

How to Support Prescriptive Statements by Empirical Research: Some Missing Parts

Christof Wecker

Published online: 16 August 2012
© Springer Science+Business Media, LLC 2012

Abstract This paper continues a discussion started in a special issue about the acceptability of prescriptive statements in educational research articles. In light of some ambiguities concerning what counts as a prescriptive statement, and the special issue's focus on causal relations as a requirement for the justification of prescriptive statements, a more detailed characterization of prescriptive statements and the structure of a complete argumentation for them is offered. This reveals two major obstacles to valid justifications of prescriptive statements that have received little attention before: the problem of normativity and the problem of generality. The proposed solution to the problem of normativity—that is, the impossibility to support prescriptive statements by empirical research alone—is to take into account that arguments for prescriptive statements target an audience that may agree on the values of many educational goals. The proposed solution to the problem of generality—that is, the necessity of well-established general causal regularities for the justification of prescriptive statements—requires appropriate designs for testing the generality of claims. Methodological suggestions include nested designs with quasi-representative samples of treatments as well as standard procedures for determining the cost and side effects on an agreed-upon set of relevant outcome dimensions for both current practice and any new intervention. If such steps are undertaken, prescriptive statements are no less justified in discussion sections than general descriptive claims as long as the final decision about them is suspended if the available normative and empirical arguments are not yet conclusive.

Keywords Prescriptive statements · Normativity · External validity · Generality · Representative design

C. Wecker (✉)
Empirische Pädagogik und Pädagogische Psychologie, Ludwig-Maximilians-Universität München,
Leopoldstraße 13, 80802 Munich, Germany
e-mail: christof.wecker@psy.lmu.de

The second issue of *Educational Psychology Review* in 2011 was devoted to the philosophical, theoretical, and methodological aspects of the question: “When is it acceptable to make prescriptive statements in educational research articles?” (“Call for papers” 2009, p. 91). This question was targeted at the conditions under which a prescriptive statement or recommendation for practice is justified (Kulikowich and Sperling 2011, p. 190).

The contributions to this special issue strikingly concurred on one necessary component of the justification of a prescriptive statement: they unanimously supported the requirement that there has to be evidence that the action recommended in a prescriptive statement causes a desired effect. Or, as one author put it: “It is not unreasonable to suggest that prescriptive statements rely to a large degree on the demonstration of causal effects” (Martin 2011, p. 236; see also Graesser and Hu 2011, p. 280; Yussen 2011, p. 288).

As a consequence, the main focus of the special issue was on methodological prerequisites for the justification of causal claims. In particular, the contributions emphasized the importance of the internal and, to some extent, external aspects of the validity of research designs as basic criteria (Marley and Levin 2011, p. 199–202; Sun and Pan 2011, p. 217). They also discussed the merits and problems of specific methodological options with respect to the demonstration of causal relations, such as the appropriateness or inappropriateness of correlational, quasi-experimental, and experimental designs (Martin 2011, pp. 237–241, 243; Marley and Levin 2011, pp. 200f.; Yussen 2011, p. 288), or stage approaches to research, such as the strategy of “CAREfully crafted intervention research” (Marley and Levin 2011, pp. 202–204). Furthermore, the contributions put forward requirements for the presentation and discussion of results (O’Connell and Gray 2011, pp. 246, 251f.; Sun and Pan 2011, pp. 217f.) that enable the reader to assess the validity of causal claims.

A symposium at the 2010 annual meeting of the American Educational Research Association in Denver, Colorado, was organized by the guest editors as a companion event to the special issue. The symposium comprised position statements by the editors of several leading journals in the field (Kulikowich and Sperling 2010, 2011, p. 194; Robinson 2011, p. 293) with the same focus on methodological prerequisites for the justification of causal claims. Concerning the treatment of causality in the special issue, there is hardly anything to disagree with or even to add.

However, as acknowledged by one of the contributors (Martin 2011, p. 236), the justification of prescriptive statements does not rest upon the demonstration of causal relations alone. Luckily, the guest editors announced that they “fully recognize that the dialog regarding the nature of prescriptive claims and their warrant in educational and psychological research shall continue” and that they believe that “there remains much to discuss” (Kulikowich and Sperling 2011, p. 194). I fully agree with them and take this as an invitation.

In light of some ambiguities in the special issue concerning *what counts as a prescriptive statement*, in the following section I shall try to clarify this question. Given the special issue’s focus on the demonstration of causal relations as one crucial requirement for the justification of prescriptive statements, I shall then draw on insights from decision theory to portray the parts of a complete body of information from which a prescriptive statement can be derived. This might be called the *logic of prescriptive statements*. Next, I shall turn to two blind spots in the discussion thus far. One is the *problem of normative premises* that are required for deriving prescriptive statements. The other is the *problem of generality* in prescriptive statements and its methodological consequences. This involves broadening the treatment of the external aspect of validity. Finally, I shall present some *methodological consequences* and *concluding thoughts*.

What Is a Prescriptive Statement?

That so much has been written in the special issue about the requirements for the justification of causal claims seems to be partly the consequence of some lack of clarity and consensus about what counts as a prescriptive statement. How the contributors conceive of prescriptive statements is illustrated by the varying semi-formal patterns and examples they offer as a characterization (Graesser and Hu 2011, pp. 279, 281; Kulikowich and Sperling 2011, p. 190; Marley and Levin 2011, pp. 197f.; “Call for papers” 2009, p. 91; Sun and Pan 2011, p. 207; Yussen 2011, pp. 287f.).

Here are some of them:

- (1) “A causes B” (Graesser and Hu 2011, p. 281).
- (2) “If x, then y” (Kulikowich and Sperling 2011, p. 190).
- (3) “If persons take Action X, then Situation Y will improve” (Marley and Levin 2011, p. 197; see also Yussen 2011, p. 287).
- (4) A certain intervention “can help struggling students in early adolescence develop at least an awareness of strategies for overcoming or at least compensating for their reading difficulties” (Sun and Pan 2011, p. 207, quoting Cantrell *et al.* 2010, p. 270).
- (5) “If children are provided with manipulatives while reading, their comprehension will increase” (Marley and Levin 2011, p. 198).

Despite some similarities among most of them, they belong to several different types on differing levels of generality. But, as shall become clear shortly, none of them is a prescriptive statement.

In contrast, part of the following sentence from one contribution, which the authors themselves labeled as a prescriptive statement, actually qualifies: “We conclude with the prescriptive statement that only after achieving these high standards of research credibility should educational researchers offer prescriptive statements” (Marley and Levin 2011, p. 197). Beyond its topic, this sentence obviously differs considerably from the examples just quoted. It almost seems that, in contrast to other contexts, as soon as a statement is about an educational intervention, some authors take “prescriptive” to mean “causal.”

