol{IR}^2.$$

Then, formula "$DM_p \S u°$" means that following conditions are fulfilled:

(1) $\{p_1, \ldots, p_5\} \subseteq D°$;

(2) $\{\ldots,m_i,\ldots\} \subseteq I°$;

(3) $s°/\{p_1, \ldots, p_5\} \text{ X } \{\ldots,m_i,\ldots\} \text{ X } \boldsymbol{IR}^2 = d.$

From the claim "$DM_p \S u°$" (which has been mathematically proven), $GP°$ now concludes:

The experiential situation *$ICES_p$* can be subsumed under theory *T_p*.

Clearly, this conclusion is not a logico-mathematically valid inference; it corresponds more to what $GP°$ understands under "subsumption of experience under a theory" or, to put it more crudely, to pointing out that theory T_p "works well" with respect to $ICES_p$.

At this point, finally, $GP°$ (as well as we ourselves, as philosophers of science analyzing scientific behavior) are allowed for the first time to engage in ontological commitments and to declare

Planets are really dots.

Or, to put it somewhat differently "What there is in the night sky *are* dots;" or still, in the traditional jargon: "The **Being** hidden behind the nightly appearances consists of dots."

My claim is that this schematic and admittedly oversimplified example is *in principle* paradigmatic for the way ontological issues are solved from a scientific point of view. I say "in principle," because the following complication may arise.

SECOND EXAMPLE

Let us suppose we are confronted with another intersubjectively controlled, experiential situation *ICES* of the following kind. $GP°$ observes a number of more or less round, medium-sized rigid bodies, which any member of $GP°$ can see and touch (and smell and taste if he/she wishes). *GP* agrees thoroughly to polish these bodies (according to some standardized procedures), to put them on an equally well polished table, and to push the round bodies so that they move in straight lines, rotate, and collide with each other. $GP°$ systematically observes the changes of direction and rotation of the round bodies, and **represents** all this in a way analogous to, though not identical with, the case of the planets (e. g., the objects are no longer represented by points but by volumes, and the representation of the motion is not two- but three-dimensional, etc.). Suppose that $GP°$ comes in this way to the construction of another "data model" DM_r. Suppose, further, that somebody invents another theory, call it "$\boldsymbol{T_r}$," essentially different from $\boldsymbol{T_p}$, whose models have the structure $x = \langle C, I, \boldsymbol{IR}, \ldots \rangle$, thereby satisfying laws different from those of $\boldsymbol{T_p}$ (say, the principles of the conservation of kinetic energy and of angular momentum). *C*'s elements are now called "chunks."

Assume, further, that, by going through a formal argument analogous to the case of the planets with respect to the theory $\boldsymbol{T_p}$ of dots, $GP°$ now comes to the inference that, for a given $v° \in M[T_r]$, the claim is valid:

$$\boldsymbol{DM_r \ \S \ v°} \, ,$$

and that, therefore, the conclusion is warranted

$$\boldsymbol{ICES_r \text{ can be subsumed under } T_r.}$$

In this case, we may go on to the ontological way of speaking and declare that **chunks really exist**, or that "the hidden Being behind the rotating and colliding round bodies on the table consists of chunks."

Because it is not the case that $ICES_r = ICES_p$, nor that $DM_r = DM_p$, and because the laws of $\boldsymbol{T_p}$ and $\boldsymbol{T_r}$ are different nonequivalent formulae, there is no reason to assume that chunks have anything to do with dots. Even if we were to assume that our experiential situations are restricted to $ICES_p$ and $ICES_r$, we would be obliged bitterly to acknowledge that there is no unified ontological constitution of reality, since Being is sometimes being a dot and sometimes being a chunk.

Suppose, however, that *GP* notices the following fact: For any model *v* of $\boldsymbol{T_r}$ that is ontologically relevant, in other words, for any $v \in M[T_r]$ really subsuming a given $ICES_r$, one can proceed according to the following steps:

(a) One reinterprets v's basic domain as being not a simple set, but rather a set of sets. Instead of having $C = \{c_1, ..., c_n\}$, we would have $C = \{\{d_1, ... d_k\}, \{d'_1, ... d'_l\}, ..., \{d''_1, ..., d''_m\}, ...\}$.

(b) One takes the great union $D = \cup C = \{d_1, ..., d_k, d'_1, ..., d'_l, d''_1, ..., d''_m\}$ as the basic domain of a structure u that, comes out as a model of $\boldsymbol{T}_p$, that is, $u \in M[\boldsymbol{T}_p]$.

(c) It is possible to prove mathematically that if you add some special conditions to $\boldsymbol{T}_p$'s proper axioms, that is if you presuppose that $u \in M_i [\boldsymbol{T}_p]$, in which $M_i [\boldsymbol{T}_p]$ is an axiomatizable proper subset of $M[\boldsymbol{T}_p]$, then u also satisfies some formulae that are equivalent to $\boldsymbol{T}_r$'s proper axioms.

In this situation, we can say that $\boldsymbol{T}_r$ (at least in the area of relevant experiential situations) is **reducible** to $\boldsymbol{T}_p$, and, in particular, that, for any (ontologically relevant) $v \in M[\boldsymbol{T}_r]$ with $D_1(v) = C$, there is a corresponding $u \in M[\boldsymbol{T}_p]$ with $D_1(u) = D$, such that $C \subseteq \wp(D)$. Therefore, it seems plausible to admit that, in spite of the difference between the theories applicable to different experiential situations, the ontological unity has been restored: Chunks are "in fact" sets of dots, and consequently, the Being "hidden behind the phenomena" still only consists in dots.

There are several directions in which this possible situation may be generalized in order to speak of a restablished unity of Being in spite of having different, non-equivalent theories that are applicable to different experiential situations.

1. One possible direction is this. The set-theoretical relationship between C and D may come out as being more complex than the simple formula $C \subseteq \wp(D)$ suggests. It could be the case that the experiential situations require several distinct basic domains $D_1, ..., D_n$, instead of a single D, but, nevertheless, allow for a reconstruction of C as a complex configuration of a *relational* kind over D_i; an example for this might be illustrated by the formula:

$$C \subseteq \wp\ (B_1 \mathrm{X}\ \wp\ (B_2) \mathrm{X}\ B_3).$$

In general, we may suppose that C can be constructed as an **echelon set** out of previously given domains $B_1, B_2, ...$ In this case too, it is plausible to say that what "really exists" is not constituted by the elements of C but rather by the basic elements of those B_i that settle the base for the echelon set C.

2. Second, it will usually be the case that the derivation of the laws satisfied by v from the laws and additional special conditions satisfied by v's counterpart u (as illustrated in Point c above) does not work exactly but only *approximatively*. To deal with this case in a serious way, a serious (i. e., a precise and plausible) notion of approximation as a kind of intertheoretical relation is needed. But its formal explication poses no particular problem (at least not as a matter of principle): The structuralist approach provided such an explication (and its illustration by means of real-life examples) some time ago; it is essentially based on the notions of a *uniform structure* and a *blur* (as a kind of "model-theoretical fuzzy-set"). For details, see Moulines (1980) and (1981), as well as Balzer, Moulines and Sneed (1987, Ch. 7). At any rate, what interests the ontologist here is that, even if there is no exact derivation of the laws of one theory from those of the other,

and therefore there is no exact relationship between their corresponding models, the notion of intertheoretical approximation makes it possible to recover the ontological unity.

3. Third, there is some hope for an ontological unification even in those cases in which the laws of one theory cannot be derived either exactly or approximatively from the laws and special conditions of the other. Imagine the following situation: There is a theory T with a model $v \in M[T]$ and a basic domain C, so that *GP* "suspects" that, for another theory $T°$, the model v with its domain C has "something to do" with $T°$. However, even by adding special conditions to $T°$, *GP* is not able to find a v-corresponding model in $T°$, $u \in M_i[T°]$, allowing for an exact or approximative derivation of T's laws from the laws and special conditions of $T°$. Still, *GP* should not necessarily despair. One might be able to establish the following intertheoretical relationship (expressed in exclusively model-theoretical terms) between T and $T°$: v can be connected with a particular $u° \in M_i[T°]$ with a basic domain D and subsuming the same experiential situation as before (or a similar one); this connection may consist, for example, in the identification "$C \subseteq \wp$ (D)" without thereby hindering v's ability to subsume the corresponding experiential situation. In this case, we could conclude that, even though the *nomological* reduction between both theories is no longer possible we may speak of an *ontological* reduction, and we (or *GP*) may claim that the elements of C of T are "nothing but" sets (or structures) of elements of D of $T°$.

Let us summarize the results of our considerations. Let us suppose we have a theory T with some models subsuming – through idealization – some experiential situations that interest us; in these models, the basic domains $D_1, \ldots, D_n$ appear. Let us further assume we find (or invent) another theory $T°$ possessing models subsuming the same, or similar, and possibly other interesting experiential situations and having basic domains $D°_1, \ldots, D°_m$. And let us finally suppose that, for any relevant model v of T and for all its basic domains D_i, we can find a relevant model $u°$ of $T°$ with one or several domains $D°_j$, such that v and $u°$ are linked together through at least one of the configurations depicted in Situations 1 to 3 above. In such a case, we may claim that T is **ontologically reducible** to $T°$, and that the only ontologically relevant commitments we have taken are those corresponding to $T°$. The only Real Being is being according to $T°$.

A final word on matters ontological. Imagine for a moment we would have a single BIG THEORY $T°$ to which all other scientific theories T_i would have the kind of relationship we have just called "ontological reducibility." In this situation, I think we would be warranted in claiming that we have a unified ontological picture of THE WORLD. Is present-day science in this situation? Certainly not! Some scientists, especially physicists, are trying very hard to make it come about; some other scientists, mainly nonphysicists, are doing their best to hinder it; and most other scientists just don't care. We, as philosophers of science, cannot decide the issue; but, at least, we know in quite precise terms what it would be like to have either a positive or a negative answer to the question.

Institut für Philosophie, Logik und Wissenschaftstheorie, Universität München

NOTES

[1] For an assessment of a somewhat modified Quinean perspective on ontological matters, see Moulines (1994) and (1998).
[2] By "science," I mean the collection of well-established, institutionally anchored scientific disciplines.
[3] In this discussion, I shall leave aside the realm of pure mathematics, not because I think that the ontological questions related to mathematics are uninteresting, quite the contrary, but rather because dealing with them would go far beyond the scope of this article. In the present context of discussion, mathematics will be considered in its purely instrumental value for empirical science.

REFERENCES

Balzer, W., Moulines, C. U., Sneed, J.D. (1987). *An Architectonic for Science*. Dordrecht: Reidel.
Moulines, C. U. (1980). Intertheoretic Approximation: The Kepler-Newton Case. *Synthese* 45, 387-412.
Moulines, C. U. (1981). A General Scheme for Intertheoretic Approximation. In A. Hartkämper, H. J. Schmidt (Eds.), *Structure and Approximation in Physical Theories*. New York: Plenum, 123-146.
Moulines, C. U. (1994). Wer bestimmt, was es gibt? Zum Verhältnis zwischen Ontologie und Wissenschaftstheorie. *Zeitschrift für philosophische Forschung* 48.2, 175-191.
Moulines, C. U. (1998). What Classes of Things Are There? In C. Martinez, U. Rivas, L. Villegas (Eds.), *Truth in Perspective. Recent Issues in Logic, Representation and Ontology*. Hants: Ashgate, 317-330.
Quine, W. V. O. (1953). *From a Logical Point of View*. New York: Harper and Rowe.

MARCO PANZA

SOME SOBER CONCEPTIONS OF MATHEMATICAL TRUTH

Abstract. It is not sufficient to supply an instance of Tarski's schema, $\ulcorner$"*p*" is true if and only if *p*$\urcorner$ for a certain statement in order to get a definition of truth for this statement and thus fix a truth-condition for it. A definition of the truth of a statement *x* of a language *L* is a bi-conditional whose two members are two statements of a meta-language *L'*. Tarski's schema simply suggests that a definition of truth for a certain segment *x* of a language *L* consists in a statement of the form: $\ulcorner \nu(x)$ is true if and only if $\tau(x)\urcorner$, where $\ulcorner \nu(x)\urcorner$ is the name of *x* in *L'* and $\tau(x)$ is a function $\tau: S \rightarrow S'$ (*S* and *S'* being the sets of the statements respectively of *L* end *L'*) which associates to *x* the statement of *L'* expressed by the same sentence as that which expresses *x* in *L*. In order to get a definition of truth for *x* and thus fix a truth-condition for it, one has thus to specify the function τ. A conception of truth for a certain class *X* of mathematical statements is a general condition imposed on the truth-conditions for the statements of this class. It is advanced when the nature of the function τ is specified for the statements belonging to *X*. It is sober when there is no need to appeal to a controversial ontology in order to describe the conditions under which the statement $\tau(x)$ is assertible. Four sober conceptions of truth are presented and discussed.

Key words: mathematical truth, conception of truth, truth-condition, statements vs. sentences, Tarski's condition.

Truth is generally considered to be a crucial matter in the philosophy of mathematics. It is quite common to define realism in mathematics as the thesis that mathematical statements can be true, but their (eventual) truth does not depend on the fact that they are proved, or even that they could be proved in our mathematical theories, being rather somehow independent of us. In other words, it is generally accepted that to be realist in mathematics means to consent to the thesis that truth or falsehood are intrinsic properties of mathematical statements.

I do not intend to discuss this thesis here. My aim is much more modest. I simply observe that, commonly, this definition assumes implicitly not only that there are mathematical statements, but also that they are homogeneous with respect to the property of them that makes them eventually true or false. This means that mathematical statements, whether true or false, are all so in the same sense or for the same sort of reasons. If one does not take this for granted but nevertheless accepts the previous definition, one can hardly admit that realism in mathematics is a well-defined and even a consistent thesis, and thus argue in favour of or against it. One evidence for this is that people who consider themselves to be realist in mathematics in the previous sense mainly look for a suitable conception of mathematical truth, and only for one.

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 335 – 347.

When mathematics is considered abstractly as a (more or less well-defined) domain of justified beliefs, or even as a system of statements of a certain sort that belong to such a domain because of their form, their content, or the modality of their justification, such an assumption is quite tenable. Explicitly or not, it can even enter the definition of mathematics itself and, when the realist thesis (as previously defined) is accepted, also contribute to the definition of mathematical knowledge. In contrast, it seems to me that when mathematics is considered as an actual practice or activity accompanying the history of humanity, or even as the system of issues resulting from such a practice or activity, this same assumption is quite doubtful. There is no doubt, I think, that from such a point of view, one may rightly speak about mathematical statements, that is, admit that the term "mathematical statement" refers to genuine objects. And I also hold that one can sensibly assign to these latter objects the property of being true. But, when mathematics is conceived in such a way, there is no guarantee that these objects are homogeneous with respect to what makes they are true or false; there is no guarantee that, if true, they are all so in the same sense or for the same sort of reasons.

There are at least two ways to argue that this is not really the case. The first one consists in seeking out various occurrences of the term "true" and their cognates in texts unanimously considered as mathematical ones, and to show, by means of textual analysis, that this term does not have the same sense across these occurrences. The problem with this strategy is that the argument it provides could be countered by arguing that, although they occur in mathematical texts, some of these occurrences are not specifically mathematical. Thus, I prefer to look beyond this objection and to proceed in the second way. I shall present different senses in which it seems to me that one should admit that a mathematical statement could be true, and argue that when a statement is true according to one of these senses it is not true for the same sort of reason as when it—or any other statement—is true according to any other one of them.

In doing this, my intention is not to argue against realism in mathematics. Far from it, I argue that if the opposition between realism and anti-realism in mathematics is genuine and crucial, and if one speaks of mathematical truth in the senses I shall consider here, then this opposition is not concerned with truth. This is the reason why I take these senses to be sober ones. My aim is simply to show that it is possible to speak of truth in mathematics without being engaged in doubtful and controversial ontology. Of course, I admit that one could also speak of truth in mathematics in other and not sober senses. It is even a matter of fact that a lot of philosophers and mathematicians have done and do this. I do not intend to make a stand against their attitude here. I simply advance the hypothesis that sober senses (the ones I shall consider or some other ones that could be added to my list) would suffice to give a satisfactory account of mathematical practice or activity (though I concede that this account might not meet the intentions and metaphysical convictions of some mathematicians).

1

Up to now, I have generically used the term "sense" in order to refer to a certain way to conceive truth in mathematics. To be more precise, I shall from now on speak of truth-conditions and of definitions and conceptions of truth. I shall speak of truth-condition and definition of truth for a certain statement in order to refer respectively to the condition that such a statement has to satisfy in order to be true, and to the statement that specifies such a condition. I shall speak instead of conception of truth for a certain class of statements in order to refer to a general condition imposed on the truth-conditions for the statements of this class. Thus, in order to present different senses in which one should admit that mathematical statements could be true, I shall advance different conceptions of truth for different classes of mathematical statements.

Before doing this, let me present a general condition that, in my opinion, any definition of truth for a mathematical statement, should satisfy.

This is Tarski's condition. I require that any definition of the truth of a mathematical statement should be expressed by an instance of Tarski's schema: ⌜"p" is true if and only if p⌝. I shall not argue in favour of such a condition. I simply argue that, taken as such, an instance of Tarski's schema is nothing but a sentence of a certain language, *i. e.* a well formed combination of terms (or formula) of this language. It is thus not sufficient to supply an instance of Tarski's schema for a certain statement in order to get a definition of truth for this statement and thus fix a truth-condition for it. In order to do that, it is also necessary to interpret such a sentence, that is, to take it as being the expression of a certain statement.