Unfortunately, a pertinent definition of the term “prescriptive” is difficult to find. In the disciplines dealing with the logic of prescriptive statements (i.e., meta-ethics and, partly, the philosophy of language), its meaning seems to be not too controversial; hence, typically no definition is provided. For example, in one of the classic texts, moral language is characterized as a special case of prescriptive language right at the outset, without any definition of the latter (Hare 1952).

Nevertheless, some indications about the meaning of the term “prescriptive statement” can be found in the literature. First, prescriptive statements are typically regarded as a counterpart of descriptive statements (von Wright 1963, p. 3). The distinction between these two kinds of statements is perhaps best explained on the basis of their so-called *direction of fit*, a term first used by Austin (1952–53/1979, p. 141): In the case of prescriptive statements, the world ought to fit the words, whereas in the case of descriptive statements, the words ought to fit the world. A classic example of the world-to-words direction of fit is a shopping list. If the purchases do not fit the list, it is not the fault of the list. In contrast, a list about a persons’ purchases contrived by a detective is an example of words-to-world direction of fit. If the purchases do not fit the list, nothing is wrong with the purchases (Anscombe 1963, p. 56; Searle 1979; see also Sun and Pan 2011, p. 209). Accordingly, prescriptive statements are used to influence people’s actions (Falk 1953, pp. 147f.; von Wright 1963, pp. 2, 30). In

contrast, causal claims are not prescriptive statements because they have a words-to-world direction of fit, and are therefore descriptive.

Next, the meaning of the term “prescriptive statement” can be characterized by the role of prescriptions in communication: “Prescribing an act to someone ... is to provide a direct and unequivocal answer to the question, ‘What should I do?’” (Taylor 1962, p. 216). Put differently, “prescribing” is the same as “telling someone what he ought to do” (Taylor 1962, p. 213). The paradigm kind of a prescription is an order or command (Falk 1953, p. 148; see also von Wright 1963, pp. 7, 71); other important kinds include guidance, recommendations, and advice (Taylor 1962, pp. 219, 222, 223). Accordingly, typical ways of formulating a prescriptive statement are imperatives (Taylor 1962, p. 225; von Wright 1963, p. 98) and so-called *deontic vocabulary* (Glüer and Wikforss 2009, sect. 1.2; von Wright 1963, pp. 100f.). Examples of “deontic vocabulary” are the modals “ought (not) to”, “should (not)”, “must (not)”, and “may”, as well as the words “prescribed”, “forbidden”, and “allowed” (Glüer and Wikforss 2009, sect. 1.2; Taylor 1962, pp. 220, 227; von Wright 1963, p. 96). Furthermore, prescriptions can be introduced by “I suggest”, “My advice is”, or “I recommend” (Taylor 1962, p. 226).

This characterization of the meaning of “prescriptive statement” is pretty much in line with several remarks in one paper in the special issue: “Prescriptive statements ... consist of recommendations about how things should be done in practice” (Sun and Pan 2011, p. 207); or “Prescriptive statements are usually concerned with what should be done to achieve the goals; thus, they are value-laden and demonstrate a distinct logic of imperatives” (Sun and Pan 2011, p. 212). The latter quote leads directly to the question of how prescriptive statements can be justified.

The Logic of Prescriptive Statements

One of the recommendations offered in the special issue was that journals should impose rules that specify when prescriptive statements are allowed in articles (Yussen 2011, p. 291). This would go far beyond current editorial practice concerning other types of statements. In light of such proposals, the logical relations of prescriptive statements to other statements that may justify them need to be considered carefully. The present section extends the treatment of this topic in one of the papers in the special issue (Sun and Pan 2011, pp. 212f.).

Usually, a rational argument in favor of a prescriptive statement is analyzed as a so-called practical syllogism with the prescriptive statement as its conclusion. Applied to prescriptive statements about educational interventions, the form of a practical syllogism would be as follows (see Anscombe 1963, pp. 58; 1989, p. 380; Kenny 1975, pp. 70f.):

- (1) The state of affairs X is an educational goal.
- (2) The educational intervention A will bring about the state of affairs X .
- (3) Therefore, the educational intervention A should be implemented.

In this type of practical syllogism, the major premise (1) expresses a positive evaluation of a certain state of affairs (Kenny 1975, p. 70; Sun and Pan 2011, p. 212). The minor premise (2) states that a certain action is *sufficient* for bringing about this state of affairs (Kenny 1975, pp. 70f., 81f.; see von Wright 1972, pp. 40f. for a different view). This condition is fulfilled if the minor premise states that the action will cause this outcome (Graesser and Hu 2011, p. 281).

According to the schema of the practical syllogism, an educational intervention is prescribed if it has a beneficial effect (Sun and Pan 2011, p. 209). However, this conclusion

may be unwarranted at times. For example, there may be other—potentially better—methods of bringing about the outcome. In this case, the justification of the prescriptive statement requires that the method is the best of these several options (Taylor 1962, pp. 225, 227f.). Such cases are more appropriately dealt with on the basis of decision theory (Kenny 1975, p. 95), which extends the practical syllogism.

Decision theory was developed around the middle of the twentieth century at the intersection of philosophy and economics (see, e.g., Jeffrey 1965; Luce and Raiffa 1957; Savage 1954). Similar to formal logic, which deals with the principles underlying sound reasoning, decision theory is a prescriptive discipline that deals with the principles underlying rational decisions. Basic concepts involved in the analysis of decisions under risk include *probability* and *utility*.

Before applying this approach to the analysis of arguments concerning prescriptive statements in educational research, I would like to deal with two potential objections against this approach right away. One objection could be: “This is schematic.” The answer is: yes, and rightly so. Schemata that are sufficiently appropriate for particular cases organize information about them and thereby help us understand them (Rumelhart and Norman 1978, pp. 43f.). In the present context, this may come at a cost—that with respect to a specific decision about an educational intervention, some piece of information that seems relevant may not fit neatly into the framework provided here. Conversely, this cost is arguably outweighed by the fact that in general the “schematic” framework of decision theory constitutes a heuristic that can facilitate the identification of issues of importance in such a decision, and structure them in a way that highlights their role in the overall logic underlying the decision.

Another objection could be targeted at the mathematical character of this approach: “We don’t know these figures, and we never will.” This also may be true. However, there are cases in which we do know, for example, that two states of affairs are about equally valuable, or that two educational interventions bring about a certain state of affairs with approximately equal probability. In such cases, a comparative assessment of options with respect to only one aspect may be sufficient to decide the issue in line with decision theory. In other cases, all aspects of the decision situation may favor the same option in a comparative evaluation, although it may be impossible to express them by exact figures. Of course, these arguments are not decisive because decision theory still might be of no use in the majority of cases. However, the “metrical” variant of decision theory presented here, which uses cardinal numbers, can be treated as the basis for deriving a “comparative” variant that will help in most cases. Accordingly, the variant involving numbers presented below should rather be regarded as a paradigm to frame real-life cases.