Suppose that a particular instance of Tarski's schema is uttered in a certain language L' in order to provide a truth-condition for a certain statement x of another language L. L' should then serve as a meta-language with respect to L. Hence L' would contain a name for x, to be used to form the first member of the bi-conditional constituting such an instance of Tarski's schema. This name should not only denote x in L', but also be, as such, functionally related to the second member of this bi-conditional. And this second member should, in turn, be a sentence of L' expressing a certain statement. It is just this latter statement that fixes a truth-condition of x.

A definition of the truth of a statement x of a language L is thus a bi-conditional whose two members are two statements of a meta-language L', the first saying that x has the property to be true, and the second being the value taken by a function $\psi: N' \rightarrow S'$ (in which N' is the set of the names of L' and S' the set of the statements of L') when its argument is the name of x in L'. Of course, not every name of L' could be an argument for such a function. In order to be so, a name of L' should be the name of a statement of L and result from a quotation. Tarski's schema would then suggest the following: Take the statement x as it is uttered in L; transform it into its same name in L' by adding quotation marks to the sentence of L which expresses it, and use this name to form a statement of L' saying that x has the property of being true; disquote such a sentence in order to obtain another statement of L'; and form a bi-conditional whose two members are given by these two statements of L'. This means that Tarski's schema would suggest to define a function $\tau: S \rightarrow S'$ (S being the set of the statements of L) by composing a function $\nu: S \rightarrow N'$, giving the name of x in L' when applied to x, with another function

$\psi: N' \rightarrow S'$. It is just the value of this function τ that ultimately would provide a truth-condition for x.

Suppose now that the function ν is simply the quotation function: applied to a statement of L, it gives a name for this statement in L' by simply adding quotation marks to the sentence which expresses this statement. Then the function ψ can not simply be the disquotation function, since when this function is composed with ν, it can only produce the identity function and not a function τ from S to S'. Thus, either ν is not simply the quotation function, or ψ is not simply the disquotation function. It follows that either there is more in Tarski's schema that the simple allegation of the quotation and the disquotation functions (jointly with the instruction to form a suitable bi-conditional by using their values), or it is not sufficient to supply an instance of Tarski's schema for a certain statement in order to get a definition of truth for this statement and thus fix a truth-condition for it. But, in Tarski's schema, there is nothing more than that; hence, it is not sufficient to apply this schema to a certain statement in order to get a definition of truth for this statement and thus fix a truth-condition for it.

A definition of truth of a certain statement x of a language L is rather a bi-conditional statement of a meta-language L', whose first member says of x that it has the property of being true, and the second one is the value taken by a function τ from S to S' when it is applied to x, under the condition that this function results from the composition of the quotation function ν from S to N' and a function ψ from N' to S' such that $\psi(\nu(x))$ is the statement of L' expressed by the same sentence that expresses x in L. This bi-conditional statement has thus the form: $\ulcorner \nu(x)$ is true if and only if $\tau(x)\urcorner$, where $\ulcorner \nu(x)\urcorner$ is the name of x in L' and $\tau(x) = \psi(\nu(x))$.

Of course, a definition of this sort can only be uttered in L' if this language contains all the terms of L entering in x. This is an obvious necessary condition. But it is not sufficient, since to say that a statement of L' is expressed by the same sentence that expresses it in L is not sufficient in order to identify this statement. The work of the function ψ is just to identify this statement.

Thus, either it is admitted that at least some ones of the sentences of L' are such that one of them may express different statements, or it is admitted that the statements of L' are not given independently of the specification of the functions ψ and τ. I favour the second possibility. After all, it seems to me very natural to conceive the language L' as being constructed on the basis of L just in order to utter truth-conditions for the statements of this latter language, rather than as an already given language fortuitously satisfying all the conditions that should be satisfied by a meta-language with respect to L where a definition of truth for a statement of L could be uttered.

It follows that, in my opinion, in order to advance a conception of truth for a certain class X of mathematical statements of a certain language L, one should simply specify the nature of the function τ entering the truth-conditions of the statements belonging to X and allowing the determination of the statements of L'. This is what I shall do.

Before to do that, let me consider some consequences of Tarski's condition. Suppose that one has specified the nature of the function τ entering the truth-conditions of the statements belonging to a certain class X of mathematical statements, and has thus advanced a conception of truth for this class of mathematical statements. One can then advance a conception of falsehood in the same way, by assuming that a definition of the falsehood of a certain mathematical statement x belonging to a certain language L and a certain class X of mathematical statements is a bi-conditional statement of a language L' working as a meta-language with respect to L, namely the statement: $\ulcorner \nu(x)$ is false if and only if non $\tau(x)\urcorner$, where $\ulcorner$non $\tau(x)\urcorner$, is the negation of the statement $\ulcorner\tau(x)\urcorner$ in L'. It is then only a question of propositional logic to derive in L' the following bi-conditional: $\ulcorner \nu(x)$ is not true if and only if $\nu(x)$ is false$\urcorner$. But what about the truth or the falsehood of the negation of x in L? If such a negation belongs to X, the previous definitions give, by substitution: $\ulcorner \nu$(non x) is true if and only if τ(non x)$\urcorner$ and $\ulcorner \nu$(non x) is false if and only if non τ(non x)$\urcorner$. However, nothing enables us in general to derive from here the bi-conditionals: $\ulcorner \nu$(non x) is true if and only if non $\tau(x)\urcorner$, or $\ulcorner \nu(x)$ is false if and only if τ(non x)$\urcorner$. To do that, we would have to admit that τ(non x) is identical with non $\tau(x)$, and nothing enables us to admit that in general. Nevertheless, nothing prevents the function τ from satisfying such a further condition. If this is the case, it is then only a question of propositional logic to derive the bi-conditional $\ulcorner \nu$(non x) is true if and only if $\nu(x)$ is false$\urcorner$, and, if either the logic in L or the logic in L' are classical, also the conditional $\ulcorner \nu$(non x) is false if and only if $\nu(x)$ is true$\urcorner$. It would follow that the truth of the negation of x is equivalent to the falsehood of x, and if either the logic in the language which x belongs to or the logic in the language in which this definition is stated are classical, it is also such that the truth of x is equivalent to the falsehood of the negation of x. This is a remarkable condition, and it seems to me that we are entitled to qualify a conception of truth that satisfies it as a correspondentist conception of truth.

2

The first conception of truth I shall advance is just a sober correspondentist conception. Thus it constitutes an example proving that a correspondentist conception of truth for a class of mathematical statements can be sober.

2.1

Take a sentence like the following one: "the primitive of $\dfrac{5x^2+4ax}{2\sqrt{ax+a^2}}$ is $\dfrac{x^2}{a}\sqrt{ax+a^2}+C$." It seems to me that one possible way of understanding the statement expressed by such a sentence, say s, is to take it as referring to the symbolic expressions that enter it. If so, the statement s tells us that if one applies the derivative algorithm to the expression $\frac{x^2}{a}\sqrt{ax+a^2}+C$, one obtains the expression $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$. It is thus a statement describing a relation between two symbolic expressions, that is, two equivalence-classes of empirical objects such as concrete signs. One is then entitled, I think, to take s as true, just because the application of the derivative algorithm to the expression $\frac{x^2}{a}\sqrt{ax+a^2}+C$ gives the expression $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$. This is simply a way to understand the term "primitive." Let us suppose that s belongs to a language L. This means that L is the language in which the sentence "the primitive of $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$ is $\frac{x^2}{a}\sqrt{ax+a^2}+C$" expresses s. The function τ entering a possible truth-condition of s should thus simply associate s to a statement $\tau(x)$ of a suitable meta-language L' expressed in such a language by the same sentence "the primitive of $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$ is $\frac{x^2}{a}\sqrt{ax+a^2}+C$" and asserting here that the application of the derivative algorithm to the expression $\frac{x^2}{a}\sqrt{ax+a^2}+C$ gives the expression $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$. In other terms, this function should associate s to a statement of L' telling us the same as that which the statement s tells us in L.

It is easy to generalise this example. Let us consider a class X of mathematical statements of a language L asserting that a certain rule of formal transformation applied to a certain expression produces or does not produce another given expression. A possible conception of truth for such a class of statements simply requires that the function τ defined over this class associates any statement x of X to a statement of L' telling us the same as that which the statement x tells us in L; that is, that the rule of formal transformation that x refers to, when applied to the first expression that x refers to, produces the second expression that x refers to. I suggest that a lot of mathematical statements can be conceived as asserting that a certain rule of formal transformation applied to a certain expression produces or does not produce another given expression, and that one is thus entitled to apply to them such a correspondentist conception of truth.

2.2

To pass to other conceptions of truth, let us consider the same sentence as before: "the primitive of $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$ is $\frac{x^2}{a}\sqrt{ax+a^2}+C$." One could also understand the statement expressed by such a sentence, say s^*, by referring the term "primitive" not to the derivative algorithm, but to the general definition of derivative. If so, s^* tells to us that the limit of $\frac{\frac{(x+h)^2}{a}\sqrt{a(x+h)+a^2+C-\frac{x^2}{a}\sqrt{ax+a^2-C}}}{h}$ when h tends toward 0 is equal to $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$. If the derivative algorithm is taken as being justified by having shown that it satisfies the general definition of derivative, the difference between the statement s^* and the statement s considered in the previous example is far from being essential from a mathematical point of view. This is not my point, however. What is important for me is that the equivalence of these two statements has to be set up by means of a proof, which is not, as such, part of these two statements. It seems to me that this allows us to consider these two statements as distinct from each other.

One could certainly hold that to say that the limit of $\frac{\frac{(x+h)^2}{a}\sqrt{a(x+h)+a^2+C-\frac{x^2}{a}\sqrt{ax+a^2-C}}}{h}$ when h tends toward 0 is equal to $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$ is nothing but asserting that the first expression can be transformed into the second one by applying certain rules of transformation in a suitable way. If so, the situation does not change with respect to the case considered in Section 2.1: Though the truth of the statement s and the truth of the statement s^* do not depend on the same reason, the reasons they depend on are of the same sort; and the truth-conditions of these statements satisfy the same conception of truth, since a certain sequence of rules of transformation is itself a rule of transformation. But one could also argue that this is not so: that when speaking of limit, we are implicitly referring to functions, whereas when operating according to certain rules of transformation, we are working with symbolic expressions, eventually with the symbolic expressions expressing these functions. When we are considering particular functions like the previous ones, the difference between a function and the expression that expresses it could appear to be minor. But it is not certainly so in general, and it is not so in particular, when we are referring to a class of functions respecting certain general conditions, like continuous functions or limited functions. Thus, it should be clear that a statement referring to these functions can not be understood as asserting that a certain rule of formal transformation applied to a certain expression produces or does not produce another given expression.

Hence, if one wanted to understand the statement s^*, or any other statement about functions, as a description of some objects and/or their relations, one should admit that these objects do not simply consist of equivalence-classes of empirical objects. This is also the case for statements about numbers, sets, algebraic structures, and a lot of other mathematical entities that cannot be understood as equivalence-classes of empirical objects. One could think that there is no way to admit that these statements can be true or false unless their truth-conditions depend on a function τ that is supposed to associate any one of these statements to a statement of a suitable meta-language telling us that the objects it refers to have the properties or relations it assigns to them. A similar conception of truth is hardly a sober one. Far from arguing here that a similar conception of truth is not tenable, I limit myself to ignoring it and advancing three sober alternative options.

2.2.1

The first two options consist in requiring that the function τ associates any statement x of the considered class of statements of a language L to a statement of a suitable meta-language L', telling us respectively that the statement x has been proved or is provable within the theory to which it belongs.
Of course, to say that x has been proved in a certain mathematical theory is not the same thing as saying that it is provable in such a theory, whatever our notion of proof. I shall not enter into this distinction here. I simply observe that in the first case, $\tau(\text{non } x)$ is certainly not identical with non $\tau(x)$, whereas in the second case, the identity of $\tau(\text{non } x)$ and non $\tau(x)$ depends on the nature of the mathematical theory to which x belongs. Thus, the conception of truth depending on the first option is never correspondentist, whereas the conception of truth depending on the second is not so in general.

However, if the first option were accepted, then it would be very unsatisfactory to define the falsehood of x by the bi-conditional ⌜$\nu(x)$ is false if and only if non $\tau(x)$⌝. The bi-conditional ⌜$\nu(x)$ is false if and only if $\tau(\text{non } x)$⌝ would be preferable. Hence, it would be only a question of propositional logic to derive in L': ⌜$\nu(\text{non } x)$ is true if and only if $\nu(x)$ is false⌝. By substitution, one also would have: ⌜$\nu(\text{non } x)$ is false if and only if $\tau(\text{non non } x)$⌝. Thus, if the logic in L were classical, it would also be only a question of propositional logic to derive in L': ⌜$\nu(\text{non } x)$ is false if and only if $\tau(x)$⌝ and thus ⌜$\nu(x)$ is true if and only if $\nu(\text{non } x)$ is false⌝. But it would not be possible to derive: ⌜$\nu(x)$ is not true if and only if $\nu(x)$ is false⌝. Thus a statement x could be not true, without being false.

One could object to these conceptions of truth by drawing on a classic argument first presented by Tarski. Using godelisation, one can associate injectively each sentence of a formal theory, that is, a mathematical one, to a natural number, and thus each set of sentences of a formal theory to a set of natural numbers. Supposing that any sentence of a formal theory is associated bijectively which a statement that is expresses, and that to prove a statement in such a theory is nothing but to derive in it the sentence that expresses it, one could then wonder whether the set of numbers associated to provable statements and the set of numbers associated to true statements are identical. Now, because of the nature of a formal theory, the first of these sets can be characterized in terms of simple arithmetical operations and relations. One could then translate the definition of provability into the language of the theory, that is the object-language. If a similar translation were also possible for the definition of truth, this language would be semantically universal and then it would be possible to use it to formulate the antinomy of the liar. Thus, either the antinomy of the liar can be formulated in the language of a formal theory, or provability within this theory and truth of its statements are not extensionally equivalent properties.

Though admitting that a mathematical theory could take the form of a formal theory and that a proof of a statement in such a theory reduces to a formal deduction of a sentence, this argument does not seem to me to be a conclusive reason to reject the previous conceptions of truth for the statements of such a theory. After all, these are sober conceptions, and they are simply supposed to correspond to two different senses in which mathematicians speak or have spoken of truth. Thus, if Tarski's argument is correct, it simply discloses that these senses make it possible to formulate the antinomy of the liar within a mathematical formal theory. Notice, moreover, that this is not the same as asserting that this theory is not consistent, since the theory itself is completely independent of the nature of the meta-language used to assign to its statements the property of being true.

2.2.2

The comparison between the statements *s* and *s** expressed by the same sentence in a language *L* suggests the possibility of defining the truth of one of these statements, say *s*, by means of a statement of a meta-language *L'*, telling us in this latter language the same as that which the other statement, say *s*, tells us in *L*. In such a way, one could also assign a sober sense to the distinction between truth and proof or provability. One could admit that in order to prove that the primitive of $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$ is $\frac{x^2}{a}\sqrt{ax+a^2}+C$, one should show that the limit of $\frac{\frac{(x+h)^2}{a}\sqrt{a(x+h)+a^2}+C-\frac{x^2}{a}\sqrt{ax+a^2}-C}{h}$ when *h* tends toward 0 is equal to $\frac{5x^2+4ax}{2\sqrt{ax+a^2}}$, though maintaining that the statement *s** is true because the application of the derivative algorithm to the second of these expressions simply produces the first. When understood in this way, the distinction between truth and proof or provability reduces to a methodological one and just concerns the internal organisation of a certain mathematical theory.

A similar view can be applied to universal statements. Let us take the example of the last theorem of Fermat. This tells us that if n is a natural number greater than 2, then there is no trio of strictly positive natural numbers x, y, and z such that $x^n + y^n = z^n$. The simpler way to understand this theorem is to take it as assuring us that no substitution of the letters "x," "y," "z" and "n" in the sentence "$x^n + y^n = z^n$" with symbols of natural numbers satisfying the previous conditions transforms this sentence into numerical identity. However, there is no way to prove it by considering such substitutions. To do that, it should be rather showed that the fact that x, y, z, and n are four natural numbers satisfying these conditions is a sufficient condition for the sum $x^n + y^n$ to be different from z^n.

The same point can be made in general. Any universal theorem can be understood in at least two ways: a distributive or extensional way and a compact or intensional way. Extensionally, it asserts that every single element of a certain domain has a certain property; intensionally, it asserts that it is sufficient to belong to such a domain to have this property. If the domain is an infinite one, there is no other way to prove this theorem than by proving the latter. However, one should maintain that this theorem is true because of the former.

In general, one could identify a theorem (of a certain sort) with an equivalence-class of statements composed of two statements, and assume that the proof of this theorem is concerned with one of these statements, while its truth is concerned with the other.

It would then be sufficient to treat an equivalence-class of statements as a statement whose utterance is nothing but the utterance of one of its members in order to be able to present a conception of truth that fits with such a view. Let us consider a class X of mathematical statements of a language L and two relations of equivalence defined over X, say R' and R'', such that R' divides X into several equivalence-classes, each of which is composed of two distinct statements, and R'' divides X in two equivalence-classes respectively composed of one and only one of the two statements making up each one of the classes of equivalence in which X is shared by R'. Let X_p and X_t be the two classes in which X is divided by R''. A conception of truth for X could require that the function τ defined over this class associates any statement x of X to a statement of a suitable meta-language L' telling us the same as the statement of L equivalent to x according to R' and belonging to the class X_t. At the same time, one could suppose that in order to prove any statement x of X, one should produce an argument showing that the things are as it is said by the statement of L being equivalent to x according to R' and belonging to the class X_t.