Several elements, according to decision theory, form the basis of a rational decision, and by analogy, the basis of the justification of a prescriptive statement (see Table 1). The first is the set of available options. These may be represented as (the heads of) rows in so-called desirability and probability matrices (Jeffrey 1965, pp. 1f., 5f.). In Table 1, both matrices are integrated into one table (as will be explained shortly). An example could be three hypothetical educational interventions to foster mathematical skills (see the entries in the first column of Table 1).

The second element is the set of potential outcomes of these options (Jeffrey 1965, p. 2). In the present example, only the simple case of dichotomous outcomes is considered (i.e., whether a positive or negative outcome of some relevance occurs or not). For example, the three intervention options may or may not foster the mathematical skills of learners (by a certain minimum degree, which may be expressed as an effect size, represented by the entries “Yes” and “No” in the two columns under outcome dimension 1 in Table 1). Of course, outcomes may also be differentiated on a polytomous or continuous scale.

Table 1 Schematic example of a decision matrix for educational interventions

Outcome dimensions	1		2		3		Cost	E(U)
	Yes	No	Yes	No	Yes	No		
Outcomes								
Outcome evaluations	9	-9	-7	5	5	-1		
Educational intervention 1	.3	.7	.6	.4	.1	.9	1	-7.2
Educational intervention 2	.8	.2	.4	.6	.3	.7	2	4.4
Educational intervention 3	.7	.3	.2	.8	.1	.9	2	3.8

The outcome dimensions in this example are: (1) the mathematical skills of the learners improve (by a certain minimum effect size). (2) the self-concept of the learners deteriorates. (3) the slope of the regression of the learners' skills on their socio-economic status decreases

Numbers in the Outcome evaluations row represent the values attached to each of the two possible outcomes on each of these three outcome dimensions. Numbers in the second to seventh columns represent the probabilities associated with each of the two possible outcomes on each of the three outcome dimensions. Numbers in the Cost column represent the cost and effort required for implementing the corresponding option. Numbers in the E(U) column represent the overall expected utility of the corresponding option. According to the last column, the hypothetical information represented in this table favors educational intervention 2

Each outcome is assigned a value; otherwise, it would not make sense to include it in the deliberation anyway. These values constitute the third element that enters the justification of the prescription. These values are represented by the numbers in the cells of a so-called desirability matrix (Jeffrey 1965, pp. 5f.). In the example, the gain in mathematical skills that constitutes the positive outcome is arbitrarily assigned a value of 9, and the state of affairs that this gain does not occur is arbitrarily assigned a value of -9 (see the values in the "outcome evaluations" row of Table 1). It can be assumed that these evaluations of the potential outcomes are the same for each of the three options. Accordingly, the whole desirability matrix can be represented by a single row of outcome evaluations.

Each outcome occurs with varying probabilities, depending on the different options. These probabilities constitute the fourth element that enters the justification of the prescription. They are represented by the numbers in the cells of a so-called probability matrix (Jeffrey 1965, p. 6). In the example, the probability of the outcome that the gain in mathematical skills occurs if the first option is taken is 30 %, whereas the probability of the complementary outcome that this gain does not occur is 70 % (see the entries in the second and third column in the row labeled "Educational intervention 1" in Table 1).

However, there may be not just one dimension of outcomes, but a whole set of such dimensions, each of which consists of a set of possible outcomes along with their values and probabilities (conditionalized on the options). For instance, the three options in our example may or may not negatively affect the learners' academic self-concept (outcome dimension 2) or lower the social gradient (i.e., the slope of the regression of the learners' skills on their socio-economic background; outcome dimension 3).

Finally, the different educational interventions may require different amounts of cost and effort to implement. This needs to be taken into account as well (see the entries in the column labeled "Cost" in Table 1).

Based on all this information, the best option can be chosen. According to the so-called *Bayesian decision rule*, this is the option associated with the highest expected utility $E(U)$ (if more than one option is associated with the same maximum expected utility, then one of these). The expected utility of an option is the (potentially weighted) sum of all products of the values and conditional probabilities for the option, across all outcomes and outcome

dimensions (Jeffrey 1965, pp. 1; 6). In our example, for educational intervention 1, this would amount to $(.3 \times 9) + (.7 \times -9) + (.6 \times -7) + (.4 \times 5) + (.1 \times 5) + (.9 \times -1) = -6.2$. If we further take into account the expected costs as a kind of certain negative consequence of each option and therefore subtract them from the probability-weighted sums of the outcome values, we get the final estimates of the expected utility of each option (see the entries in the column labeled “E(U)” in Table 1). Accordingly, the expected utility of educational intervention 1 would be $-6.2 - 1 = -7.2$. Given the expected utilities of all three options in the present example, the second educational intervention should be chosen, and a statement that prescribes this action would be the one justified by this body of information. This option is most likely to bring about an improvement of the learners’ mathematical skills, (arguably) sufficiently unlikely to negatively affect their academic self-concept, and most likely to reduce the social gradient. In addition, option 2 requires no more cost and effort than educational intervention 3.

As argued above, it is not strictly necessary to determine the exact values of the variables of this pattern of inference. Rather, this formal characterization of the elements on which the justification of a prescriptive statement is based highlights what kinds of information need to be considered to fully justify a prescriptive statement. Accordingly, several types of statements that express these kinds information can be distinguished. As I will elaborate shortly, some of them can be provided by empirical research. If a statement of any of these types is considered in isolation, however, its consequences for the prescriptive statement may be unclear (Black 1989, p. 414). Therefore, statements of all the following types are required to provide a complete justification of a decision:

- Type 1: *Educational intervention A is likely to cause outcome X* (with a sufficient degree of probability). For example: “Educational intervention 2 is likely to bring about a satisfactory gain in the learners’ mathematical skills.” This kind of statement could be a finding from an empirical study.
- Type 2: *Outcome X has a certain positive value*. For example: “A gain in the learners’ mathematical skills is something good.” or “A gain in the learners’ mathematical skills is an important educational goal.” This kind of statement cannot be derived from an empirical study (at least considered in isolation).
- Type 3: *The implementation of educational intervention A is associated with a certain expected amount of cost and effort*. This kind of statement could be a finding from a certain kind of empirical (non-intervention) study.
- Type 4: *Apart from education interventions A, B, C, ..., there are no further options that are likely to cause outcome X* (with a sufficient degree of probability). This is a universal hypothetical assumption that cannot be firmly established by empirical research. Rather, in the more favorable case, it will be falsified by scientific progress that leads to the discovery of such new options.
- Type 5: *Apart from the outcomes on outcome dimensions x, y, z, ..., there are no further potential outcomes of educational interventions A, B, C, ... with relevant positive or negative values*. This is also a universal hypothetical assumption that integrates descriptive as well as evaluative aspects and cannot be firmly established by empirical research, but only falsified as science progresses.
- Type 6: *Compared to all other known options that cause outcome X with sufficient probability, the expected positive (side) effects, the expected negative side effects and the expected necessary effort for the implementation of educational intervention A as a whole are more favorable*. This constitutes a rather comprehensive “optimality assumption” (Black 1989, p. 414; Taylor 1962, pp. 227f.) based on a

whole body of empirical findings from more than a single study (in combination with a whole body of evaluations). Under ideal circumstances most of these findings would be available from prior research.

Mainly because a statement of type 6 is a necessary component in a complete justification of a prescriptive statement, prescriptive statements are hardly ever justified on the basis of empirical research (Marley and Levin 2011, p. 205; Robinson 2011, p. 294), let alone on the basis of a single empirical study (Sun and Pan 2011, p. 216). Nevertheless, the difficulty in collecting all the different kinds of empirical findings that—given certain values—may jointly justify a specific prescriptive statement does not free educational research from the obligation to work hard to provide these kinds of information (Sun and Pan 2011, p. 208).

The purpose of this brief sketch of the application of decision theory to the justification of prescriptive statements in education is not meant to provide an exercise in some formal branch of philosophy and economics, but to lay bare the kinds of information that, according to the formal study of rationality, are required for the justification of a prescriptive statement about an educational intervention. I next discuss two major obstacles to valid justifications of prescriptive statements based on empirical studies that have received little attention in the special issue.

Two Blind Spots in the Discussion

The problem of normativity

As discussed in the preceding section, the justification of prescriptive statements involves value judgments (see, e.g., type 2 above; Taylor 1962, p. 213; see also Sun and Pan 2011, p. 212), and their validity is crucial for the soundness of the justification of a prescriptive statement (Sun and Pan 2011, pp. 212, 216). Many scientists think, however, that value judgments have no place in scientific inquiry. One of the roots of this view is the so-called is–ought question originating from the following famous passage by David Hume:

In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning ... when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, *is*, and *is not*, I meet with no proposition that is not connected with an *ought*, or an *ought not*. This change is imperceptible; but is, however, of the last consequence. For as this *ought*, or an *ought not*, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it. (Hume 1739—40/1978, p. 469)

Couched in more contemporary vocabulary, according to Hume, prescriptive statements can be deduced only from a set of premises of which at least one is prescriptive or normative itself. In the context of the justification of prescriptions for educational practitioners, this constitutes a problem because the methods of empirical research yield only justifications for descriptive statements. Hence, the justification of prescriptive statements appears to transgress the professional responsibility of scientific inquiry.

Related to this view is the claim that science should be free of value judgments (see Sun and Pan 2011, p. 215), which has been advocated by the sociologist and philosopher Max

Weber. Weber acknowledges that a main goal of scientific research is to provide information for practical decisions about interventions (Weber 1904/1988, p. 148) and that the valuations underlying the identification of practical problems influence the selection of research topics and questions (Weber 1904/1988, p. 158; 1917/1988, p. 511). However, according to Weber, the actual decision-making has to take place outside of a scientific discipline (Weber 1904/1988, p. 150). As Weber puts it: “An empirical discipline cannot teach anyone what he should do, but only, what he can and—under certain circumstances—what he wants to do” (Weber 1904/1988, p. 151, my translation). His main reason is that the evaluation of interventions is unambiguous only if the goal is given and the available interventions differ exclusively with respect to the certainty and rapidity of the attainment of the goal as well as the quantitative yield. Otherwise, further value judgments about the different aspects of the interventions play a role in the decision (Weber 1917/1988, p. 529).

The is-ought problem and Weber’s ban on value judgments may make the justification of prescriptive statements by means of empirical research appear hopeless. However, some arguments suggest that this conclusion may be premature. The first is based on the assumption that recommendations for practitioners are best understood as so-called *hypothetical imperatives*. A hypothetical imperative is a prescription that is conditional on a certain goal, such as in the sentence: “If educational goal X is to be attained, educational intervention A should be implemented”—as opposed to a so-called *categorical imperative*, which is an unconditional prescription, such as: “Educational intervention A should be implemented” (see Kant 1785/1956, BA 39f.). According to this assumption, the latter is to be regarded as shorthand for the former. From this perspective, the problem of the justification of prescriptive statements disappears because hypothetical imperatives are assumed to be equivalent to descriptive statements. The hypothetical imperative just mentioned is assumed to be equivalent to “If educational intervention A is implemented, then educational goal X will be attained.” According to Weber, a hypothetical imperative is nothing but the conversion of a causal statement (Weber 1917/1988, p. 538). Hence, it could be justified without reliance on any normative or value judgment.

However, a hypothetical imperative and such a corresponding descriptive statement are not equivalent. A statement such as: “If educational goal X is to be attained, educational intervention A should be implemented” rests on further (implicit) value judgments that are not “outsourced” to the conditional phrase containing the hypothetical value judgment about the goal. These further value judgments include, for example, assumptions about the optimality of intervention A with respect to the effort required and its likely side effects. Accordingly, although it is true that prescriptions for practitioners in discussion sections of empirical studies may be best understood as hypothetical imperatives, and that empirical research can contribute to their justification, further normative preconditions for their justification usually need to be fulfilled. In sum, a recommendation cannot be reduced to a descriptive statement.

The second argument is partly complementary to this first one, but at the same time it does provide independent support for the view that empirical research can provide argumentative support for prescriptive statements: The audience that (empirical) arguments for prescriptive statements have to convince comprises real people with their actual normative beliefs. There is a strong and uncontested consensus that many dispositions that are typical outcomes of educational interventions, such as knowledge, understanding, skills, abilities, competences, achievement motivation, broad interests, or a strong academic self-concept, but also certain socio-structural outcomes, such as a flat social gradient, are desirable for everyone. Of course, the degree of agreement in these issues may differ among countries, and the weighting of different educational goals may be more contested if they are partially

in conflict with each other and cannot all be fully attained at the same time. Nevertheless, consensus about important educational goals constitutes a fulcrum for empirical arguments even in favor of unconditional prescriptive statements about educational interventions because a categorical imperative is appropriate if it recommends an action that leads to a known legitimate goal of the person to whom it is directed (Edwards 1955, pp. 132f.). Similarly, implicit optimality assumptions in hypothetical imperatives can be justified in light of information about further value judgments of the recipients.