In the case of the last theorem of Fermat, the role of the statement x is taken by any one of the two statements telling us respectively that: no substitution of the letters "x," "y," "z," and "n" in the sentence "$x^n + y^n = z^n$" with symbols of naturals numbers satisfying the given conditions transforms this sentence into a numerical identity; and that the fact that x, y, z, and n are four natural numbers satisfying these condition is a sufficient condition for the sum $x^n + y^n$ to be different from z^n. The role of the statement of L equivalent to x according to R' and belonging to the class X_t is taken by the first of these two statements, and the role of the statement of L

equivalent to x according to R' and belonging to the class X_p is taken by the second one.

If X is a class of universal statements, the relations R' and R'' can be defined on this class in such a way that R' associates one to each other two statements telling us respectively that every single element of a certain domain has a certain property, and that it is sufficient to belong to such a domain to have this property, while X_t and X_p coincide respectively with the class of the elements of X having an extensional form and the class of the elements of X having an intensional form. It is then possible to generalise to any statement of X the previous way to fix the truth-condition of the last theorem of Fermat. However,, this is nothing but a particular case of the conception of truth I am presenting here, since nothing prevents us from defining the relations R' and R'' on a certain class of mathematical statements, either universal or not, in a different way. The example of the statements s et s^* suggests a way to do it in a certain case. Other strategies could be followed in other cases.

I shall not enter into a discussion of these strategies. I limit myself to observing that such a conception is a correspondentist one, and that it is, as such, sober. Of course, one could associate it with an understanding of distributive statements that depends on some strong ontological or epistemological condition and thus transform it into a non sober conception. Nevertheless, this is not necessary. In order to refer to such a conception saying that a certain mathematical statement x is true and distinguishing between its truth and its proof or provability, it is sufficient to have at one's disposal a procedure able to decide, in any specific case covered by this statement, whether the things are or are not as claimed by the statement of L equivalent to x according to R' and belonging to the class X_t. In the case of the last theorem of Fermat, it is, for example, sufficient to have at one's disposal a procedure able to decide, for every set of four natural numbers $<x, y, z, n>$ satisfying the given conditions, whether $x^n + y^n = z^n$ or not. And there is no doubt that all who are acquainted with elementary arithmetic and have a sufficient computational capacity have such a procedure at their disposal. Of course, if the specific cases one should consider in order to exhaust the domain covered by the statement under examination are infinite in number, one would never know in this way whether this statement is true or not. But, far from being an unsatisfactory consequence of the present conception of truth, such a circumstance shows how close this conception is to any acceptable conception of truth for empirical universal statements.

2.3

The last conception of truth I consider here is a very classic and traditional one. I shall limit myself to observe that, under suitable conditions, this conception can be understood as being sober, and to present it in an uncustomary frame.

Let us take the example of the Bolzano-Weierstrass theorem: If $f(x)$ is a continuous function from $\boldsymbol{R}$ to $\boldsymbol{R}$ defined both in a and b, K is a real value, and $f(a) < K < f(b)$, then there is a real value c such that $f(c) = K$.

In classical analysis, this theorem is proved by reduction to absurd, by showing that its negation contradicts the axiom of the superior upper bound. This is a non-constructive proof that (from a classical point of view) only warrants that c exists without exhibiting it. Its legitimacy has been thus the object of several discussions. I do not want to return to such an issue here. I simply observe that the principal reason for a mathematician to be not disposed to renounce to this theorem is not a logical one, being rather concerned with the expressive power of mathematics.

A theory of real continuous function in which it is not possible to prove the Bolzano-Weierstrass theorem could hardly pretend to be useful for explaining a large class of real phenomena. Suppose that John and Mary are in Paris and that John walks from the *Arc de Triomphe* to the *Place de la Concorde* along the left side of the *Champs Élysées*, while Mary is somewhere on this same side of the *Champs Élysées*, drinking a glass of *champagne*. They will probably meet there. A theory or real continuous function in which it is not possible to prove the Bolzano-Weierstrass theorem cannot be used to explain this trivial phenomenon. To conclude that such a theory would thus be unsatisfactory is the same as admitting that a mathematical theory should have an expressive power and that this power is part of its mathematical legitimacy.

Once this has been admitted, it is very natural to speak of the truth of certain mathematical statements by referring to their expressive power rather than to their intra-theoric content. Let us suppose that X is a class of statements of this sort belonging to a certain theory T and a certain language L. For advancing a conception of truth which justifies thus way of speaking, it is sufficient to define a function φ that associates any statement x of X with another statement $\varphi(x)$ referring to objects that are not part of the domain of T and require that the function τ defined over X associates any statement x of X to a statement of a suitable meta-language L' telling us the same as the statement $\varphi(x)$.

This is what one does when defining the truth of a mathematical statement with respect to a certain model of the theory this statement belongs to. A classic example is the definition of the truth for the statements of Peano's arithmetic with respect to one of its set-theoretical models. Notice, however, that the possibility to define the truth of a mathematical statement in this way does not depend on the possibility to define a model for the whole theory this statements belongs to according to the constraints of the logical theory of models. The statement expressing the Bolzano-Weierstrass theorem could, for example, be associated by φ to a statement concerning suitable curves traced on a Cartesian plane, without need for these curves to belong to a model (in the sense of the logical theory of models) for the whole theory of real functions.

It seems to me that in order to understand this conception of truth for a certain class X of mathematical statements as being sober, it is not necessary to suppose that the statements associated to the statements of X by the function φ are not mathematical in turn, or do not refer to objects that are involved as such, or could be involved with a controversial ontology. What it is needed is simply that these latter statements or any one of their particular instances (if they are universal statements of a distributive form) are somehow decidable, that is, that there is an effective procedure to decide whether these statements or any one of their particular instances can be asserted or not.

3

This last remark, together with the example given by the conceptions of truth presented above, should make clearer what I mean by "sober". Taking the Tarski's condition for granted, I maintain that a conception of truth for a class X of mathematical statements of a language L is sober if and only if there is no need to appeal to a controversial ontology in order to describe the conditions under which the statement $\tau(x)$ of L' is assertible.

As there is neither a generally accepted criterion to decide what is a controversial ontology nor a general accepted definition of assertibility, and my formulation of Tarski's condition does not involve any strict constraint on the nature of the statement $\tau(x)$, it would be easy to object that my characterization of the general notion of a sober conception of truth for a class of mathematical statements is so large that one could arbitrarily suggest many other sober conceptions of truth. I accept the point, but I do not think that this is an argument against my views.

I have two reasons for maintaining this.

First, it seems to me that however large it could be, my characterisation of the general notion of sober conception of truth for a class of mathematical statements sets up a general form that a sober conception of truth should satisfy. I claim this very useful because I think that the only proper way to speak of truth in general is to fix a form that a certain predicate should satisfy in order to be taken in certain contexts as the predicate "to be true." This is the same as arguing that the term "true" should be taken in general—that is, independently of any specific and contextual definitional clause—as referring to an equivalence-class of predicates rather than to a single and well-defined predicate.

My second reason is also a justification of this attitude: I argue that as far as philosophy may lead us, it cannot do more than provide us with some general categories to be used to study real phenomena. Philosophy of mathematics should provide us with some general categories that can be used to study mathematics as a given reality. Though philosophy certainly has a history, a method, and a disciplinary content to which it refers, these categories should not be shaped abstractly. What ultimately decides whether they are the good ones or not is neither philosophy nor logic. It is rather the reality to which they should be applied.

REHSEIS, CNRS and Université de Paris 7

JEAN PAUL VAN BENDEGEM

CAN THERE BE AN ALTERNATIVE MATHEMATICS, REALLY?

Abstract. David Bloor, already in 1976, asked the question whether an alternative mathematics is possible. Although he presented a number of examples, I do not consider these really convincing. To support Bloor's view I present three examples that to my mind should be considered as genuine alternative: (a) vague mathematics, i. e., a mathematics wherein notions such as 'small', 'large' and 'few' can be used, (b) random mathematics where mathematics consists (almost) solely of a practice, and (c) a mathematics where infinitesimals can be used without any problem, on the assumption that one is willing to work with local models only and to resist looking for global models. Finally, I argue that these examples support Otte's thesis that an ontology is constituted by a practice and not vice-versa.

Key words: alternative mathematics, mathematical practice, non-compactness, randomness, sociology of mathematics, vagueness.

1. INTRODUCTION

In the first edition of *Knowledge and Social Imagery*, David Bloor, one of the founding fathers of the *Strong Programme* in the philosophy and sociology of the sciences, raises the important question of the possibility of an alternative mathematics. The fact that this question needs to be dealt with is rather obvious: if (the production of) knowledge is a social process, then this must also apply to that part of human knowledge that seems to resist most strongly this sociological turn: mathematics. Especially among Western mathematicians and philosophers there is a deep belief shared by most, if not all, that mathematical knowledge is (a prime candidate for) necessary knowledge. Surely one of the possible roads to follow to attack this necessity view is to show that alternatives are possible, leaving open the matter why it is that we happen to have the mathematics today we actually practice. However this task is actually a very tricky matter. As Bloor himself remarks:

> To decide whether there can be an alternative mathematics it is important to ask: what would such things look like? By what signs could they be recognised, and what is to count as an alternative mathematics? (Bloor 1991, 107)

It is worthwhile, I think, to have a brief look at the specific examples Bloor presents in the chapter that carries the same title as this paper (apart from the addition of 'really', which, of course, is a reference to Reuben Hersh's book). He presents four cases, here grouped into three[1]:

(a) *The nature of numbers*. Here Bloor argues that in different historical periods numbers such as one and zero were interpreted in quite different ways compared

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 349 – 359.

to today's practices. To give one small illustration: there is indeed a world of difference between saying that "the equation $x^2 - x - 6 = 0$ has two solutions $x = 2$ and $x = -3$, but I only use the positive solution" and saying that "the equation $x^2 - x - 6 = 0$ has exactly one solution, viz. $x = 2$," simply because negative entities can be anything you like, but they are not numbers. In turn, this goes together, as Bloor shows, with visual representations. Thus, making a diagram such that the parabola, representing $y = x^2 - x - 6$, can be drawn, it becomes inevitable to talk about two solutions on a par, so to speak, because they have a shared identity as geometrical points.

(b) *The metaphysics of numbers.* Here Bloor deals with the Pythagorean view of numbers as part of a larger metaphysical framework. These metaphysical considerations are part and parcel of mathematical activities. In support of this claim, he discusses the great classic of all (mathematical) times: the irrationality of $\sqrt{2}$. Actually by formulating the theorem in this way, one is already subscribing to a particular metaphysics, viz. there are other numbers besides the rational numbers, so the rational numbers do not exhaust all numbers. But a quite different conclusion can be drawn: $\sqrt{2}$ is not a number. Perhaps a geometrical entity, but definitely not a number.

(c) *The case of the infinitesimals.* This, of course, is the best-known example of an alternative approach in differential and integral calculus. Bloor here quotes the work, among others, of John Wallis, more specifically the case of the surface of a triangle. Let me briefly present this case. Wallis slices up the triangle into an infinite series of very thin layers, parallel to the base of the triangle. If the height of the triangle is h, and the number of layers is ∞, then a layer has height h/∞. The length of each layer varies from b, the base of the triangle to 0 at the top. This forms an arithmetical series, the sum of which is equal to the product of the average of the terms and the number of terms or, in this case, $\frac{b}{2} \cdot \infty$. Hence the surface of the triangle is $\frac{h}{\infty} \cdot \frac{b}{2} \cdot \infty = h \cdot \frac{b}{2}$.

It is obvious that different standards of rigour are at work here, but that precisely is what is at stake: the standards of rigour are not fixed once and for all, but are susceptible to (deep) changes.

In short these examples support the thesis that alternative mathematics is possible. At the same time, it seems clear to me that these examples are rather "mild." For a critic of Bloor's view[2] it is easy to remark that the cases presented are historical cases and that historical progress means precisely that: we used to believe half-baked ideas, these helped us forward, but in due time, when we found out the rubbish we believed in, we changed our minds, and so we come closer to (eternal) truth. Compare it to the fact that most mathematicians and (I hate to say) philosophers accept readily that in order to understand the mathematics of a non-Western culture, the social, the anthropological, the metaphysical, ... dimensions are required. However, for Western mathematics this is not necessary.

The purpose of this paper is to try to push the boundaries a bit further, i. e., to show that there are possibilities for alternative mathematics that go beyond the "mild" examples of Bloor, thus making his case even stronger. Probably the most important element is that for "real" alternatives, more has to be taken into account than mathematical concepts and theories on their own. At the same time, this paper can be read as a plea for the importance of mathematical practice. A consequence of looking at the matter from this point of view is that I will not talk about intuitionistic mathematics or any other form of constructivist mathematics, including strict finitism, or about inconsistent or paraconsistent mathematics or about alternative set theories (e. g., non-well-founded ones). Although some can indeed claim to be alternatives compared with standard classical mathematics, they share too many properties: they all focus on mathematical theories and mathematical proofs, there is an underlying logic implying a standard picture of the nature of a proof, all concepts are sharply delineated. No wonder that multiple translations are possible.

2. SOME EXAMPLES

I will present in this paper three alternatives: vague mathematics, random mathematics, and open (or non-compact) mathematics. The first one, vague mathematics, is perhaps the "softest" one, as I will show that it is translatable into a classical framework. On the other hand, the translation only serves the purpose to show that vagueness does not exclude in any case clear and sharp reasoning. Once one is convinced of this fact, one can simply forget about the translation and simply do vague mathematics.

2.1. Vague mathematics[3]

Vague mathematics can be considered as an attempt to make sense of statements such as "Small numbers have few prime factors." Why the emphasis on vagueness? I guess that even a rather superficial look at the history of mathematics indicates that a story can be told that centers around the gradual elimination of vague concepts from mathematics and their replacement by sharply defined concepts instead. Thus the notions of small and large, few and many numbers disappear quite soon from mathematics (see, e. g., the now rarely mentioned, yet famous text of Archimedes, *The Sand Reckoner*, dealing with large numbers[4]), and the notion of infinitesimals – perhaps one of the vaguest concepts ever – has been eliminated and replaced by the notion of limit (see Boyer (1959) for a classic, but see an 'alternative' alternative treatment of infinitesimals below in 2.3).

Let me start with elementary number theory. The language of that theory (in its first-order formulation) involves the use of constants (usually 0), variables (x, y, z, ...), predicates (P, Q, R, ...), and logical connectors (&, v, $\supset$, $\sim$, $\equiv$, $\forall$, $\exists$) and the usual formation rules. To interpret the language we need a model M, being a triple $<D, I, v>$, where D is the domain of the model, I the interpretation function that maps constants and variables on D and predicates P of rank r on a subset of D^r, the

r-cartesian product of D. Finally v is the valuation function that maps sentences A onto $\{0, 1\}$ according to some set of semantical rules.

Let me now focus on the predicates in relation to M. The interpretation of a predicate P (of rank r) will be $I+(P) \subseteq D^r$. Call $I+(P)$ the *positive* extension and $D^r \backslash I+(P) = I-(P)$, the *negative* extension. Thus to a predicate P in the language will correspond a "full" interpretation $I(P) = < I+(P), I-(P) >$ such that $I+(P) \cup I-(P) = D^r$ and $I+(P) \cap I-(P) = \emptyset$.

The crucial step is to consider not just one model, but a series $\mathbf{M}$ of models M_1, M_2, ..., M_n such that, if in M_i the predicate P has an interpretation $I_i(P) = < I_i+(P), I_i-(P)>$ and such that, if $i < j$, then $I_i+(P) \subseteq I_j+(P)$ (and, hence, $I_j-(P) \subseteq I_i-(P)$). Thus we obtain a "nested" set of models. The core idea is to apply a supervaluational method to the set $\mathbf{M}$. This means quite simply that if A is true in all models M_i, then it is True (the capital T serves to indicate the supervaluational function), if A is false in all models M_i then it is False (likewise), and in the remaining cases the truth value is Undetermined. This opens up the possibility to introduce new predicates that cannot be defined classically.

A specific example to illustrate the procedure is to introduce the predicates "small" or $S(n)$ and "large" or $L(n)$. Both are rank 1. As to their interpretation, what one has to do is to specify the set $\mathbf{M}$ of models. We can select two numbers S_1 and S_n, $S_1 < S_n$, such that, for any number S_i, such that $S_1 \leq S_i \leq S_n$, there is a corresponding model M_i, where:

$$I_i(S) = <I_i+(S), I_i-(S)> = < \{\mathbf{0}, \mathbf{1}, ..., S_i\}, \{S_{i+1}, S_{i+2}, ...\} >.$$

It is easy to see that the condition: if $i < j$, then $I_i+(S) \subseteq I_j+(S)$, is satisfied. The same procedure can be carried out for $L(n)$. Again, we can select two numbers L_1 and L_m, $L_1 > L_m$, such that, for any number L_i, such that $L_l \geq L_i \geq L_m$, there is a corresponding model M_i, where:

$$I_i(L) = <I_i+(L), I_i-(L)> = < \{L_{i+1}, L_{i+2}, ...\},\{\mathbf{0}, \mathbf{1}, ..., L_i\} >.$$

It is easy to see that the condition: if $i < j$, then $I_i+(L) \subseteq I_j+(L)$, is satisfied in this case as well.

Finally, we have to say something about how to combine the set of models for "small" with the set of models for "large." Many possibilities can be explored: each classical model for "small" can be combined with every classical model for "large", or, more restrictedly, for every "small"-model, there is one "large"-model and vice-versa. In other words, in terms of the description above, $n = m$. Let me continue with this simplified approach.

Within this framework it is easy to prove that:

Theorem: $(\forall n)(S(n) \vee \sim S(n))$ & $(\forall n)(\mathrm{L}(n) \vee \sim L(n))$. (Any number is either small (large) or not small (large)).

Suppose that we add further that for every M_i, $S_i < L_i$ (or equivalently, $S_n < L_n$). Then it is easy to prove the following:

Theorem: $(\forall n)(L(n) \supset \sim S(n))$ & $(\forall n)(S(n) \supset \sim L(n))$. (If a number n is large (small), then it is not small (large)).