To be sure, a researcher advancing a prescriptive statement may be mistaken in assuming a consensus about a value judgment that is a prerequisite of this conclusion (Weber 1904/1988, p. 153; 1917/1988, p. 502). However, this can easily be detected if the normative assumptions underlying the prescriptions are made explicit—no matter whether as arguments or, maybe more appropriately, as part of the conditions of a hypothetical imperative—because this subjects them to scrutiny and criticism. In cases of disagreement about such normative assumptions, the justification of prescriptive statements about educational interventions also requires arguments for the normative assumptions. Ultimately, these will have to be based on ethical reasons, and the discussion of such issues will certainly benefit from contributions from moral philosophy and philosophy of education. But again, empirical research will help in this respect. One contribution may be information about the relations among different educational goals or their relations to major life goals (for instance, information about the extent to which a weak academic self-concept can interfere with knowledge acquisition, or whether achievement motivation or domain knowledge is more important for employability, or whether broad interests are associated with higher satisfaction and happiness). Another contribution may be information that certain outcome dimensions with contested value (such as the social gradient) are unaffected by some educational intervention.

As a consequence, it appears that if the goal is to arrive at justified prescriptions, in cases of disagreement about normative prerequisites these issues need to be resolved through discussion. There are, however, rational methods for resolving such disputes, and empirical research is certainly among them.

The problem of generality

Beyond normative prerequisites, prescriptive statements are based on conditional predictions that certain outcomes will occur with some probability if a certain option is taken (type 1 and, in part, types 4 to 6); for example, that under the present circumstances the implementation of a specific educational intervention is highly likely to increase learners' mathematical skills by a certain amount. To derive such conditional predictions, general statements about robust empirical laws are indispensable; for example, that a certain type of educational intervention increases learners' mathematical skills. These general statements must hold for real-world settings in the field of practice rather than just the laboratory to be of any worth for the justification of a prescriptive statement for practice (Shaw *et al.* 2010, pp. 983f.; Sun and Pan 2011, p. 209), and they are the hypotheses that are tested in empirical studies.

Therefore, the methodological requirements for evidence in support of the generality of a hypothesis warrant thorough consideration despite the fact that these issues are often regarded as less important than the internal aspect of validity (e.g., Campbell 1957, p. 310; Campbell and Stanley 1966, p. 5). I will take a particular empirical study as a starting point for the discussion of these issues. This study belongs to a line of research concerning the question of whether an intervention that emphasizes the utility value of some content can increase its *perceived utility value* or personal relevance, and thereby foster both learners'

interest in the topic and their *performance* (Hulleman *et al.* 2010, pp. 882f.; Hulleman and Harckiewicz 2009, p. 411). Apparently, this research is of high practical relevance. It has been conducted in line with high methodological standards, and it has been published in highly competitive outlets (*Science* and the *Journal of Educational Psychology*).

The authors of this study conjectured: “In the classroom, one way to highlight utility value could be to ask students to describe the relevance of course material to their own lives” (Hulleman *et al.* 2010, p. 882). Accordingly, they tested the hypothesis “that a situational intervention that encourages individuals to make a connection between a task and their lives (i.e., a relevance intervention) will increase perceptions of utility value for the task” (p. 882).

One of the studies by these authors involved 107 undergraduates in a control group design comparing the utility value intervention to writing an essay about a random topic. The participants were required to learn a specific technique for mentally solving two-digit multiplication problems. Amongst other variables, the learners’ perceived utility values of the technique and their interest levels before and after the intervention were measured by means of Likert-type scales (Hulleman *et al.* 2010, p. 883). The experimental treatment was implemented as follows:

Next, the experimenter handed the participant a folded sheet of paper (to ensure that the experimenter was blind to condition) that contained instructions for writing either a relevance or control essay. ... Participants in the relevance condition were asked to “type a short essay (1–3 paragraphs in length) briefly describing the potential relevance of this technique to your own life, or to the lives of college students in general. Of course, you’ll probably need more practice with the technique to really appreciate its personal relevance, but for purposes of this writing exercise, please focus on how this technique could be useful to you or to other college students, and give examples”. (Hulleman *et al.* 2010, pp. 883f.)

Given this operationalization of the independent variable, apparently the application of the treatment in future situations should be rather unproblematic because the authors have not only given an exact description of it, but even provided the instructions word for word.

When discussing the appropriateness of this methodology with respect to the amount of support it lends to the generality of the hypothesis under investigation, it is advisable to look closely at the formulation of the hypothesis (quoted above). Plausibly, the persons whose perceptions of utility value according to the hypothesis will increase, are the “individuals” who are encouraged “to make a connection between a task and their lives” (Hulleman *et al.* 2010, p. 882). Hence, a (rather broad) general term (“individuals”) is used to delineate a population of persons for whom the intervention is hypothesized to be effective. According to textbook wisdom in statistics, the generality of a hypothesis with respect to a population of persons can be tested and supported by combining an appropriate sampling technique (preferably random sampling) with the application of appropriate inferential statistical tests (Fisher 1925/2003, pp. 6f.; 1935/2003, p. 3; Winer 1962, pp. 4f.).

A well-known problem is that typically the samples used in educational psychological intervention research are not random (Marley and Levin 2011, p. 201). It has been argued that under these circumstances, the application of inferential statistical tests secures the generality of a hypothesis for a population of persons “like those observed” (Cornfield and Tukey 1956, p. 913). As this problem has already been discussed sufficiently in the past (Cornfield and Tukey 1956, pp. 912f.; Bracht and Glass 1968, pp. 440–442; Snow 1974, p. 270), I shall not further elaborate on it here.

There are, however, other “dimensions of generality” in the hypothesis tested in this example. In the same manner as the general term “individuals” in the hypothesis characterizes a population of people, the general expression “a situational intervention that encourages individuals to make a connection between a task and their lives” characterizes a class of interventions that is hypothesized to be effective (Hulleman *et al.* 2010, p. 882)—or a “universe” of interventions, to use a term introduced by Brunswik (1955, pp. 198, 202; 1956, p. 37) and also used by Cronbach and his colleagues (e.g., Cronbach *et al.* 1963, pp. 144ff.; Cronbach *et al.* 1972, p. 18). In addition to the dimension of generality explicitly mentioned in the hypothesis, the hypothesis has to be regarded as implicitly general with respect to such aspects as the content domain, the teacher, or the institutional environment simply because the hypothesis is left unrestricted in these respects (Cook 2004, p. 91; Cronbach and Shapiro 1982, pp. 82f.). Apparently, the hypothesis claims effectiveness not only for the specific instructions reprinted verbatim in the article, but for any intervention fulfilling the specification in the hypothesis, and likewise for different content domains and teachers who administer the intervention in different institutions (Campbell 1957, p. 309; Cook 1993, pp. 39f.; 2000, pp. 4f.; 2002, p. 6037; 2004, p. 93; Cronbach and Shapiro 1982, pp. 80–82, 93–97).