Theorem: $(\exists n)(\sim S(n)$ & $\sim L(n))$ or, equivalently, $\sim(\forall n)(S(n) \vee L(n))$. (There is a number n that is neither small nor large, or, equivalently, it is not the case that every number is either small or large).

For a further development of this approach, I refer the reader to my (2000), where a proof is given of the theorem "Small numbers have few prime factors." Let me repeat once more that it is clear by this presentation that it is possible to translate this vague mathematics into classical mathematics. The important point, however, is that we can leave out the semantics and continue with the vague notions on their own. Or, put differently, if someone says that (a) small numbers have few prime factors, (b) this number has a lot of prime factors, therefore (c) it is most likely not a small number, then, if there are doubts about the correctness of the argument, one can always pull in the semantics and present a classical detailed analysis and determine under what precise circumstances the argument is definitely correct. Note that this line of handling the subject is obviously analogous to the deep belief mathematicians share that every concrete proof (as it appears in the journals, is written down on blackboards or presented at conferences) could be rewritten in an exact formal format satisfying the most stringent demands of the logician.

To round off this first alternative, let me just mention that nothing prevents a useful combination of classical and vague mathematics. This opens up the possibility that, on the one hand, we can have a (classical) theorem stating that for objects satisfying conditions C_1, C_2, ..., C_n a certain property P holds, and, on the other hand, a vague counterpart that says that for most objects P holds (if such happens to hold in vague terms). Or, one might say that we have a combination a "rough" mathematics and a "detailed" mathematics. To quote the case that was the topic of Lakatos' *Proofs and Refutations*: from a vague perspective, it should be alright to claim that for *most* polyhedra, V(ertices) – E(dges) + F(aces) = 2. But, if the game is played in more detail, then the exceptions have to be dealt with.

2.2. *Random mathematics*

The next alternative is a radical departure from mathematics as we know it presently. I will try to show that it is possible to arrive at a mathematical theory, taken in the sense of a set of true statements, *without the notion of proof.* As this presentation does not necessitate heavy formal machinery, let me present it in the form of a story.

Imagine a culture where arithmetic is developed in the following way. Let us assume that they have some notions of numbers in the sense that they can generate names for numbers that are locally ordered. By this I mean that they know that 3 follows 2 and comes before 4, but 3 compared to 1000 does not necessarily make sense. They are not incomparable as such, but it is uninteresting to do so. They also have some idea of addition in the following sense. When they are presented with an equation of the form $n + m = k$, then they can also generate the equations $(n - 1) + m = (k - 1)$, $n + (m - 1) = (k - 1)$, $n + (m + 1) = (k + 1)$ and $(n + 1) + m = (k + 1)$. Call these four equations the neighbours of the given equation. Do note that the numbers between brackets are meant to be names for those numbers.

Imagine further that empirically they discover (e. g., through the manipulation of objects) that $2 + 3 = 5$. They then accept this equation as correct and, this is the most determining element in the story, also the neighbours. Such equations are put on a list, entitled "things we know for sure." In the course of time, people ask questions of the type "What is the outcome of $n + m$ (for some specified n and m)?" To find the answer, they check the list. If it is on the list, that is the answer. If not, any answer will do. In the former case, note that all the neighbours are added to the list as well. In the course of time, this list will grow.

What will happen is that, under the assumption that, given sufficient time, any equation is likely to turn up, at the end of times the list will contain all true arithmetical statements involving addition. A simple argument to see why this must be the case is to consider a two-dimensional lattice, where each square corresponds to a couple (n, m). One starts (as in the example presented here) at (2, 3) and the neighbours are (1, 3), (3, 3), (2, 2) and (2, 4). This corresponds neatly to a random walk in the plane and, as is well known, given sufficient time, all squares will be visited. Hence the label *random* mathematics[5]. Formulated in those terms, if we call T the set of all true arithmetical statements involving addition, then in the limit T will be reached. It is perfectly acceptable therefore to say that this culture knows how to add.

This approach can be (rather) easily generalized. Take any classical mathematical theory T. As the set of sentences is countable, it is always possible to write down a listing of all sentences. The only tricky part is to define a relation of neighbourhood, that in some sense or other can be connected to either empirical data or concrete practices. From a formal point of view, one could argue that any mathematical theory can be coded into arithmetic and it is clear that for addition and multiplication such a connection can be established. Of course, the weak point of this argument is that, according to the code used, the corresponding arithmetical statements might be difficult to connect to a practice[6].

However, what this perhaps somewhat bizarre example show, is that it is absolutely not necessary to have a notion of proof (as weak or as strong as one would like to have it) to arrive at a set of true arithmetical statements. What happens here is that local bits and pieces are glued together as time goes on and all that is required is (a) some knowledge of the local bits and pieces (but this can be learnt quite easily in an empirical fashion) and (b) a "desire" for consistency (i. e., when local bits are glued together, make the least amount of changes to what you already have "for sure").

Compared to Bloor's examples, I claim that this type of mathematics truly deserves the label 'alternative'.

2.3. *Open or non-compact mathematics: infinitesimals once more*[7]

In this paragraph I will outline how one can have "genuine" infinitesimals on condition that one is willing to accept the following:

(a) in terms of models, only local models (in a sense to be specified in what follows) are considered, or, alternatively, there are no global models[8],
(b) all local models are essentially finite.

I realize that these conditions run counter to everything that is cherished by logicians, mathematicians and philosophers, but I *am* looking for *real* alternatives.

If, however, one is willing to make these "sacrifices", then matters become rather easy, if perhaps a bit tedious. What follows presents a rough outline and not a full-blown theory. Let us start with the standard theory T of real numbers. The first change that has to be made is that two sets of distinct variables will be used:

(i) variables for "standard" real numbers: $x, y, z, \ldots$
(ii) variables for "infinitesimals": $\varepsilon, \varepsilon', \varepsilon'', \ldots$

Suppose we now have a finite series of formulas $\boldsymbol{F} = \{F_1, F_2, .., F_n\}$, all expressed in the language of T. The intuitive idea is that $\boldsymbol{F}$ could, e. g., represent a calculation of the value of a function in a particular point. Further suppose that if all formulas are interpreted in $\boldsymbol{R}$ such that all variables are treated in the same way, then they are true in the standard model of real numbers.

Example:

$\boldsymbol{F} = \{F_1, F_2, F_3, F_4, F_5\}$

F_1: $((x+\varepsilon)^3 - x^3)/\varepsilon = ((x^3 + 3x^2\varepsilon + 3x\varepsilon^2 + \varepsilon^3) - x^3)/\varepsilon$

F_2: $((x^3 + 3x^2\varepsilon + 3x\varepsilon^2 + \varepsilon^3) - x^3)/\varepsilon = (3x^2\varepsilon + 3x\varepsilon^2 + \varepsilon^3)/\varepsilon$

F_3: $(3x^2\varepsilon + 3x\varepsilon^2 + \varepsilon^3)/\varepsilon = 3x^2 + 3x\varepsilon + \varepsilon^2$

F_4: $3x^2 + 3x\varepsilon + \varepsilon^2 = 3x^2 + (3x + \varepsilon).\varepsilon$

F_5: $((x+\varepsilon)^3 - x^3)/\varepsilon = 3x^2 + (3x + \varepsilon).\varepsilon.$

(I consider all the formulas universally quantified both over x and ε, taking into account that $\varepsilon \neq 0$, i. e., every F_i is preceded by $(\forall x)(\forall \varepsilon)(\varepsilon \neq 0 \supset \ldots))$

Obviously, if $\boldsymbol{F}$ is finite, then so is the number of variables, both "standard" and "infinitesimal," so is the number of constants, and so is the set of terms occurring in the members of $\boldsymbol{F}$.

This set of terms can be split up in different types:

(t1) some terms involve only constants and variables for standard real numbers,
(t2) some terms involve only infinitesimal variables,
(t3) some terms are mixed such that the term consists of the sum of a term of type (t1) and a mixed term,
(t4) some terms are mixed such that the term consists of the product of a term of type (t2) and a mixed term.

I will make one further assumption, namely, that, although (t1) – (t4) do not exhaust the set of terms, yet, any term can be transformed into one of these categories[9]. Thus, e. g., if the term is $(x + \varepsilon).\varepsilon + \varepsilon.(x - \varepsilon)$, it is mixed but neither of type (t3) or type (t4). But the term is easily transformable into $\varepsilon.(x + \varepsilon + x - \varepsilon)$, and thus of type (t4).

Example (continued): The terms occurring in the calculation are:

(t1) $x, x^2, x^3, 3, 3x, 3x^2$,

(t2) $\varepsilon, \varepsilon^2, \varepsilon^3$,

(t3) $x + \varepsilon$, $(x + \varepsilon)^3 - x^3$, $x^3 + [3x^2\varepsilon] + [3x\varepsilon^2] + [\varepsilon^3]$ (where the square brackets mean that the bracketed term is either present or not, at least one term being present),

(t4) $3x^2\varepsilon$, $3x\varepsilon^2$, $[3x^2\varepsilon] + [3x\varepsilon^2] + [\varepsilon^3]$ (as this term is the same as $([3x^2] + [3x\varepsilon] + [\varepsilon^2]).\varepsilon$, the brackets have the same meaning as above).

Finally, we arrive at the interpretation of the formulae F_i. Again, the procedure here is rather unorthodox. There are several stages that have to be executed consecutively. Throughout, *Int* is an interpretation function that interprets the variables, constants, terms and formulae in the ordinary real number model. The resulting model, if it exists, will be called a *local model.*

(S1) Let *Int* fix the values of the standard variables and of the constants. This also implies that all terms of type (t1) are thereby fixed, as we follow standard procedures, i. e., $Int(t_1 + t_2) = Int(t_1) \oplus Int(t_2)$, where $\oplus$ is addition in the real number system, likewise for multiplication.

(S2) Consider the following set $Dist = \{|\text{Int}(t_1) - Int(t_2)|, |Int(t_3)| \mid t_1, t_2$ are terms of type (t1) and t_3 is the term that has the smallest non-zero absolute value$\}$. In short the set Dist looks at all the differences between the standard numbers in order to determine a lower limit, which is why t_3 has to be taken into account[10].

(S3) Let δ be the smallest non-zero element of *Dist*. Take a number $\delta' \ll \delta$. Consider all terms of type (t_2) and type (t_4). Choose $Int(\varepsilon)$ in such a way, that, for all those terms t, $|Int(t)| < \delta'$. As both sets of terms are finite, this is always possible. For terms of type (t2), this is obvious and for terms of type (t4), note that it is a product of a pure infinitesimal term and a mixed term.

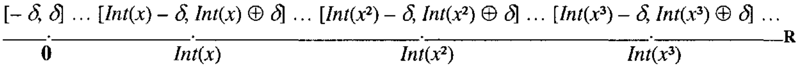
(S4) All remaining terms can now be interpreted in the usual way.

The formulas can be evaluated according to standard principles, e. g., if v is a valuation function based on Int, then $v(t = t') = 1$ iff $\langle Int(t), Int(t')\rangle \in Int(=)$. Do note that the clause for the universal quantifier is restricted to local models and not to all standard real number models.

Example (continued):

Suppose that $Int(x) = \mathbf{2}$. Then terms of type (t1) are evaluated as: $Int(x^2) = \mathbf{4}$, $Int(x^3) = \mathbf{8}$, $Int(3) = \mathbf{3}$, $Int(3x) = \mathbf{6}$, $Int(3x^2) = \mathbf{24}$. The minimum distance is $\mathbf{1}$, so take, e. g., $\delta' = \mathbf{0.001}$. The largest term of type (t4) we can encounter is $3x^2\varepsilon + 3x\varepsilon^2 + \varepsilon^3$. A direct calculation shows that $Int(\varepsilon)$ = **0.00001** will do the job.

Graphically, what is being proposed here, produces the following picture. The line represents the real numbers **R**, where only a finite number of elements are interpreted, such as x, ε, and so on.



Furthermore one has to make sure that in the interval $[-\delta, \delta]$ one finds the interpretations of all infinitesimal expressions, such as, ε, ε^2, ε^3, ... and that for any term t of type (t1), all expressions $t + \varepsilon.t'$ (where t' is any term) are interpreted in the interval $[Int(t) - \delta, Int(t) \oplus \delta]$. Because we deal with a finite number of statements, this procedure can always be executed.

On the one hand, it is obvious that I treat infinitesimals, such as ε in the example, as an ordinary real number and, semantically, it is in fact interpreted as such. But, on the other hand, it is also the case that, in every local model of a set of formulae $\boldsymbol{F}$, for any "standard" variable x and for any infinitesimal variable ε, $x \neq \varepsilon$, thus expressing that no "standard" number equals an infinitesimal. Likewise, for all the constants n named in the set $\boldsymbol{F}$, $n \neq \varepsilon$. Thus infinitesimals are at the same time different. Finally, let me mention that this approach is quite distinct from such theories as Robinson's non-standard analysis or synthetic differential geometry[11] (which, as said before, I would consider to be "mild" alternatives, if at all).

3. THE IMPORTANCE OF PRACTICES

Perhaps the reader is wondering whether this contribution has been published in the right book, viz. a *Festschrift* for Michael Otte. Should this paper not belong in a similar book dedicated to David Bloor? My answer is no. Bloor was the starting point I needed to develop the idea of (forms of) a "truly" alternative mathematics (compared to the "mild" forms). Granted that (if the reader is so willing) such "real" alternatives are genuine possibilities, it is important to ask the next question. What does it tell us about the ontology of mathematics? This, no doubt, has been and still is a major concern in Otte's thinking about mathematics and philosophy. I have always been and still am very sympathetic to the idea that mathematics is an activity and that the task is to show how ontological considerations flow from this activity. To let Otte and co-author Panza speak for themselves:

> From our point of view, mathematics (like science in general) has to be understood as a human activity, namely the activity of producing mathematical (...) theories (...). The aim of logic is not merely to study the internal structure of such theories, or even the formal nature of their propositions (...). Rather it implies the study of the modalities of human activity that produces them. Such an activity is a concrete and historical phenomenon. It is in terms of this phenomenon only that, we believe, it becomes possible to explain all other phenomena or entities. (Otte & Panza 1997, 269-270)

Referring to the third example I presented above, I would like to make a modest suggestion[12]. If we are asked what are the basic constituents of mathematical practice in its most reduced form, then my answer would be: a triple $<m, s, P>$, where m is a particular mathematician, s a mathematical statement, and P a finite set of formulae (interpreted as a proof of s or a calculation leading to s). If one is now willing to make a "double abstraction" – exclude the first element: who wrote the proof is of no importance, and exclude the third element: proofs only help us to see what statements are true – then we are left with a set s of statements, that might carry the label of a mathematical theory T. Then, and only then, can questions be asked about the ontology of T, what it is about. As the other way round does not work, i. e. to "undo" the double abstraction, starting from T, it becomes obvious that practice generates ontology and not vice versa. It is a kind of "simple-minded" argument in support of Otte's views. But as we all know, such arguments often have a high rhetorical impact. Let this be the core of my contribution to this *Festschrift*.

Vrije Universiteit Brussel, Centrum voor Logica en Wetenschapsfilosofie, Universiteit Gent

NOTES

[1] To be precise, the chapter that follows the chapter being discussed here presents a fifth case, but that case is tied strongly to the necessity and certainty of the underlying logic and I am not addressing that problem here.

[2] In the second edition of Bloor's book, there is an afterword where Bloor replies to his critics (163-185). The paragraph entitled "Mathematics and the Realm of Necessity" (179-183) discusses mainly this problem of the necessity of mathematics.

[3] This chapter is a summary of Van Bendegem (2000).

[4] An English translation is to be found in Newman (1956), vol. I, 420-429.

[5] The other source of inspiration is the theory of cellular automata and, especially, the result that shows that Turing machines can be easily translated into such automata.

[6] It is not that difficult to see that many problems will arise (that however do not really challenge the basic idea). The example given here concerns statements involving only constants. But, if the language is to be more expressive, quantifiers and variables should be included. What then happens if someone asks the question: "Is it the case for all x and y, $x + y = y + x$?" Several approaches are possible: if all the specific cases on the list agree with the general statement, then answer yes; if not, answer no, and present the counterexample. In the former case, one risks to give wrong answers. However in due time these will be corrected. This makes clear that classical logic will not be the underlying implicitly given logic, but rather (some kind of) a non-monotonic or default logic.

[7] This presentation is a summary of my (2002). In that paper the focus is more on the inconsistent behaviour of infinitesimals, whereas here I am more interested in the non-compactness.

[8] Yet another formulation is that one should be willing to give up compactness.

[9] A simple proof by induction on the length of terms will do. Suppose that all terms of length n are of type (t1) – (t4). A term t of length n+1 is either of the form $t' + t''$or $t.t''$ As both t' and $t''$are shorter than n, they are of type (t1) – (t4). It is now sufficient to check all possible cases to see that t is of either one of the four types. E. g., suppose t' is (t3) and t'' is (t4), then t' is of the form $t_1 + m$, where t_1 is of type (t1) and m is mixed, and t'' is of the form t_2, m', where t_2 is of type (t2). But then $t = t' + t'' = t1 + m + t_2, m' = t_1 + m''$, where m'' is obviously a mixed term.

[10] A straightforward example: suppose F consists of one formula, viz., $1 + \varepsilon = 3$. $Int(1) = 1$ and $Int(3) = 3$, so the minimum distance is 2, but then $Int(\varepsilon) = 1$ is possible, since $1 < 2$, but then $Int(\varepsilon) = Int(1)$, which is precisely what needs to be avoided.

[11] See my (2002) for more details on these approaches.

[12] This suggestion is inspired by Philip Kitcher, who in his (1983, 163-164) proposed to model mathematical practice by a five-tuple <L, M, Q, R, S>, where L is the language, M the set of metamathematical views, Q the set of accepted questions, R the set of accepted reasonings, and S the set of accepted statements.

REFERENCES

Bloor, D. (1991[2]). *Knowledge and Social Imagery*. Chicago: University of Chicago Press (London: RKP, 1976[1]).

Boyer, C. (1959 <1949>). *The History of the Calculus and its Conceptual Development*. New York: Dover.

Hersh, R. (1997). *What is Mathematics, Really?*. London: Jonathan Cape.

Kitcher, P. (1983). *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.

Otte, M. & Panza, M. (1997). Mathematics as an Activity. In M. Otte & M. Panza (Eds.), *Analysis and Synthesis in Mathematics. History and Philosophy*. Dordrecht: Kluwer Academic Publishers, 261-271 (BSPS 196).

Newman, J. R. (Ed.) (1956). *The World of Mathematics, volume I*. London: George Allen and Unwin.

Van Bendegem, J. P. (2000). Alternative Mathematics: The Vague Way. In D. Krause, S. French, & F. A. Doria (Eds.), Festschrift in Honour of Newton C.A. da Costa on the Occasion of his Seventieth Birthday. *Synthese* 125(1-2), 19-31.

Van Bendegem, J. P. (2002). Inconsistencies in the history of mathematics: The case of infinitesimals. In J. Meheus (Ed.), *Inconsistency in Science*. Dordrecht: Kluwer Academic Publishers, 43-57 (Origins: Studies in the Sources of Scientific Creativity, volume 2).

ROLAND FISCHER

AN INTERVIEW WITH MICHAEL OTTE

Alpbach, Tyrol on August 16th, 2002

Alpbach is a very nice village in the Tyrolean mountains, about 100 metres above sea-level, surrounded by meadows and woods. It is a good place for walking, climbing (and skiing in winter). It is also a good place for thinking and dialogue. Each summer since 1945 the "European Forum Alpbach" takes place, an event of discussions between scientists, politicians, businessmen, journalists etc. Michael Otte was in Alpbach several times in the last 3 years and we discussed questions of mathematics, education, politics and many other fields.

Fischer: Michael, was there any precise point in time of which you remember having decided to become a scientist?

Otte: No, there was no such point in time, and this due to a certain kind of schizophrenia (which lies in my own person). On the one side, and this is also one of the reasons why we are here in Alpbach, I have always been loyal to some kind of peasant realism, that is I have always felt the necessity, even as a child, to earn my own money, and the values bound to arise from that have been strongly formative for me. On the other side, *ideas* have enormously fascinated me since I was a child. My focus of interest on mathematics and physics developed because this is precisely where *thinking can be observed at its closest and most transparent.* I am, for instance, no natural scientist in the proper sense, as would have been suggestive, coming as I did from the country, from the mountains, where agriculture was then absolutely non-mechanized, quite in contrast to today, that's how you develop an interest in such a kind of environment; but I have not become a natural scientist (*Naturwissenschaftler*), I am now rather more what the English formerly called "a naturalist." There is almost a dispute today in the biology departments between those who carry out field expeditions, and those who are sitting in front of computers doing gene analysis. In English, you have two different terms for this; naturalist vs. natural scientist. But, as has been said, how man thinks, and why he thinks, that's what has always interested me in the first place, and that's what has continued quite steadily and has led me towards leading

M. H. G. Hoffmann, J. Lenhard, F. Seeger (Eds.), Activity and Sign – Grounding Mathematics Education, 361 – 378.

some kind of double life. Even a "naturalist" will measure and classify, while at the same time seeking to experience nature.

Fischer: And why did you choose this quite unusual combination of studying mathematics on the one hand, and German language on the other?

Otte: Because for me, poets are people quite similar to mathematicians. While for me everything is subsumed under the category of the *image*. On the one hand, I am familiar with a convoluted reality, a complex process of calculation whose end cannot be anticipated – I have always liked to do numbers – it is at the same time always somehow weird. I also like mountain meadows better than the forests there because you can't look about yourself in the woods, and this is something you can stand but for a while. The basic need people formerly shared was that of an ordered vision, a clear form. The evaporation of general truth is typical for the crisis of modernity which has been brought about by the scientific revolution of modern times and by everything it entailed. Descartes was this revolution's first champion. He deemed mathematics important because it exercised your vision, and not because its theorems or truths were so interesting. His concern was the perception of truth which was to be enhanced by extreme concentration and appropriate means. Rationalist science would seem to be nothing but the self-scrutiny of human thought projected to the external world.

Fischer: That's how it looks in retrospect. Did you see that you liked this as you now express it quite as clearly that time already?

Otte: I never really liked something on my own, but *was rather driven* to it, I have always been on the defensive, that's how it is. I have always felt that I have to master some kind of survival struggle while at the same time having been aware of the freedom of imagination. In the village, you were isolated because you belonged among the have-nots, and later you were isolated because you came from a village and were not the same as the others. That was quite intimidating. I wanted to study history with Franz Schnabel, but with him you did not have any chance to apply for participation in one of his seminars if you did not wear a tie. I did not quite have impressive social skills either. Nevertheless, I have pursued the same goals, and the same ideas as well, over a long time, actually all my life. I realize this over and again whenever my entire structure of ideas all of a sudden shifts. Such a shift generates a meta-view. Before, you were firmly stuck in the deep, impenetrable forest of meanings, and all at once this became a clear form and quite flat. In mathematical thought, this process of development is very pronounced. Ever again, one is led, after long experimentation, to exclaim: Ach, it is that simple! There is nothing behind it which would involve the soul. *X* means no more than *X*! Mathematics, and the arts as well, are characterized by the fact that the ideas only become apparent in realization or application.

Fischer: What then, were your goals, your ideas which you have always pursued? Or, let's just begin with one of these.

Otte: Well, if you are interested in thinking, I'll say it this way: Thinking should be understood as a kind of perception. And in perception, there is always a mix of interpretative, structuring, and receptive moments. That is, you don't perceive reality in itself, but all this is actually a recursively intertwined process between what is your own activity of interpretation and that which suggests itself. While I cannot say that everything I perceive is only my own fiction, – that is you, as you are sitting before me, will not disappear even if I blinked my eyes most forcefully, – but an "objective" description of what I perceive seems just as impossible on the other hand. Reality has a certain resistance and stubbornness, things are active, not as strongly as human beings, but certainly somehow as well, and this means in the last instance that they are something like basic units. While I can analytically take apart such a basic unit, like some kind of black box, the real process disappears by this activity. In perception, besides the indication that something exists, there is also what I project into it, or the categories I use to make a judgment of perception whenever I interpret a perceptual impression, and I will never be able to acutely distinguish between what comes from outside and what is being added from within. The same thing is repeated for every other activity, not only for the activity of perception. Within a process of activity, these complementary elements will always be in action, either the side of means and problems, or the object side. If I considered this in a static way, I would always get entangled in paradoxes. Proof is a means to explain reality, but at the same time, proof is the goal of my cognition, its object. When a student asks what a certain concept means, say, *vector*, then I'll say well, lets try and begin to develop a linear algebra, and in the end we shall know what a vector is. On the other hand, the student will not be able to participate in the process of developing a linear algebra if he/she doesn't already have an intuitive idea what this is all about.

Fischer: Your concern in this context is, and has ever been, to convey such insights to people. Or did you want to realize it yourself in different concretizations, or what is your concern now?

Otte: My concern, properly speaking, or what I always have believed is that *reasoning is a form of life*, and actually one that is just as essential as eating and drinking. And only after I had come to Brazil, for instance, did I realize that in a society as particularized as the Brazilian, epistemology is something eminently political. One of the fundamental problems of *Brazilian society*, for instance, is the weakness of all the institutions, resulting in the fact that one is compelled to recur, in the last instance, to life within the family and the "tribe," this weakness again having to do with the fact that the general, in this case the societal general, is not felt to be something real. In Brazil, people think rather more like Margaret Thatcher who said "I don't know any society, I know only individuals." In this sense, I think that *epistemology is naturally a means to enhance one's personal life, and to solve problems as well.* However strongly single individuals strive for self-realization, they would at the same time like to become "more general," and to contribute to some general ideas and works. When I happened to visit the church here in Alpbach, I say a woman kneeling in one of the benches. On a small piece of paper, she wrote: "Lord, I beg, protect me from loneliness." Here again, we sense that there is no individual existence, and that generalization and connection are the most important things in life. This connection is then fed from the sensation of immediate familiarity with others, or it is what we have in common in our thinking and in our ideas.

Coming back to didactics in the narrower sense, saying that concepts are important means of reasoning, we have to know at some point what kind of entities concepts actually are, in the same sense as I would ask what chairs are, or what apples are, that is in principle, to what category of our reality something like that belongs. This means that we must connect the two levels of the general just named, the immediate, the common on the one hand, and the structural order of ideas on the other.

Fischer: Is this to mean that one of your motives was to tell the others that that's how it is?

Otte: Where the telling is concerned, I have missed a lot, because I wanted to understand myself first, and that takes some time, actually. This is why I have always been more comfortable with the people I was able to work together for a long time, with candidates for doctoral degrees, or in work groups on projects.

Fischer: You seem to have offered the solution to the problem already which consists, for instance, in pointing out these complementarities. But what was before? When you say that you only realized that this was the case, then there must have been some pressing problems before.

Otte: I grew up in very poor times, 1945 – 47 in mountain Bavaria, where there was a social wilderness including the struggle of everybody against everybody about everything, and in particular of the native farmers against the refugees who had come from the east. For me, ideas are *part of a genuine humanism.* People should contribute to a larger connectedness however important it remained in every moment to get hold of the money and of the things essential for life. To teach people to think, or to help him, respectively myself, to attain a clear view of themselves and of the world, that is for me in a certain sense a humanist concern. It is even the kind of help other people can offer – perhaps therapists, or a medical doctor, – each of these in their own way. On the other hand, I have never grown completely beyond feeling the pressure of a struggle of life determined by economics, and I have hence always felt bashful when confronted with others, a thing which is not ideal for an educator or teacher (in this respect, you have always been a model for me).

Fischer: Before you came to Bielefeld, where you clearly turned to didactics on the outside, what was there, how did that what you now have sketched as a fundamental problem become concrete at that time?

Otte: I had come to didactics before that, from a practical reason. I obtained my PhD. in Münster in pure mathematics, and Münster was a very large university. They had done a survey when people began to reflect on the number of semesters spent in studies and the like. One of the findings was that there were an incredible number of students at the university, I believe more than three hundred *students of mathematics in their twelfth semester or more without having had any examination and no topic for their paper*. I wrote a letter to them all, and quite a number responded, almost 80, 90. These I then organized into groups of three to five people, than giving projects into these groups, and these projects had to be something altogether different. It made no sense to me to offer them just some bit more new subject matter, they had been sitting there for six years already, having heard lots of things going into their one ear, and out by the other. These were projects where a subject matter which one was supposed to know was taken up again from another perspective in a new and problem-oriented way. Elementary mathematics from a higher point of view, this was actually the task. In any case a subject matter they had already had, but now to be worked upon from a new perspective and in a novel way, in cooperation and with a clear goal in mind. And I believe I made about more than sixty people from this group of people, who were to all intent and purpose dropouts, successfully pass their examinations. I thus did not begin my work in didactics theoretically-scientifically, but rather practically and with the conviction that there are wonderful things in mathematics which everybody should know. Moreover, I was a good organizer, if left alone, and if there was not such a lot of political and socio-psychological debate going on. The rector of the University of Bielefeld got to know all this, and that the reason proper why I came to Bielefeld.

Fischer: This means that your approach was one concerning the didactics of higher education one, starting from a problem with which you were confronted in teaching.

Otte: Precisely, yes, that was quite practical. At that time, however, I had already begun to hold additional seminars in Münster, for instance about Piaget. I used the book "*Mathematical Psychology and Epistemology*" written by Piaget jointly with E. Beth.

Fischer: And that's how you came to concerning yourself with didactics theoretically as well.

Otte: That has always come natural to me. It is simply impossible for me to do something without having general thoughts about it, well perhaps with the sole exception of painting. In all the years when I worked in farming and in handcraft, this also impeded me.

Fischer: When you came to Bielefeld then, you were also confronted with a didactics of mathematics already existing in Germany. What was your idea of that then, and how did you see yourself in relation to what was there?

Otte: I did not acknowledge the didactics that was there already. That had to do partly with the arrogance of the mathematician with regard to these theories of method: "Either one masters mathematics, or one becomes a mathematics educator." I don't want to enlarge upon all these prejudices here again. That's the one side, and the other side was that I felt rather constrained between two colleagues who were relatively prominent in German mathematics education, the two of them being both more than ten years older than me, and thus having already occupied certain fields for themselves: Bauersfeld curriculum theory and primary school mathematics education, Steiner the didactics of secondary school and of the Gymnasium, that is to mean what is the didactics of the subject matter in the classical sense, and beyond that the corresponding theories of didactics. And properly speaking, I strove for something quite different, *to insert what was offered in the shape of classical subject matter didactics in mathematics somewhat within an interdisciplinary context.* I always told the mathematicians: mathematics education, properly speaking, represents the objective historical interests of mathematics as a science. I was convinced from the very outset that mathematics education is always an institution concerned with teaching, also in the sense that it must take note of what is being achieved in its most diverse *sciences it relates to* (theory of cognition, psychology, philosophy, sociology, of course mathematics as well) and try to implement this in a kind of "*applied basic science*." Nobody was really willing to do that, the term was held to be some kind of invective, the intention being called "sociologism" in the IDM's advisory body. The only institutional model proper was the Max-Planck Centre for Educational Research. That time, of course, was a difficult period politically as well, the era of the Cold War and of the Berufsverbot.

Fischer: What were the first projects you initiated in this sense at the IDM?

Otte: The first project I had already begun in Münster, even before I had taken up work at the IDM, because I had reflected upon what I could do at all. Something must be said on this: to begin with, we had rather a lot of free positions at the IDM, for instance 16 positions for delegated teachers, and at least fourteen, fifteen positions more for assistant professors. And these were virtually impossible to fill, for the didactics of mathematics itself had not yet produced any academic rising generation. Even established mathematics educators who otherwise did not give me the lightest credit recognized that I have been the one person who initiated workshop projects for doctoral theses in the sense of establishing a rising generation. The doctoral theses workshop at the IDM was important in this frame. And the second project which I started in Münster already had to do with my – somewhat naïve intention – to establish a good connection to the mathematicians as well. It was a volume published by Springer: "Mathematiker über die Mathematik" (mathematicians on mathematics). Today, books of this kind are nothing remarkable. But at the time, this was quite unusual, and the publishing house of Springer, conservative and successful as it is, hesitated for a long time whether they would edit something like that, but then created a series which saw two more volumes. The series was called "Wissenschaft und Öffentlichkeit" (science and the public). Only at the time you had to deliver a complete printable typescript; in a sense, this was a typed book. Half of it consisted of contributions collected and translated from the English. I travelled, for instance to Moscow to see Kolmogorov about it. The other half was from people I had convinced to contribute original German essays, such as Jürgen Brieskorn in Bonn, Andreas Dress in Bielefeld, or Hermann Dinges in Frankfurt. This volume, I have heard, has later been used over years and decades in seminars, of philosophy as well. It has been selling for 30 years. And that was the very first project.

Fischer: The workshop projects for doctoral theses, what was their contents? What were the themes you gave there?

Otte: When I remember and look at the first themes, these were things which had to do with mathematics itself, and which were intended to open new perspectives on mathematics. Jahnke, for instance, was to work on the problem of proof, using the previous work of Sneed and Stegmüller in structuralist theory of science. These are concerned with empirical theory, and there is so-to-say from the outset this *duality between a structural nucleus and intended applications*. The theory itself was to be seen only as a dual edifice, this is simply the more reasonable and natural point of view, and I reasoned that this must work in mathematics as well. If I had been more seasoned at the time, and if I had known Abraham Robinson's work, that is the work of scholars who from the very outset had extended the purely proof-theoretical orientation of mathematical foundational research by a model theoretical component in the thirties already (Abraham also was an applied mathematician who worked in the airplane industry in World War II), this would certainly have been better. Our approach at the time was a bit difficult to access and problematical. We tried to legitimize it mainly epistemologically.

Other results concerned applications, and I would like to quote two examples, the doctoral theses submitted by Biehler and Steinbring. Steinbring's thesis was on probability theory, that is on the history of probability theory, continuing Hacking's books of the period – among them Hackings own doctoral thesis. Biehler continued Tukey. What I liked in this was that *explorative data analysis* from its very outset is about quite another type of knowledge. That is, different from the hermetical character of the Newman-Pearson theories where theorems of pure mathematics can still be used, a mode not quite overcome, but supplemented by a conglomeration of methods of analysis and strategic rules. What was lacking at the time was the semiotic knowledge, otherwise one would have seen more: it is always a matter of interpretation and of a new elaboration. What is interpreted is not only reality, but all the representations – that is *stem-and-leaf* and however all these methods are called – become an object themselves.

Somehow, things were as a rule about mathematics itself, and about robbing mathematics of its alien character which it acquired so-to-say inadvertently by the predominance of a linguistic culture, and by certain forms of pure mathematics itself. I recall a debate, however, one of the first I had with you. Your critique of me was that I was always embellishing mathematics, somehow trying to make it look less obtuse than it really is, while you saw that situation always more relentlessly and realistically. I have later often had occasion to think back to this.

Fischer: Well, I later formulated it to say that you were trying to overcome this narrow view from within mathematics, while I was trying to do the same from without. This is then rather more a question of definition, of what you are willing to count among mathematics.