In analogy to statistical textbook wisdom about how to demonstrate the generality of a hypothesis with respect to a population of persons, the generality of a hypothesis with respect to classes of interventions, content domains, instructors, institutional environments, and the like would have to be tested and supported by combining an appropriate sampling technique from these universes (Snow 1974, pp. 272f.; to be accurate, this was already stated by Fisher 1925/2003, pp. 2f.) with the application of appropriate inferential statistical tests. As a direct consequence of this line of reasoning, no support for the generality of the hypothesis with respect to these aspects (Marley and Levin 2011, p. 202)—and hence also no support for a prescriptive statement dependent on this hypothesis—is provided by the study in the present example. To be sure, this is not a weakness specific to this particular study, but a feature that besets a substantial, maybe even major, share of educational intervention research.

Of course, most researchers are aware of this problem, although they may frame it differently, and there are at least two standard solutions for it. One is to conduct replication studies across which one or more of the aspects mentioned are varied (e.g., teachers or subjects; see Campbell 1957, p. 310; Levin 2004, p. 176; Levin and O’Donnell 1999, p. 190; Marley and Levin 2011, p. 202; Robinson 2006, p. 116; Rosenshine 1994, p. 250; Snow 1974, pp. 277f.; Sun and Pan 2011, pp. 214f.; Yussen 2011, p. 288). In fact, this was actually done in the present example, which further emphasizes that this study comes from a solid research program (Hulleman *et al.* 2010, pp. 887f.; Hulleman and Harckiewicz 2009, p. 1411). What is puzzling about the replication approach, however, is the fact that hardly anybody would conduct a study with one participant and, if “successful”, move on to “replicate it” with one or several other participants. Why are people and interventions not treated alike?

The other standard solution is meta-analysis (Cook 1993, p. 65; 2000, p. 25; 2002, p. 6041; 2004, p. 109; Sun and Pan 2011, p. 216). However, this approach helps only if the studies in the sample can be plausibly regarded at least as an approximation of a random sample of instances of the intervention, content domains, teachers, and institutional environments (see Hedges 1994, p. 35; Matt and Cook 1994, pp. 514; 516f.; Rosenthal and DiMatteo 2001, p. 66). It is no secret, however, that in many areas of research, any sample of studies often will be severely biased toward one or two laboratories using a rather limited set of materials.

A different approach would be to deal with generality in terms of aspects other than populations of people in a way that parallels how generality with respect to people is tested and supported. This approach acknowledges that the assessment of generalizability requires an estimate of components of variance in the variability of outcomes that can be attributed to different facets (Cronbach *et al.* 1972, pp. 16f.; see also Cook 2004, p. 99; Snow 1974, p. 285). This is possible in so-called representative (Brunswick 1955, p. 198) and quasi-representative designs (Snow 1974, pp. 271, 273): A representative design involves samples of situations (or, more specifically, instances of interventions, content domains, teachers, etc.) and the application of inferential statistics to them (Brunswick 1955, pp. 198, 202; 1956, p. 37). If this is not feasible, employing some kind of “quota” or stratified sample of situations that mirrors the distribution of attributes in the universe has been suggested (Brunswick 1955, p. 204). This may be combined with tests for interactions between the independent variable and the stratification factors, which has been termed “quasi-representative design” (Snow 1974, pp. 271, 273; see also Cook 1993, pp. 58f., 76f.; 2000, pp. 19, 36f.; 2002, p. 6040; 2004, p. 108). An advantage of a representative design is that it allows for natural covariation in multivariate distributions within the actual ecology (Brunswick 1955, p. 199; Snow 1974, p. 273). It is striking that such designs are quite common in other fields, such as communication research (Brashers and Jackson 1999, pp. 460f.).

It would be inappropriate, of course, to demand of every single study that it provide evidence for the generality of the hypothesis under investigation with respect to any aspect involved by implementing a fully representative design that employs random sampling from all of these dimensions. It is appropriate, however, to demand some kind of sampling from such dimensions in later studies in a research program (see Burkhardt and Schoenfeld 2003, p. 8; McDonald *et al.* 2006, p. 17; Snow 1974, p. 285) because a hypothesis may be wrong in at least two different ways. Current practice is very much preoccupied with the failure to eliminate potential alternative causes (in line with Campbell 1957, p. 310; Campbell and Stanley 1966, p. 5). A hypothesis may be no less wrong if the terms used to refer to independent or dependent variables in them are overly general, or if important boundary conditions are left unspecified. Unfortunately, there is only one way for a hypothesis to be correct: to avoid both of these defects (McDonald *et al.* 2006, p. 20), which is the reason why it is misleading to speak of so-called *internal* and *external validity* (because a study can only be valid or fall short of it) rather than maybe “internal” and “external” invalidity.

Methodological Consequences

The preceding discussion of two problems for the justification of prescriptive statements in educational research has some implications for research methodology as well as the reporting of empirical studies, which are briefly compiled in this section.

- (1) *Evidence for the generality of hypotheses needs to be provided not only with respect to persons, but also with respect to treatments (and other aspects).* In terms of research design, the kind of “representative design” required could include drawing (random) samples of teachers or instructional designers who then create their own versions of an intervention based on a written specification, and each use their own versions with a sample of learners. The analysis of such a design would then have to take into account the multi-level structure of this kind of data and treat the teachers or instructional designers as random factors nested within the conditions of the design (Fontenelle *et al.* 1985, pp. 103f.; Serlin *et al.* 2003, p. 527). In terms of reporting, complete characterizations of the

treatment need to be provided, both on the level of the general specification of the type of treatment from which the treatment instances employed have been sampled, and on the level of these instances used in the study (see Rosenshine 1994, pp. 248f.; Harris and Pressley 1994, p. 204). All boundary conditions held constant also need to be described because they limit generalization. The opportunities to make supporting online material available on the websites of journals might be very helpful in this respect.

- (2) *The effects of current practice with respect to all relevant outcome dimensions, and the cost and effort associated with this practice, need to be determined.* According to the logic of prescriptive statements, any recommended intervention must be demonstrated to be superior to all other available options, and current practice is certainly among them. This would be very much facilitated if there were a more uniform current practice, as seems to be the case in medicine—but this opens up another issue that cannot be discussed satisfactorily in the present context. In any case, a reference point is needed that has to be overtopped by any alternative educational intervention recommended in a prescriptive statement.
- (3) *Intervention studies need to include routine assessments of the cost and effort required to implement the intervention as well as the side effects on a fixed set of agreed-upon relevant outcome dimensions (e.g., interest, academic self-concept, social gradient).* According to the logic of prescriptive statements, this information is of vital importance for a holistic evaluation of an intervention that may justify its prescription, and it would qualify as a standard item under the heading “what to write in a results section” in any article about an empirical study designed to support a prescriptive statement. A fixed set of agreed-upon instruments would be helpful in this respect (Brown and Wilson 2011, pp. 221, 232; Burkhardt and Schoenfeld 2003, p. 9).
- (4) *Contested normative premises need to be supported based on contributions from fields such as moral philosophy, linguistic analysis, and the philosophy of education.* This is inevitable against the backdrop of the logic of prescriptive statements, but at the same time transgresses the capacity of our discipline. If we maintain the goal to provide support for prescriptive statements, even in cases in which pertinent values are disputed, these issues need to be subjected to discussion. What needs to be further debated is the proper place for such discussion.

These methodological requirements illustrate what has been noted by some contributors to the special issue on prescriptive statements (e.g., Marley and Levin 2011, p. 203): A single study will almost certainly not be sufficient for justifying a prescription. If we stick to this goal, comprehensive research programs specifically designed to support such statements are necessary (Sun and Pan 2011, pp. 214, 216). However, point (3) implies that these will probably have to comprise theoretically heterogeneous sub-projects that contribute all the pieces that make up the whole puzzle.

Concluding Thoughts

The preceding discussion suggests that in some cases our field would benefit from more attention to how statements are phrased. For example, maybe most of the articles in the special issue and certainly the present one contain more prescriptive statements than the average discussion section of an empirical study. When discussing whether—or under which conditions—a certain kind of statement can be justified, our own meta-linguistic statements

need to be justified as well. This requires that we focus on the level of the particular statements under scrutiny and provide precise characterizations of the patterns underlying these statements. This was the basis for the analysis of the logic of prescriptive statements offered in the present article.

The same applies, however, to the object level of research. At times it seems that we do not say precisely what we mean, and do not take statements to mean precisely what they say. To convince oneself of this diagnosis, I recommend a quick look at a couple of empirical studies: It is not difficult at all to find an article in which hypotheses are formulated in the present tense (indicating that they are meant to be general) throughout the introductory pages, whereas statements throughout the discussion section occur only in the past tense (suggesting that they refer to nothing more than the sample of the study). It is difficult to believe that researchers do not want to say something on a general, theoretical level after they have presented their findings, and I suspect that most of us typically read those statements as general and theoretical (for evidence that recipients are vulnerable to stronger interpretations than warranted by the methods employed; see Shaw *et al.* 2010, pp. 984f.). Stating in a discussion section what one really wants to claim on a general, theoretical level might help to increase the perceived demand for a methodology that warrants conclusions concerning the generality of a hypothesis.

How can we finally answer the initial question: are prescriptive statements in discussion sections appropriate or not? As already indicated, a particular prescriptive statement will certainly not be justified based on the findings from a single empirical study in isolation. The main reason for this is not that replications are needed before firm conclusions can be drawn, but that, as shown above, different kinds of information are needed as premises to justify a prescriptive statement. If, however, other studies, which may be rather diverse in their research questions and methodology, and normative arguments are taken into account as well, a particular prescriptive statement can be justified, and it may therefore occur in a discussion section.

This position receives further support from an analogy with theoretical statements in discussion sections. Many articles are strong in discussing the consequences of the study for (potentially competing) theoretical assumptions, without necessarily settling the issue completely. Instead, they clearly delineate the need for further research that would contribute to a clarification of the open questions. It would be no less feasible to discuss the consequences of a study for prescriptions for practitioners in the same manner (i.e., narrowing down the promising approaches), without necessarily settling for a specific recommendation. Instead, future research could be delineated that would be pertinent to the same practical issues, but may be concerned with aspects that are completely different when considered from a theoretical point of view (e.g., effects of an intervention on another outcome dimension). Under such circumstances, it would still be possible to support the reader in drawing correct practical conclusions, even if no definitive recommendation is justified at this point (Nolen and Talbert 2011, p. 271). At least this can be expected from educational research as a practical discipline (Sun and Pan 2011, p. 208). Despite the high demands on appropriate support for prescriptive statements, we should not dismiss them from our scholarly discourse, but rather strive to increase the quality of the evidence we can provide for them.

Acknowledgments I would like to thank Andreas Hetmanek and Anna Wecker-Delatrée for helpful comments on the first draft of this article.

References

- Anscombe, G. E. M. (1963). *Intention* (2nd ed.). Cambridge, MA: Harvard University Press.
- Anscombe, E. (1989). Von Wright on practical inference. In P. A. Schilpp & L. E. Hahn (Eds.), *The philosophy of Georg Henrik von Wright* (pp. 377–404). La Salle, IL: Open Court.
- Austin, J. L. (1952–53/1979). How to talk: Some simple ways. In J. L. Austin, *Philosophical Papers*. Oxford, UK: Oxford University Press.
- Black, M. (1989). Some remarks about 'practical reasoning'. In P. A. Schilpp & L. E. Hahn (Eds.), *The philosophy of Georg Henrik von Wright* (pp. 405–416). La Salle, IL: Open Court.
- Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal*, 5(4), 437–474.
- Brashers, D. E., & Jackson, S. (1999). Changing conceptions of 'message effects': A 24-year overview. *Human Communication Research*, 25, 457–477.
- Brown, N. J. S., & Wilson, M. (2011). A model of cognition: The missing cornerstone of assessment. *Educational Psychology Review*, 23(2), 221–234.
- Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62(3), 193–217.
- Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.
- Burkhardt, H., & Schoenfeld, A. H. (2003). Improving educational research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, 32(9), 3–14.
- Call for papers. (2009). *Educational Psychology Review*, 21, 91.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54(4), 297–312.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.
- Cantrell, S. C., Almasi, J. F., Carter, J. C., Rintamaa, M., & Madden, A. (2010). The impact of a strategy-based intervention on the comprehension and strategy use of struggling adolescent readers. *Journal of Educational Psychology*, 102, 257–280.
- Cook, T. D. (1993). A quasi-sampling theory of the generalization of causal relationships. In L. Sechrest & A. G. Scott (Eds.), *New directions for program evaluation: Understanding causes and generalizing about them* (pp. 39–82). San Francisco, CA: Jossey-Bass.
- Cook, T. D. (2000). Toward a practical theory of external validity. In L. Bickman (Ed.), *Validity & social experimentation: Donald Campbell's legacy* (Vol. 1, pp. 3–43). Thousand Oaks, CA: Sage.
- Cook, T. D. (2002). Generalization: Conceptions in the social sciences. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 9, pp. 6037–6043). Oxford, UK: Elsevier.
- Cook, T. D. (2004). Causal generalization: How Campbell and Cronbach influenced my theoretical thinking on this topic. In M. C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 88–113). Thousand Oaks, CA: Sage.
- Cornfield, J., & Tukey, J. W. (1956). Average values of mean squares in factorials. *Annals of Mathematical Statistics*, 27(4), 907–949.
- Cronbach, L. J., & Shapiro, K. (1982). *Designing evaluations of educational and social programs*. San Francisco, CA: Jossey-Bass.
- Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology*, 16, 137–163.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Edwards, P. (1955). *The logic of moral discourse*. New York: Free Press.
- Falk, W. D. (1953). Goading and guiding. *Mind*, 62, 145–171.
- Fisher, R. A. (1925/2003). Statistical methods for research workers. In R. A. Fisher, *Statistical methods, experimental design and scientific inference: A re-issue of statistical methods for research workers, the design of experiments, and statistical methods and scientific inference*. Oxford, UK: Oxford University Press.
- Fisher, R. A. (1935/2003). The design of experiments. In R. A. Fisher, *Statistical methods, experimental design and scientific inference: A re-issue of statistical methods for research workers, the design of experiments, and statistical methods and scientific inference*. Oxford, UK: Oxford University Press.
- Fontenelle, G. A., Phillips, A. P., & Lane, D. M. (1985). Generalizing across stimuli as well as subjects: A neglected aspect of external validity. *Journal of Applied Psychology*, 70(1), 101–107.
- Glüer, K., & Wikforss, Å. (2009). The normativity of meaning and content. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Stanford, CA: The Metaphysics Research Lab, Center for the Study of