I recall that at that time there was still the EPAS project. How do you see that in retrospect?

Otte: The EPAS project was formed on the basis of a different belief. I have already said that I thought that to teach someone to think, or also to convey knowledge about cognition and knowledge to him or her should be some kind of humanist concern. What I did not like in traditional didactics, in particular in subject matter didactics, was how the teachers are kept in leading strings. This began with the textbooks, the textbooks being not textbooks at all for the pupils, but rather for the teachers. And the teacher, vice verso, refrained from reading anything else; that is he had the most diverse school textbooks in his cupboard, trying to distil recipes from these for his next lesson to be held.

Consequently, the idea was the following: As a scientist, you cannot act as if you knew more about teaching than the teacher, and this is why we externals must in my opinion refrain from trying to act directly on the pupils, somehow circumventing the teacher. The consequence from this is that you must focus on educating the teachers. And in order to focus on teacher education and training, you must cooperate with those people who have a most supportive role in professional teacher training during the first two trainee years in teaching service preceding the second state examination. In the university studies of mathematics, a maximum of 8%

percent of the curriculum until the first state, purely academic state examination was then assigned to the didactics of the subject matter. The professional training proper took place during the two-year traineeship. And there were the teacher seminars, where *the directors (Fachleiter) assigned to the various subjects were the true trainers.* The idea was to cooperate with these people. While they don't have the teacher's problem in the same extent as the teachers themselves, that is they are not in over their heads, are compelled to talk about teaching, must teach others how to teach, and insofar have certain problems to reflect on the one hand, they are practitioners, are involved in teaching, know how school looks like from within etc. on the other. Hence, we applied for some funding from third parties, the application was written in 1975, I remember precisely, the secretary had to type it even on Easter Monday in order to ensure that it arrived in time at the Volkswagen Foundation. We won over three federal states, North-Rhine Westphalia, Berlin, and Schleswig-Holstein, to participate in this project and for delegating teacher seminar directors. I visited everywhere, in the classrooms as well as in the teacher training seminars from Kiel to Wuppertal. It was still easier then, North-Rhine Westphalia, the biggest federal state, did not even have an institute for school and in-service training then which was later founded in Soest, and the ministry was much more willing to accept offers to cooperate from the university. I don't remember the name of the ministry official, but above all there were some outstanding seminary directors. We wrote texts on central professional problem fields, these were then distributed to teachers, in their tentative version, and there followed an annual conference on experience had with them, for an entire week in Bielefeld. Everything was extensively discussed on this occasion, and we attempted to integrate suggestions made. Later, the texts were published in several volumes.

Fischer: Were you as content with this, this kind of practical activity, as with the doctoral theses?

Otte: I was content on the one hand, because there were really quite a number of very outstanding persons among the seminar directors. If all the participants seriously understand their own craft, and show mutual respect, something good will result. What was actually missing with me, but also for the university in general, was a scientific preparation. Let's take, for example, the volume "Text Wissen Tätigkeit" which we authored at the time. In the entire world, we were unable to find a text analysis system oriented towards mathematics of which we could have said, well, let's take this and analyze the mathematics textbooks on this basis and see what comes out. There were only content-indifferent systems of the kind developed by psychology. One of the famous among these in Germany was then that of the Hamburg psychologists Thun and Götz. The inevitable outcome is that while the text is improved and reads quite beautifully, the mathematical object has simply disappeared. With regard to some problems, we simply bit off more than we could swallow. I think this has persecuted me anyway for my whole life: I have always tackled problems which were much too difficult.

Fischer: Another undertaking I remember in which I participated some time myself was the BACOMET project in which you actually attained a significant achievement in organizing science within an international context. Does it still exist at all?

Otte: While the BACOMET project has not been pronounced dead yet officially, it as been put a bit to sleep since we are no longer organizing the annual meetings from the IDM. This is also due to the fact that nowadays the DFG and other funding organizations do no longer provide any resources for this kind of educational research. This seems to be different in Austria, as can be seen from your own work.

Fischer: These meetings continued until two years ago?

Otte: Yes, they did, and the conferences were for a long time held in Melle, in the vicinity of Bielefeld. To get funding for that is even more difficult in England and in Italy, where we tried, and has hitherto failed. In 2003, another volume of the project will be published. The BACOMET project evolved from something else, actually from organizing the Karlsruhe Congress of 1976. This Third International Congress on Mathematics Education was to have a quite different structure, one aligned to themes. This was prepared by 13 expert reports. I had volunteered for the report *on teacher education and training* and held a preliminary conference with persons whom I had chosen myself or who had been designated by IMUK the year before in Bielefeld in order to prepare the report. This pleased people like the then UNESCO representative Bent Christiansen from Copenhagen, and he subsequently asked me whether I would like to continue something similar. This is how the BACOMET project came into being. The intention was to invite institutions – we also invited the Klagenfurt team as a whole – in order to have some continuity of project work in the time between congresses by having several members in the same place. This was an overambitious structure which soon disintegrated because not everybody was interested with equal intensity. Nevertheless, BACOMET provided a program of orientation for many, for almost 20 years.

Fischer: This is then something like a red thread: in the case of the doctoral theses the candidates, then the seminar directors in the EPAS project, the colleagues from the different nations in the BACOMAT project: All these were at the same time *projects of education for a new generation of educators.*

Otte: That's correct, according to the motto: specialization and cooperation will yield the first steps forward from the stone-age. The academic discipline of mathematics education was called to create itself in a first step, actually. But it also had to do with how I worked, with some deficit in myself. I do not write well. Or better, I write in such a way that most people find it too condensed. In the English, this is even worse as easy readability has an even more important role. "How can such a bright person write such boring things"? Or they say something like that. Formerly, I was no very well able to hold a lecture in front of people I did not know. But I think that I am good if I can work together with the same group for a longer period of time. BACOMET was a big success, everybody liked to come up to the close.

Fischer: How many people came usually?

Otte: Always between 21 and 23 people.

Fischer: An international class of school.
Another line in your work is *history, the history of science, the history of mathematics*. That was one aspects you always had in mind as well, but I had the impression that this became particularly intense in the eighties where you got onto the historical track with some of your co-workers as well.

Otte: That's right, actually. For me, the formation of the new educational generation had a certain priority, and doctoral thesis projects had to find recognition in the established sciences as well. I have been working at a reform university, and in a most ambitious one, in Bielefeld, and I had the good fortune of encountering a very pronounced tradition in the fields of history of science and research into science there. When von Weizsäcker was pensioned, the Max Planck Institute for researching the condition of life in the scientific civilization situated in Starnberg was dissolved. Because of this, some additional excellent people such as Wolfgang Krohn and Wolfgang van den Dale came to Bielefeld because of a contract between the Max Planck Society and the University of Bielefeld.
As for myself, I've never been a historian quite simply. While I have myself worked historically, about Hermann Grassmann (1809 – 1878) quite intensely, I have never quite liked to do history in a purely descriptive or narrative way. This has then led to some disagreements in our work group, because of the complicated individual problems of orientation within a field of work of such interdisciplinary dimensions. And it has led to some splits, three different directions being established in Bielefeld: the purely historical one represented by Schubring, my own, which is the most philosophical one, and Jahnke in between. I was also co-applicant for post graduate college for research into science and technology, its present name is completely different. I have been with that from the very beginnings, it consisted of twelve professors if different faculties and departments of the University of Bielefeld who cooperated to write this application.

Fischer: Over which period did that take place, do you recall that?

Otte: This, the post-graduate college?

Fischer: Yes, this cooperation with colleagues.

Otte: That certainly kept going for 10 or 15 years, perhaps even longer.

Fischer: Well, you can perhaps quote an example to describe what this focus on the history of mathematics may still contribute, from your point of view, for your other concerns, that is for didactical concerns, the role of mathematics in society?

Otte: At present, I have begun to concern myself with the history of geometry. In Brazil, I have begun to work with *computer geometry*. For the students in Brazil, I had to look for something immediately relevant for my lectures, linking mathematical subject matter, or mathematical knowledge immediately with historical and philosophical backgrounds. The computer has generated so many illusions about thinking. And as I generally assume that mathematics is an activity, and that activity represents a system of relations between means and objects, thinking cannot remain unconcerned if a new means like the computer is placed at the disposal of geometrical activity. And when I began with this, it was at once visible that a very strong element has become relevant again which had an important role in mathematics until the end of the 18^{th} century, but then fell into oblivion, namely *the principle of continuity*. In natural philosophy, too, it had played an important role, just think of the presentation in Lovejoy's famous volume "The Great Chain of Being." The fascination in the principle of continuity is that it is not simply a principle, but rather it are the principle's changes from which the historical stage of development of mathematical and natural science thought can be discerned. I think the mathematicians have erected some kind of interpretative dogmatism there. They say, for instance, what about it, this simply means the theorem of identity for homologous functions, or this should be expressed in the terms of Zariski topology of algebraic varieties. This is nonsense. I think the principle of continuity is so fascinating because it indicates, so-to-say, the emergence of the theoretical view from perception, from crude empirics. It is situated at the point of intersection between the empirical and the conceptual which has occupied reasoning for centuries. It may even show us how the good theories about the essence of the relationships in nature have arisen.

Fischer: Could you try to formulate this principle?

Otte: Properly spoken, the principle of continuity is the following: The world is infinitely complex, it changes every moment, and if we desire to recognize anything, we must look for *invariants or similarities*. And in this connection, it is quite interesting to see how this is being done empirically in the first place, that is how a purely phenomenal concept of similarity is taken as the basis, and how this concept of similarity changes. Galileo, for instance, took the catenary and the parabola to be the same. If you hang a chain on the wall and then project the oblique shadow of a circle onto a parabola, you actually see that they are quite similar, at least in the vicinity of the vertex. It was Huygens who discovered this "error." What does that mean, he discovered an error? He used another principle of relationship. Since then, we look at a similarity of genus which is based on the algebraic-analytical mode of generation. Parabola and catenary are empirically so similar that they were still taken to mean the same by Galileo. On the other hand, while circle, ellipse, and parabola are similar in genus, or in family, they are quite dissimilar empirically. In analytic geometry, they are nevertheless considered to belong together. The new conception of the principle of continuity marks an essential element of the Scientific Revolution of the 17^{th} to 19^{th} centuries. During that period, thought and intuition began to transcend from the things to the laws which determine them. These laws, or relational structures, however, also determine the thinking in analogies or metaphors, that is they provide the aspect under which thinking approaches the existing world. Using a metaphor means to find the basis of relational similarity in the first place. More important than empirical similarity in this case here is the generic similarity or of family likeness, which implies the transition from the purely empirical to the theoretical. And this is precisely what the principle of continuity is conceptually aiming at.

Fischer: My question was how useful history is in this, and now I would like to interpret you answers to say that you actually know quite precisely what you desire to see in history, and that you then succeed in finding that, using it to illustrate your ideas even more clearly, but less that the intention now is to tumble on some things empirically in history.

Otte: I would like to say two things in answering that: Firstly, I think you are a nice fellow, but secondly you insinuate with that, well, it could also be said that I am a falsifier of history.

Fischer: no.

Otte: I would like to say to that: what is the predicament of the philosophy of mathematics? It is hung with two problems it has no hope of solving: One is the problem of the *ontological status of mathematical objects*, or the question whether mathematics has any objects at all, and the second is *how proof as the tautological process which it seems to be can lead to new insights*. I think that in modern mathematics, in contrast to the Aristotelian science, proof not only has a rhetorical function, is not intended to convey truth to somebody, but rather has a constitutive function. And if it has a constitutive function, it presents itself quite differently as a problem. Using the perspective of these problems, I approach history, perceiving much more with it; you see things you simply do not see when analyzing with the logic of mathematics. That's the first thing I wanted to say. And the other thing concerns the limits of the philosophical perspective. History will always yield total surprises, for instance if you take the problem of doubling the cube. The most suggestive thing is to double the side length – the Athenians are said to have been so stupid to do this – but then you obtain the eightfold. And somewhere between the plain one and the eightfold, there must be, according to the principle of continuity, the double volume somewhere. Another great problem of antiquity was treated in the same way, that of squaring the circle. Somewhere between the inscribed square and the circumscribed square, there must be a square which is equal in area to the circle. This, for instance, is a seemingly evident argument, and the surprise is that people like Nicolaus von Kues (1401 – 1464) radically rejected it. That reality is something continuous, the very "Great Chain of Being," appears somehow so suggestive that it was a puzzle for me how somebody cannot accept it. In the 15th century, however, the harmonious world view crumbled, and the particular and the general were no longer so statically compatible. That is to say, you learn from history as well, and in this case you learn which mathematical conceptions can have fundamental cultural-historical impact.

Fischer: One of the particularities of your historical contributions I have read is that they do not confine themselves to the history of science in the narrower sense, like you described it just now quoting examples, but that you establish references to the history of culture, to the economy of the respective period, for instance to technology and history of technology, showing the history of mathematics, or of the natural sciences, embedded into the history of technology, at least since the 19th century.

Otte: I would like to say two things to this: The first is that I had abandoned this kind of work for some time, because it is so difficult to communicate about it. Those people who have a larger sociology of science and history of culture perspective either show no interest at all for mathematics, or merely a very superficial one. If I look at the work I wrote today, on the other hand, I often wonder, dear me, what you've written that twenty years ago, you cannot improve upon even today. There seems to have been no progress with me on the general level during the last twenty years. This is why I prefer working with very precise examples which you use to develop your own perspective, that is where you can see something much more exactly, but cannot see something as large. Nevertheless, I believe heuristically or in orientation, that there was a *fundamental rupture, and this is really the Industrial Revolution.* Kant's philosophy advanced a point, a central point: That we need not harbour the illusion of being wholly transparent, clear and manageable to ourselves, and only the impenetrable mystery of the external world confronted us, as Cartesianism saw things, but that objectivity is subjective, and vice versa. If the world is nothing but a social construction anyway, one must ask: does the constructor really know himself? Or better, this illusion of (social) constructivism which says that we immediately understand what we have done ourselves, for this very reason, simply evaporates. Activity and social practice as fundamental realities, but also as scientific object for thinking and for science, this has been brought along by the Industrial Revolution of the 18^{th} and 19^{th} centuries. In this sense, this is some kind of watershed which you will always use to find your direction.

Fischer: Philosophy and theory of science have always been important to you as well. Has there been a development? You have already mentioned Sneed as a theorist of science, and I know that you often quoted Quine, you have for a long time been occupied with Russell or Peirce, or still are. Has there been a development during the last thirty, forty years?

Otte: Yes, there has been, and this year I have joined the *Deutsche Gesellschaft für Semiotik.* I hesitated for a long time, but joining simply expresses a certain development. I recall there was a time when it was said that you could properly not use Marxist concepts like "activity" or "practice" in (didactical) research. Independent of the polemic stance it is quite clear that analytical philosophy – you mentioned Quine – that is the one which is mainly concerned with language and logic, assumed a paradigm of language ("linguistic turn"), not a *paradigm of activity*. Here again, there is a fundamental insight by Kant which has been continued both by Marxism and by Peirce's pragmatism. Mathematics is not a science of concepts, although there is no explicit mention of object in particular in modern axiomatized mathematics with its attributive use of concepts. But nevertheless mathematics is, as one could state quite banally, in my opinion mainly *a technological and at the same time fine art.* I found semiotics interesting for the reason that it treats this. Language and linguistic symbols belong to the realm of signs, but they are not the only ones. To quote an example: you take a triangle, you take any point, you construct the point symmetrical to the first triangle corner, the point symmetrical to the second triangle corner, the point symmetrical to the third triangle corner, and you take the median of this and the initial point. Because of the way this last point has been introduced here, it is of course dependent on the first point, so that you would expect that if the first point varies, the last one must vary as well. It does not, however. What does this mean? This means that the point, so-to-say, has a determination which is independent of its mode of construction. I have told the students that every fixed point signifies a theorem, "find the theorem, find the proof." And this, these are typical things which I am unable to express in language like that, by speaking of this or of that point. What does this mean, this point? In language, I would have to tell how it has been introduced. It has not been introduced, however. I am still trying to find it. That is to say I suddenly have something of which I claim it is there, but whose properties I am precisely trying to find. This is exactly the path in the opposite direction to that we usually walk when we begin with the definitions.

Fischer: That is, one might consider that you see in *pragmatism* the possibility to overcome those problems which in a –

Otte: linguistic orientation

Fischer: cannot be coped with.

Otte: That's right. What matters is the activity. Kant's philosophy is a philosophy of activity, and who continued it? Porperly speaking, only Marxism and pragmatism, all the others went into the linguistic direction, into the direction of conceptual language –

Fischer: – which then runs in the direction of differentiation, with hierarchies, of types, for instance, or of others, but does not do justice to mathematics.

Otte: Exactly, mathematics, that's what it always is, that is I hypostasize things, that is I take a process again as object of another process, that's what mathematicians do, ever again. Mathematicians always operate within particular model realities, and do not claim to have any insights about the "real" world. Common positivism, against that, identifies knowledge with reality, and then wonders when something unheard of appears (e. g. mad cow disease).

Fischer: In 1994, in the volume editied by Suhrkamp "Das Formale, das Soziale und das Subjektive," you drew a balance with regard to didactics, presenting your considerations in a quite extensive frame. This book is not only about teaching mathematics, but about the role of science in society, and about the role of education, about the teacher as an intellectual. This goes far beyond what one expects of a mathematics educator. How do you see that today, is that balance still correct, or how would you place the accent now?

Otte: Well, today I would make the book shorter. What the point was? Firstly, it was some kind of a collection of what I always wanted to say about didactics. It was a situation when didactics as a science was on the decline. It suffices to look at the situation in Bielefeld. At one point in time, Bielefeld had seven or eight tenures for university professors in mathematics education, including those of the Pädagogische Hochschule training elementary school teachers, and now one of these will remain. At the same time, however, I thought there is a new theory of activity, the complementarity of object and means. The volume you mentioned circles mainly about the idea of complementarity, presenting it in ever new contexts. I consider this idea to be fundamental, and it was very significant, heuristically and philosophically as well, that I found it almost 30 years ago on the occasion of a debate on teaching analysis (see annex).

Fischer: And later, since 1994, what have you been doing mainly?

Otte: I would like to elementarize all that's totally incomprehensible in the technical mode of expression of common philosophy of mathematics in order to show that it is really useful for learning mathematics as well. If I were called to hold a course, and if I had to say what belongs to didactics, I would say that at least *three seminars* or three lectures form the beginning. A first one to discuss the *concept of number* (see annex). Russell's "Introduction to Mathematical Philosophy" would be a suitable object of study in this.

A second lecture would have to treat the *principle of continuity and its changes*. I am presently writing a number of contributions where I demonstrate theorems of elementary geometry in a different way, but not at once from so high a point of vantage. Well, I'll take now the theorem about the Euler line in the triangle, that the circumcentre, the orthocentre, and the line of centres of gravity lie on one straight line. OK with you?

Now you can say that nothing metric occurs there, and hence there must also be a projective proof for that, etc. Nice and good! you can try that anyway, and if you are a genius, more of a genius than a normal high school or teacher student of mathematics, you will certainly find something. But what I've done now is using the fact that I can rapidly draw and deform diagrams on the computer, and on this basis I generalize the theorem step by step. I execute seven steps of deformation until I arrive at Desargues' theorem, or vice versa. Afterwards, I will interpret the new theorem at first glance into the original diagram.

This is sometimes the problem of didactics: either you give the student the result, then he/she will have learnt nothing, or you always tell them "look here, look here." There is a French play by Marcel Aimé in which a teacher of French language called Topaze gives his pupil a dictation containing the sentence "des moutons étaient dans un pré." The pupil ignores the silent plural "*s*" in "mouton*s*," obstinately hearing and writing down the singular form. The teacher, looking over his shoulder, observes this error, and in his exasperation resorts to pronouncing "moutonssses" ever more exaggeratedly, thereby defeating his own pedagogical intention (the so-called "Topaze effect"), and moreover merely confusing the pupil who remains unable to find the solution. This describes the dilemma of the didactical contract which must be continuously broken by the teacher in order to be kept, by the teacher's avoiding the easy way of sparing the pupil's mental effort, or by his keeping the pupil ignorant by refraining from handing him/her the solution pat (Brousseau/Otte 1991, 15).

And the third lecture would be one about the *importance of model theory*. I try, for instance, to do some quite different types of work on the problem of non-standard interpretations, for instance on non-standard analysis. In my opinion, it is not convincing why this epsilon-delta-analysis should be the only way to teach analysis. Learning really works much better of you have alternatives. All these well-established tenets start to shiver once you suddenly see that there are new possibilities even in the infinitely small. Laugwitz has always propagated this here in Germany, but despite all his brilliance he has always done so in a – how do you call it – misanthropic way, you are so uncomfortable reading his texts. I have to agree with him in almost every case, but the atmosphere always seems so bad-tempered. That's understandable, perhaps, in view of the dogmatic traditionalism in science. In any case, there is a lot to learn from this topic, once you drop the belief that mathematics is a science which says how it is. Nothing is how it is.

Fischer: A good closing remark for a halfway balance of an educator's career.

INDEX

—A—

abduction 20, 40, 67, 69, 70, 71, 73, 203, 208, 209
Ach, N. 18, 21, 362
acquisition 33, 67, 172, 181, 184, 192, 195, 204, 205, 230
activity 2, 6, 9-11, 13, 15, 16, 19, 21, 23, 27, 32, 33, 45, 47, 52, 53, 66, 68, 73, 77-79, 86, 89, 105, 118, 119, 121, 123, 130, 132, 137, 138, 142-144, 146, 148, 149, 159, 174, 181-184, 187, 231, 232, 235, 236, 288, 313, 315, 321, 336, 357, 363, 369, 372, 375, 376
Adda, J. 156, 157, 158
Adler, H. 205, 213
Adorno, T. W. 194, 199
aesthetics 203, 211, 213
agency 23, 27-33, 289, 291
Alexander, P. A. 192, 196, 287, 290, 294
Alleau, R. 159, 168
alternative mathematics 349, 350, 351, 357
Ampère, A. M. 275, 276, 278, 282, 284
analysis-synthesis 241
Anderson, M. 194
Anderson, R. C. 194
Andrade, A. M. R. 289, 298
Apel, K.-O. 207, 213, 314, 320, 322
application 3, 10, 12-14, 21, 36, 78, 149, 155, 164, 177, 192, 204, 206, 221, 223-226, 241, 246, 247, 252, 254, 255, 257, 264, 284, 317, 318, 321, 322, 340, 343, 362, 369, 371
appropriation 23, 27-29, 31, 68, 80, 84, 162, 185, 189, 212
Aristoteles 21
arithmetic operations 266, 275
arithmetization 10, 263, 264, 267, 272
Arnold, V. I. 21, 169, 311
Arzarello, F., 146
Atiyah, M. 311
Aveline, C. 168
axiomatics 11, 12, 263-265, 267, 269, 270-272, 305
axiomatization 64, 216, 264-272, 310

—B—

Bach 21
Bachelard, G. 24
Bachtin, M. 77, 79, 89
Bajaj, A. 194
Bakker, A. 55
Balacheff, N. 168
Balzer, W. 331, 333
Barthes, R. 67, 74
Bartolini Bussi, M. G. 77, 79, 81-84, 87-90
Bateson, G. 71, 203, 213
Bauer, L. 93, 104
Baumgarten, A. G. 212, 213
Bäumler, A. 212, 213
Baxter Magolda, M. B. 194
Becker, O. 215, 226
Bendegem, J. P. van 349, 358
Berkeley, G. 35, 38, 40, 42, 43, 110, 237, 266
Beth, E., W. 141, 365
Bézout, Etienne 275-278, 282, 283
Biagioli, M. 149
Bishop, A. 26, 168
Bloch, E. 22
Bloom, B. S. 32
Bloor, D. 349-351, 354, 357, 358
Bochner, S. 53, 55
Boero, P. 77, 89, 90
Bohm, D. 11, 21

Bond, J. A., 214
Bondy, J. A. 59, 66
Boni, M. 89
Booker, G. 156, 157, 158
Borovcnik, M. 104
Bos, H. 281, 282
Bourbaki, N. 2, 79, 265, 292, 304, 305
Bourdieu, P. 203, 213
Bowlby, J. 74
Boyer, C. 351
Brakel, L. A. W., 214
Brandes, H. 205, 206, 214
Brent, J. 314, 322
Bretherton, I. 74
Bromme, R. 14, 22, 191, 198
Brousseau, G. 134, 159, 164, 165, 168, 377
Brousseau, N. 134, 159, 164, 165, 168, 377
Brouwer, L. E. J. 216, 221, 223, 224
Brown, G. S. 175, 178
Brunschvicg, L. 140
Bucci, W. 205, 213
Buehl, M. M. 192, 196
Burks, A. W. 22, 214

—C—

Calvert, C. 194
Cantor, G. 65, 216, 218-220, 263, 299, 307
Carnap, R. 234
Carton, A. S. 190
Cassirer, E. 9, 10, 12, 21, 216, 229, 230, 235
Castonguay, C. 13, 21
Cavaillès, J. 271
Cavell, M. 205, 213
certainty of knowledge 191, 194, 195, 246
Cerulli M. 89
Chandler, M. J. 197
Chandrasekaran, B. 319, 322
Chevallard, Y. 134, 163
Chick, H. 89
Chiurazzi, G. 138, 139
Chomsky, N. 26, 168, 185, 189
Christiansen, B. 153, 158, 370
Church, J. 94, 103, 171
c-knowledge 159, 162, 163, 164, 166, 167
classroom 2, 5, 26, 29-33, 77-80, 82-85, 87, 88, 91, 104, 111, 113, 120, 124, 125, 131, 132, 143, 148, 149, 154, 156, 163
Cobb, G. W. 306, 311
Coffa, A. 229
cognitive science 10, 80, 157, 203, 205, 212
Cohen, H. 229, 230, 257
Cohen, I. B. 229, 230, 257
Cohn-Vossen, S. 54, 55
commens 105, 111, 113, 114
complementarity 4-6, 15, 52, 67, 68, 73, 74, 117, 118, 120, 123, 133, 169, 170, 376
complexity 3, 26, 58, 117, 124, 134, 157, 165, 171
conception of truth 335, 337-342, 344-347
consciousness 39, 40, 46, 48, 69, 78, 119, 169, 170, 171, 173, 177, 182, 184, 203, 204, 206-213, 256
consciousness of society 169, 171
consciousness, self-awareness 207
content 2, 12, 17, 26, 91, 93, 122, 137, 138, 141, 147, 148, 155, 170, 179, 181, 183, 191, 192, 196, 197, 205, 224, 232, 233, 250, 275, 314, 336, 346, 347, 369
content, domain specificity 191, 192, 196
content, epistemological belief 191, 192, 194-199
context of science 215
conventionalisation 23, 28-31
Cooper, M. 168
Corfield, D. 322
Corte, E. de 192
creativity 5, 23, 27, 45, 46, 48, 54, 74, 105, 209, 213

cultural artefacts 77
cultural semiotics 137
Curry, H. B. 215

—D—

Dalen, D. van 26
Damerow, P. 279
Dapueto, C. 89
data model 325, 327, 328, 330
Davidson, D. 205, 213
Davis, C. 25
Davis, J. W. 25
de-contextualisation 279
Dedekind, 63, 263-267, 272, 291
deduction 13, 36, 40, 46, 50, 51, 53, 55, 57, 69, 70, 99, 147, 220-222, 251, 269, 313, 315, 321, 343
Deleuze, G. 73, 180, 186, 189
Demidov, S. 272
Descartes, R. 15, 77, 107, 137, 203, 206, 207, 241, 244, 246, 249, 257, 281-283, 362
Detel, W. 19, 21
Dewey, J. 203, 210, 211, 213
diagram 19, 35-43, 45-50, 53, 55, 57-65, 70, 72, 93, 106, 111, 113, 128, 316, 318, 319, 350, 377
diagrammatic reasoning 35-37, 45-48, 55, 57-60, 63, 65, 313, 316, 319
diagrammatic thinking 62, 63, 66, 169
didactical laws 159, 161
didactical research 153, 225
Diderot, D. 168
Dieudonné, J. 2, 265, 287, 290
Dinges, H. 305, 310, 367
division of cognitive labor 191, 197-199
Dokic, J. 168
domain dependency 191
Donald, M. 71, 263
Dörfler, W. 55, 58, 60, 63, 65, 66
Doria, F. A. 359
Dorier, J.-L. 270
Dormolen, J. van 168
Dressel, G. 178
dualism 182, 191, 231
Dubinsky, E. 25
Duell, O. K. 192
Duval, R. 92, 103

—E—

Eagle, M. N. 205, 214
early hominids 299, 301
Ebbinghaus, H. 17, 21
Eisele, C. 22
Ekeland, I. 168
Elby, A. 198
Ellenberger, H. E. 205, 214
embodiment 77, 83, 207, 208
Endler, O. 287, 292, 293, 294, 298
Engel, F 315, 321, 322
Engel-Tiercelin, C. 315, 321, 322
Engestroem, Y. 77-79, 89
English, L. D. 90
Englund, R. K. 279
epistemological belief 191, 192, 194-199
epistemological obstacles 24, 117, 118, 119
epistemological triangle 91, 93, 100, 101, 103, 120
epistemology 2, 11, 17, 35, 43, 91, 114, 117, 131, 134, 138-142, 148, 154, 171, 191, 192, 196-199, 203, 204, 211-213, 224, 229-232, 234, 236, 237, 316, 363
Ernest, P. 23-25, 27, 33, 154, 158
Escher, P. 21
evidence 6, 32, 35, 36, 41, 42, 87, 194, 196, 198, 205, 213, 243, 246-248, 251, 264, 268, 270, 271, 335
Evra, J. van 322
experiment 15, 36, 45, 47, 50, 77, 82, 83, 84, 88, 117, 131, 143-145, 147, 163, 168, 206, 223, 241, 253-255, 319, 321

—F—

Fann, K. T. 313, 322
Favero, M. L. 298
Ferrari, P. 89
Ferrero, E. 89
Ferri, F. 77, 89, 90
Feuerbach, ? 182, 190
Feynman, R. P 11, 21
Fichtner, B. 179, 181, 189
Field, J. V. 81, 89, 91
Fischer, R 169, 178, 361- 378
Fleck, L. 197
Forman, P. 215- 218, 221, 226
Fornel, M. de 168
foundations crisis 215, 216, 217, 218, 223
foundations of arithmetic 263, 265, 267, 271, 272
foundations of linear algebra 263
foundations of probability 305
Fourier analyses 299
Freeman, E. 55
Frege, G.. 93, 104, 266, 267
French, S. 159, 163, 168, 275, 359, 377
Frenkel-Brunswick, E. 194, 199
Freud, S. 85, 89, 204, 205, 207, 210, 212-214
Friedman, M. 229, 268, 320, 322

—G—

Galileo, G. 20, 144, 149, 246, 373
Gariglietti, G., 194
Garuti, R. 89, 90
Gauss 62, 234, 265, 268
general education 169, 172, 180, 181
generalization 5, 11, 14, 19, 20, 35, 36, 38, 40, 42, 53, 103, 155, 181, 183, 206, 208-210, 212, 219, 235, 255, 284, 293, 364
geometric algebra 275, 280
geometry 1, 3, 16, 20, 25, 36, 47, 55, 77, 82, 95, 105, 109, 111, 138, 232, 235, 236, 242, 249, 251-253, 256, 263, 265, 266, 269, 272, 280-282, 284, 290, 294, 305, 308, 316, 318, 319, 357, 372, 373, 377
gestures 24, 32, 82, 83, 85-87, 94, 137, 141, 142, 146
Giere, R. N. 319, 322
Gillies, D. A. 263, 264, 267
Gödel, K. 12, 13-15, 21, 215, 219, 220, 225, 265
Goldschmidt, G. 89
Gomes, J. M. 298
Gowers, W. T. 10, 21
Graßmann, G. 263-272, 284
Graßmann, H. 263- 272, 284
Grattan-Guinness, I. 281, 313, 315, 322
group theory 25, 268, 272, 303
Guattari, F. 73, 180, 186, 189

—H—

Haack, S. 237
habit 40, 108, 149, 203, 206-212
Hallett, D. 197
Hammer, D. 198
Hanson, N. R. 257
Harré, R. 27
Hartkämper, A. 333
Hartshorne, Ch. 22, 214
Heath, Th. 275, 284
Heidegger, M. 142
Helmholtz, H. v. 17, 212-214
Hendry, J. 226
Hersh, R. 349
Hertel, R. K. 214
Heyting, A. 26
Hilbert, D. 11-13, 54, 55, 64, 215-218, 220-226, 235, 264, 266, 267, 272, 302
Hintikka, J. 50, 51, 55
Hirzenbruch, F. 298
history of mathematics 6, 88, 109, 177, 263, 281, 284, 287, 295, 296, 297, 351, 359, 371, 372, 374
history of science 122, 205, 229, 241, 295, 371, 374

Hitt, F. 115, 322
Hofer, B. K. 192, 193, 195, 196, 197, 198
Hoffmann, M. H. G. 1, 45, 55, 57, 60, 66, 69, 319, 322, 323
Hofmann, J. R. 276
Hofstadter, D 14, 21
Holt, R. R.. 204, 214
Hopkins, J. 213
Houser, N. 322
Howell, K. L. 195
Howson, G. 153, 158
Hoyles, C. 89
Høyrup, J. 279, 280
Hughes, M. 30
Hughes, R. I. G. 30
Huizinga, J. 168
Hull, K. 46, 48, 313, 315, 321, 322
Humboldt, W. V. 180, 181, 182, 189
Husserl, E. 142
hypostatic abstraction 5, 45, 49, 50, 52, 53, 55, 321
hypothesis 19, 40, 69, 70, 82, 83, 164, 165, 208, 210, 241, 243-245, 299, 317, 318, 321, 336

—I—

icon 10, 13, 39, 41, 42, 46, 47, 57, 105, 107-109, 114
index 10, 41, 105, 107, 109
infinitesimal 230, 254, 355, 356, 357

—J—

Jahnke, H. N. 2, 89, 95, 104, 215, 224, 226, 367, 371
Jameson, Fr. 190
Jehng, J.-C. J. 194
Johnson, M. 93, 94, 104
Johnson, S. D. 194
Jones, A. 250

—K—

Kadunz, G. 178
Kaiser-El-Safti, M. 214
Kamenka, E. 190
Kant, I. 9, 10, 12, 15, 17, 21, 35, 39-41, 45, 55, 66, 137-142, 144, 147-149, 203, 206, 207, 213, 231, 267, 268, 272, 316, 318, 319, 320, 322, 374, 375, 376
Kapadia, R. 104
Kardash, C. M. 195
Kargon, R. 257
Kaufmann, F. 220
Keil, F. 198
Keitel, C. 2
Keller, R. 73
Kerkhove, B. van 322, 323
Ketner, K. L. 50, 214
Kihlstrom, J. F. 204, 214
Kilpatrick, J. 89, 154, 158
King, P. M. 192, 194
Kinkel, W. 231
Kirshner, D. 115
Kitchener, K. S. 192, 194
Kitcher, P 359
Klaua, D. 65, 66
Klein, J. 1
Knoche, N. 89
knowledge 4, 9, 10, 13, 14, 16, 17, 21, 24, 28, 35-40, 43, 45, 47, 49, 69, 79, 92, 93, 108, 118-120, 122, 131, 137-140, 142, 147, 148, 154, 155, 159-162, 164-168, 171, 172, 174, 177, 179, 181, 191-199, 206, 207, 229-236, 289, 295, 297, 319, 322, 349, 354, 368, 376
knowledge domain 195, 196, 197
Koyama, M. 103
Krause, D. 359
Kreisel, G. 215
Krull, W. 287, 290-294, 298
Krutetskii, V. A. 114
Kuhn, D. 15, 21, 153-155, 157, 158, 171, 192, 288, 326