- Language and Information, Stanford University. Retrieved from <http://plato.stanford.edu/entries/meaning-normativity>.
- Graesser, A. C., & Hu, X. (2011). Commentary on causal prescriptive statements. *Educational Psychology Review*, 23(2), 279–285.
- Hare, R. M. (1952). *The language of morals*. Oxford, UK: Clarendon.
- Harris, K. R., & Pressley, M. (1994). Increasing the quality of educational intervention research. *Educational Psychology Review*, 6(3), 191–208.
- Hedges, L. V. (1994). Statistical considerations. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 29–38). New York: Sage.
- Hulleman, C. S., & Harckiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, 326, 1410–1412.
- Hulleman, C. S., Godes, O., Hendricks, B. L., & Harckiewicz, J. M. (2010). Enhancing interest and performance with a utility value information. *Journal of Educational Psychology*, 102(4), 880–895.
- Hume, D. (1739–40/1978). *A treatise of human nature* (2nd ed.). Oxford, UK: Clarendon.
- Jeffrey, R. C. (1965). *The logic of decision*. New York: McGraw-Hill.
- Kant, I. (1785/1956). Grundlegung zur Metaphysik der Sitten [Foundations of the metaphysics of morals]. In I. Kant, *Werke in sechs Bänden. Bd. 4: Schriften zur Ethik und Religionsphilosophie* (reprinted 1998; pp. 7–102). Darmstadt, Germany: Wissenschaftliche Buchgesellschaft.
- Kenny, A. (1975). *Will, freedom and power*. Oxford, UK: Blackwell.
- Kulikowich, J. M. & Sperling, R. A. (2010, April/May). *Or so you claim! What claims do our methodologies allow? A town meeting*. Symposium conducted at the Annual Meeting of the American Educational Research Association “Understanding Complex Ecologies in a Changing World”, Denver, 30th April to 4th May 2010.
- Kulikowich, J. M., & Sperling, R. A. (2011). Prescriptive statements: Philosophical, theoretical, and methodological considerations: Introduction to the special issue. *Educational Psychology Review*, 23(2), 189–195.
- Levin, J. R. (2004). Random thoughts on the (in)credibility of educational-psychological intervention research. *Educational Psychologist*, 39(3), 173–184.
- Levin, J. R., & O’Donnell, A. M. (1999). What to do about educational psychology’s credibility gaps? *Issues in Education*, 5, 177–229.
- Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Wiley.
- Marley, S. C., & Levin, J. R. (2011). When are prescriptive statements in educational research justified? *Educational Psychology Review*, 23(2), 197–206.
- Martin, A. J. (2011). Prescriptive statements and educational practice: What can structural equation modeling (SEM) offer? *Educational Psychology Review*, 23(2), 235–244.
- Matt, G. E., & Cook, T. D. (1994). Threats to the validity of research syntheses. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 503–520). New York: Sage.
- McDonald, S. K., Keesler, V. A., Kauffman, N. J., & Schneider, B. (2006). Scaling-up exemplary interventions. *Educational Researcher*, 35(3), 15–24.
- Nolen, A., & Talbert, T. (2011). Qualitative assertions as prescriptive statements. *Educational Psychology Review*, 23(2), 263–271.
- O’Connell, A. A., & Gray, D. L. (2011). Cause and event: Supporting causal claims through logistic models. *Educational Psychology Review*, 23(2), 245–261.
- Robinson, D. H. (2006). Editorial. *Educational Psychology Review*, 18, 115–117.
- Robinson, D. (2011). Editor’s note. *Educational Psychology Review*, 23(2), 293–294.
- Rosenshine, B. (1994). The conduct and reporting of intervention studies. *Educational Psychology Review*, 6(3), 245–254.
- Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52, 59–82.
- Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37–53). Hillsdale, NJ: Erlbaum.
- Savage, L. J. (1954). *The foundations of statistics*. New York: Dover.
- Searle, J. R. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge, UK: Cambridge University Press.
- Serlin, R. C., Wampold, B. E., & Levin, J. R. (2003). Should providers of treatment be regarded as a random factor? If it ain’t broke, don’t ‘fix’ it: Comment on Siemer and Joormann (2003). *Psychological Methods*, 8, 524–534.
- Shaw, S. M., Walls, S. M., Dacy, B. S., Levin, J. R., & Robinson, D. H. (2010). A follow-up note on prescriptive statements in nonintervention research studies. *Journal of Educational Psychology*, 102(4), 982–988.
- Snow, R. E. (1974). Representative and quasi-representative designs for research on teaching. *Review of Educational Research*, 44, 265–293.

- Sun, S., & Pan, W. (2011). The philosophical foundations of prescriptive statements and statistical inference. *Educational Psychology Review*, 23(2), 207–220.
- Taylor, P. W. (1962). Prescribing and evaluating. *Mind*, 71, 213–230.
- von Wright, G. H. (1963). *Norm and action: A logical enquiry*. London, UK: Routledge.
- von Wright, G. H. (1972). On so-called practical inference. *Acta Sociologica*, 15(1), 39–53.
- Weber, M. (1904/1988). Die 'Objektivität' sozialwissenschaftlicher und sozialpolitischer Erkenntnis ['Objectivity' in social science and social policy]. In M. Weber, *Gesammelte Aufsätze zur Wissenschaftslehre* (pp. 146–214). Tübingen, Germany: Mohr.
- Weber, M. (1917/1988). Der Sinn der 'Wertfreiheit' der soziologischen und ökonomischen Wissenschaften [The meaning of 'ethical neutrality' in sociology and economics]. In M. Weber, *Gesammelte Aufsätze zur Wissenschaftslehre* (pp. 489–540). Tübingen, Germany: Mohr.
- Winer, B. J. (1962). *Statistical principles in experimental design*. New York: McGraw-Hill.
- Yussen, S. R. (2011). When is it acceptable to make prescriptive statements in educational research articles? Commentary on the special issue papers. *Educational Psychology Review*, 23(2), 287–292.