—L—

Laborde, C. 88, 89
Lakoff, G. 80, 89, 93, 94, 104
Latour, B. 197
Lax, P. D. 311
learning 2-6, 16, 17, 19, 23-27, 45, 46, 48, 51, 60, 64, 66-68, 70-72, 74, 79, 105, 111, 115, 117, 118, 120, 123, 130, 132, 133, 154, 159, 161-167, 170, 172, 179, 180, 181, 184-186, 189, 191, 194-196, 203, 204, 207, 377
Lederman, N. G. 192
Leibniz, G. W. 18, 20, 21, 109, 137, 149, 206, 208, 270, 307, 321
Lektorsky, V. A. 144
Lemut, E. 89
Lenhard, J. 1, 313, 320, 322
Leuzinger-Bohleber, M. 205, 214
Levinson, D. J. 194, 199
Levy, S. H. 55
Lewis, A. C. 14, 269, 270
Locke, J. 35, 36, 37, 38, 40, 42, 43, 137
logic of situations 159, 163
logical types 169, 174
Lorenzen, P. 220
Lorenzer, A. 204, 214
Lotman, Y. 149
Lütkehaus, L. 205, 214
Lutz, D. J. 198

—M—

MacGregor, M. 95, 104
Machover, M. 25
macrodidactique 159, 162, 163
Maier, H. 93, 104
Malt, B. 198
manipulatives 117, 123, 124, 130, 131, 133
Marburg school 229, 230, 233, 235, 236
Mariotti, M. A. 77, 81, 82, 88-90
Mark, G.-C. R. 323
Marquard, O. 206, 213, 214
Martin, G. 267, 268
Martinez, A. 89, 333
Martinez, C. 89, 333
Marx, K. 182, 188, 190
mathematic 174, 241
mathematical activity 4, 5, 9, 11, 12, 23, 24, 27, 52, 79, 92, 317
mathematical existence 57
mathematical knowledge 5, 9, 14, 16, 26, 33, 38, 40, 52, 91, 92, 93, 108, 161, 164, 165, 167, 192, 291, 336, 349, 372
mathematical practice 64, 102, 313, 336, 349, 351, 358, 359
mathematical truth 13, 335, 336
mathematics and organization 169
mathematics as a cultural system 215
mathematics education 1, 2-4, 6, 10, 25, 26, 55, 58, 105, 111, 117, 118, 120, 123, 130, 133, 134, 153, 154, 156-158, 161, 192, 295, 296, 322, 366, 371, 376
mathematics in Brazil 287, 296
mathematics institutions 287
mathematization 132, 242, 243, 244, 252, 256
Matte-Blanco, I. 204, 214
Mazur, B. 311
McDowell, J. 237
McKinnon, J. 204, 214
Mead, G. H. 203, 210, 211, 212, 214
meaning 3, 5, 9, 10-13, 16, 23-25, 27-30, 32, 41, 42, 46, 54, 58, 61, 64, 68, 70-74, 88, 91-94, 100, 101, 106, 109, 118, 120, 123, 137, 145, 146, 148, 165, 184, 185, 188, 190, 193, 196, 198, 206-208, 211, 215, 216, 221, 224, 225, 233, 242, 245, 251, 252, 256, 269, 275, 278, 279, 281, 306, 356
mediating 3, 40, 53, 67, 68, 74, 138, 153, 155, 158, 183, 185, 218
Medvedev, F. A. 272
Meheus, J. 359
Mellin-Olsen, S. 28, 168

metaphor 20, 31, 68, 71, 91, 95, 101, 105, 108, 109, 111, 114, 229, 230, 232, 233, 236, 237, 244, 304, 373
metonymy 105, 107, 109, 111, 114
microdidactique 159, 162
Mies, Th. 203, 205, 206, 214
Miller, G. A. 168
Miller, R. P. 168
Miller, V. R. 168
mind 20, 36-42, 45, 48, 49, 54, 71, 73, 79, 85, 89, 97, 100, 110, 111, 119-121, 125, 137, 138, 142, 147, 173, 176, 179, 190, 203, 205, 206, 208, 211, 212, 220, 223, 232, 244, 266, 307, 316, 319, 349, 365, 371
modeling 159, 162, 313, 317, 318, 320
modernity 362
Moore, D. S. 306, 311
Moore, E. C. 322
Moore, G. H 270
Moore, W. S. 193
Moser, U. 205, 214
Moulin, H. 168
Moulines, C. U. 325, 331, 333
Mulholland, K. A. 74
multiplication 58, 163, 264, 268, 270, 275, 276, 278-284, 354, 356
Mumford, D. 307, 311
Murphey, M. G. 313, 314, 322
Murphy, P. K. 196
Murty, U. S. R. 59, 66

—N—

Nakahara, T. 103
Nasi, D. 89
Natorp, P. 229-237
natural philosophy 210, 242, 247, 257, 372
Neokantianism 229, 230, 231, 232, 236
Newman, J. R. 358, 368
Newton, I. 11, 109, 198, 224, 241-257, 307, 329, 333, 359
Nichanian, M. 139, 140
Nissen, H. J. 279
Nnadozie, A. A. 118
non-commutativity 275, 284
non-compactness 349, 358
normal science and scientific paradigms 153
number 11, 12, 13, 15, 19, 20, 24, 25, 35, 37, 38, 50, 52, 57, 60- 65, 68, 85, 87, 88, 99, 100, 102, 138, 145, 155, 157, 160, 161, 163, 168, 170, 174, 176, 194, 197, 198, 219, 220, 222, 224, 234, 244, 246, 249, 254, 266, 267, 268, 275, 276, 279, 281, 295, 299, 302, 313, 327, 330, 342, 343, 345, 349-353, 355-357, 365, 369, 377
Nunez, R. 80, 89

—O—

Ogden, C. K. 93, 104
ontological commitment 5, 12, 42, 325, 326, 328, 329
ontological reduction 325, 332
ontology 37, 154, 197, 199, 225, 232, 233, 325, 326, 328, 329, 335, 336, 346, 347, 349, 357, 358
Op t' Eynde, P., 192
optics, Newton 241, 242, 246, 247, 248, 249, 252, 253, 255, 256, 257
Ossimitz, G. 178
Otte, M. 1-6, 9, 10, 12, 14, 21, 22, 35, 39, 40, 43, 52, 60, 66, 68, 74, 77, 79, 89, 90, 91, 93, 104-108, 113, 114, 117-120, 122, 134, 148, 153, 158, 159, 165, 168-177, 189, 191, 192, 197, 226, 264, 267, 272, 287, 295-298, 309, 314, 319, 321, 322, 349, 357-359, 361-377
Ozkar, M. 89

—P—

Pappus of Alexandria 248
Parenti, L. 89

participation 27, 29, 31, 67, 68, 85, 362
Pea, R. D. 144
Peano, 265, 267, 270, 271, 346
Pears, D. 205, 214
Peckhaus, V. 263, 270, 271
Peirce, C. S. 5, 6, 10, 13, 18-20, 22, 31, 35-43, 45-55, 57-60, 64, 66-74, 105-111, 114, 138, 142, 149, 174, 203, 206-211, 213, 214, 272, 313-323, 375
Peixoto, M. C. 290, 297, 298
perception 19, 20, 29, 36, 37, 41, 48, 77, 80, 119, 128, 138, 156, 174, 186-188, 209, 222, 232, 313, 315, 320, 362, 363, 372
Pergola, M. 77, 88, 89
Perrig, W. J. 204, 212, 214
Perrig-Chiello, P. 214
Perry, W. G. 193, 194, 197, 198
Peschek, W. 178
Peschl, E. 293, 298
Pfeifer, R. 204, 205, 214
phenomenology 137, 223
philosophy of consciousness 203, 206, 211, 213
physics experiment 15, 36, 45, 47, 50, 77, 82, 83, 84, 88, 117, 131, 143-145, 147, 163, 168, 206, 223, 241, 253-255, 319, 321
Piaget, J. 2, 66, 119, 137, 138, 140-142, 144, 147, 193, 197, 365
Pinker, S. 94, 104
Pintrich, P. R. 192, 193, 195-197
Poincaré, H. 17, 218, 219, 221, 225
polyphony 77, 79, 82
polysemy 77, 79, 82, 83, 87, 88
Popper, K. 237
Porter, W. 89
Possel, R. de 168
practical thinking 117-121, 123, 132, 133
pragmatic 173, 181, 210, 211, 313, 315-318, 320-322
pragmatism 20, 106, 203-212, 313, 314, 375, 376
pragmatism, American 203, 206, 210, 211
Presmeg, N. C. 43, 105, 108, 109, 115
primary school 1, 77, 89, 160, 366
principle 2, 5, 11, 17, 18, 19, 20, 59, 62, 65, 71, 138, 160, 164, 165, 169, 175, 180, 205, 216-218, 221, 243-245, 250, 251, 253-255, 263, 269, 282, 309, 321, 330, 331, 364, 372, 374, 377
production of knowledge 159, 181
psychoanalysis 186, 203, 204, 205, 209, 210, 211, 212, 213
Pyenson, L. 288, 298
Pythagoras theorem 117, 131

—Q—

Queiroz, J. 55
Quéré, L. 163, 168
Quine, W. V. O. 233, 234, 325, 333, 375

—R—

Rabardel, P. 141, 142
Radford, L. 137, 142, 143
Radu, M. 263, 272, 322, 323
Rambow, R. 198
random variables 305, 307, 308-310
randomness 308, 349
reflexiveness 218
Reich, K. 282
relativism 191, 193
representamen 46, 105-111, 113, 114
representation 5, 6, 9, 10, 16, 18, 19, 32, 40, 46, 47, 49, 54, 57, 58, 61, 72, 73, 78, 80, 82, 83, 92, 98, 102, 105, 107, 119, 128, 132, 133, 138, 139, 141, 142, 176, 189, 216, 220, 234, 303, 316, 327, 330
restrictedness of multiplication 275
Reynolds, C. 323
Richards, F. A. 93, 104

Rickert, H. 230
Rieber, R. W. 190
Rivas, U. 333
Roberts, D. D. 322
Robin, R. S. 22, 36, 43
Robutti, O. 146
Rolnik, S. 189
Rorty, R. 229, 237
Rotman, B. 26, 31, 82, 90, 102-104
Rowe, D. 226, 280, 281, 333
Rowe, D. E. 226, 280, 281, 333
Russell, B. 109, 110, 115, 218-220, 267, 320, 375, 377
Ruthven, K. 156, 158

—S—

Sacks, O. 9, 22, 94, 95, 104
Sáenz-Ludlow, A. 43
Salimbeni, L.. 282, 283
Salomon, G. 53
Sandkühler, H. J. 214
Sanford, R. N. 194, 199
Santos, G.T. 298
Sarita, A. 298
Saussure, F. 24, 105, 115, 141
Scali, E. 89
Scheier, V. 204, 214
schema 40, 58, 137-144, 146-149, 186, 189, 335, 337, 338
Schleiermacher 17, 270
Schmidt, H. J. 333
Schneider, E. 178
Schneider, H. 205, 214
Schoenfeld, A. H. 33, 192
Scholz, E 205, 214, 226
Scholz, R. 205, 214, 226
Schommer, M. 192, 194, 195, 196, 197
Schommer-Aikins, M. 195, 196
Schraw, G.. 196
Schubring, G. 270, 275, 296, 371
Schwarzenberger, R. . 156, 158
Seeger, F. 1, 2, 67, 104, 182, 187, 190
self 1, 3, 6, 18, 27, 29, 31, 33, 40, 51, 68, 95, 119, 170-173, 179-189, 192, 203, 206, 207, 211, 216, 218, 225, 257, 264, 315, 317, 319, 362, 364
self-reference 192, 207, 218
self-referentiality 179, 180, 187, 188, 189
semiotic mediation 73, 77-79, 84, 87-89
semiotic systems 23, 24, 25, 141, 145, 149
semiotics 5, 6, 24, 26, 35, 39, 45, 67, 70, 105, 106, 107, 111, 114, 137, 186, 313, 319, 375
Sfard, A. 25, 106, 115
Shank, G. 69
Shapiro, A. E 257
Sheets-Pyenson, S. 288, 298
Shevrin, H., 205, 214
Sierpinska, A. 24, 104, 117, 118, 120, 123, 154, 158
sign 5, 6, 9, 10, 11, 13, 14, 19, 23, 24-33, 37-42, 46, 50, 52, 60, 67-74, 84, 85, 91-95, 101, 102, 106, 108, 110, 114, 120, 122, 141, 145, 170, 175, 183, 187, 188, 203, 207, 210, 211, 266, 278, 279, 294
sign language 70, 91, 92, 94, 95, 188
sign transformations 23, 30
Sigurdson, S. 226
Silva, A. E. G. 287, 289, 292, 296, 298
Silva, C. M. S. 287, 289, 292, 296, 298
simplicity of knowledge 191, 194, 196
Skemp, R. R. 28, 29, 30
s-knowledge 159, 162, 163, 166, 167
Skovsmose, O. 89
Sneed, J.D. 224, 331, 333, 367, 375
sociology of knowledge 169, 171
Sokol, B. W. 197
Spengler, O. 216, 217, 221
Stacey, K. 89
Stachowiak, H. 272

statements vs. sentences 335
statistical teaching 305
Stegmüller, W. 224, 367
Steinbring, H. 91-93, 104, 120, 122, 367
Steiner, H. G. 6, 153, 158, 366
Steinhart, E. Ch. 237
Stjernfelt, F. 55, 319, 323
subsumption 140, 148, 325, 329
Sulloway, F. J. 205, 214
Sutherland, R. 168
symbol 9, 10, 39, 41, 42, 46, 52, 78, 91, 93, 100, 101, 102, 105, 107, 108, 113, 114, 139, 141, 144, 188, 204, 329
synthesis 4, 23, 39, 40, 43, 45, 48, 49, 143, 156, 172, 177, 241, 247, 248, 249, 250, 252, 256

—T—

teacher 2, 3, 24, 27-33, 74, 77, 79, 82, 84, 86-88, 95-100, 102, 117, 119, 123, 127-131, 133, 134, 153, 154, 156, 157, 161, 166, 167, 170, 172, 191, 199, 257, 309, 310, 365, 368, 370, 376, 377
teacher education 130, 153, 154, 156, 368, 370
teaching 2-6, 23, 26, 43, 67, 68, 74, 77, 79, 83, 87, 88, 91, 102, 104, 109, 117-119, 120, 123, 128, 130-134, 153, 160-168, 191, 195, 276, 292, 305, 307- 309, 365, 366, 368, 376
teaching and learning of mathematics 23, 43, 88, 123
teaching experiment 77, 83, 87, 88
theorematic deduction 36, 45, 46, 50-53
theoretical model 133, 325, 346
theoretical thinking 117, 118, 120-122, 123, 130-133
theoric transformation 45, 46, 50, 51, 53, 54
theory of situations 159, 162, 163, 164, 166
Thiel, Ch. 216
Thom, R. 3, 9, 15, 171, 189, 203, 229, 326
thought experiment 31, 117, 123, 130, 131, 133, 134, 143, 144, 146, 149
Trabant, J. 69
Troelstra, A. 26
Tropfke, J. 275
truth 12, 14, 15, 39, 41, 50, 57, 73, 94, 110, 122, 143, 171, 172, 193, 194, 197-199, 203, 234, 237, 244, 245, 249, 255, 315, 316, 335-347, 350, 352, 362, 373
truth-condition 335, 337, 338, 340-342, 345
Turrini, M. 89

—U—

Ulam, S. 314, 323
unconscious 29, 119, 170, 186, 187, 203-206, 208-214
unconscious inference 203, 209, 213
Unger, H. 293, 298
Unguru, S. 280, 281
universals 10, 105, 109, 110

—V—

vagueness 349, 351
variables 46, 91, 99, 165, 170, 175, 233, 234, 305, 307, 309, 325, 351, 355, 356, 358
Veresov, N. 68
Verillon, P. 141, 142
Verschaffel, L. 192
Villegas, L 333
Vincent, J. 89
virtual worlds 299, 301-304
visual pyramid 77, 81, 82, 85, 86
Vithal, R. 4
Voigt, J. 104, 190

Vygotskij, L. S. 27, 67-69, 73, 74, 77, 78, 80, 90, 122, 180, 182-187, 189, 190

—W—

Waerden, van der 279
Walker, K. 196
Wang, H. 264, 266, 267, 272
Warfield, V. 168
Wartofsky, M. 77, 78, 87, 90
Waschescio, U. 104, 190
Webb, J. C. 12, 13, 22, 264, 265, 266, 272
Weiss, P. 22, 214, 313, 323
Weyl, H. 215, 217, 218, 220-226
Whitehead, A. N. 91, 104, 220, 270
Whiteside, D. T 249
Whiteside, D. T. 249
Williams, W. J. 214
Winkelmann, B. 178
Wippich, W. 204, 214
Wittgenstein, L. 10, 64
Wollheim, R. 213
Woolgar, S. 197
work of art 169, 179, 187, 188
Wussing, H. 268, 269

—Z—

Zanoli, C. 89
Zeppelin, I. v. 205, 214
Zermelo, E. 216
Ziehen, T. 17, 18, 22