Cognition, consensus and consciousness: my replies

Keith Lehrer

Published online: 28 April 2012 © Springer Science+Business Media B.V. 2012

The papers in this volume, some constructive, some critical, some both, are illuminating discussions of my work, and I am honored that they were written and that I have the opportunity to offer my reflections upon them. I begin with simple thanks to Drs. Martina Fürst and Guido Melchior for proposing the conference and to Prof. Stewart Cohen for offering the pages of this distinguished journal.

1 Overview

I begin with an overview of where I now stand. My goals at this point are systematic as recounted in my recent Dewey Lecture because I seek to maximize explanation. It takes a system to maximize explanation. It takes a loop to complete the explanation, even in a theory of truth, which is part of the system. A theory of truth must explain why the theory itself is true or that will be left unexplained. The system and theory contained therein must connect experience with representation to ground the system. The form of connection is exemplar representation using the exemplar of experience to represent things, including itself reflexively. Exemplarization allows us to use experience to represent the world, ourselves, ourselves in the world and the world in ourselves. I propose that to maximize explanation it is essential to leave open how we use the individual exemplars of experience to form the general conceptions of theory. My ontology is based on the existence of individuals and individual qualities that give rise to our explanatory conception of general properties. I acknowledge the utility of properties in explanation but reject the assumption of their existence. For, assuming their existence blocks explanation of the connection between mind and body, theory and experience, self and world. Without appeal to properties, how can we construct explanation? In terms of general

K. Lehrer (🖂)

Department of Philosophy, University of Arizona, Tucson, AZ, USA e-mail: lehrer@email.arizona.edu

conceptions of the world we form by exemplarizing experience. How are we justified? It requires self-trust, aggregating trust of others, in a way that makes us worthy of it. Our trustworthiness loops back onto itself in an explanatory loop. This is only a wordbite of a system, and is not an argument, but laying my assumptions on the page may help to explain why I say what I do in reply to my discussants and critics.

2 My replies: theoretical knowledge

2.1 Reply to Marian David

When I first read the remarkable critical essay by Marian David, I thought he was missing the point of the connection I was trying to articulate about the relationship between being worthy of one's trust, being reasonable, and obtaining truth. But as I look at what he said and what I wrote more carefully, I find his account and his concern to be just. My writing may, quite correctly, be considered a quest for an understanding between what a person accepts, when a person is trustworthy in what he accepts or when a person is reasonable in what he accepts, and when he is successful in obtaining the goals of acceptance. I was mainly interested in putting these notions together to offer a theory of knowledge.

David notes changes in my views about reasonable acceptance. I now think that there are levels of reasonable acceptance, some more restrictive than others. There is a very weak notion of reasonable acceptance of p. It requires only that it not be more reasonable for the subject to accept the denial of p than p, and a strong one that requires that it is more reasonable for the person to accept p than not to accept p. I now favor the latter, but once a distinction between levels of reasonableness is drawn, the matter is settled by a distinction. Those who consider it reasonable to be bold in seeking explanatory power will favor the weak constraint, and those who favor caution to avoid error in the life of reason will favor a stronger constraint. A more substantial question concerns what notion of reasonableness results from the goals of acceptance itself. I use acceptance in a restricted way to represent the objective of accepting that p if and only if p in a way that is worthy of the subject's trust. Such acceptance implies, in a defeasible implication, that the person is reasonable to accept p. The implication from acceptance to reasonableness is explained by the principle of trustworthiness of acceptance. Moreover, the failure of trustworthiness is a defeater for the implication. So the addition of trustworthiness to acceptance simpliciter supports the implication by excluding the defeater of failure of trustworthiness. Trustworthiness of acceptance is the principle that explains why acceptance is reasonable.

My only critical comment of substance about what David writes concerns the issue of the explanation of reasonableness in terms of trustworthiness. I do not define trustworthiness in terms of reasonableness because the notion of being trustworthy is more general and is connected with a person being worthy of trust in many diverse ways. I concede ambivalence of my attitude toward the connection between trustworthiness and reliability. The ambivalence was based, as David notes,

on a view that a person can be worthy of his trust in how he proceeds to obtain truth and avoid error, and yet not be successful in obtaining truth and avoiding error. This will result when the subject is invincibly deceived. So the notion of being trustworthy is one I finally thought of as epistemic following Chisholm, as David notes. I sought to distinguish a normative component of proceeding in a way that is worthy of trust, leading to personal justification, from reliable success in obtaining truth, which must be added as a further condition to obtain the undefeated justification of knowledge.

Of course, as I insisted, to get from the trustworthiness of acceptance to knowledge, one must assume that trustworthiness is successfully truth connected, though again such success is not entailed by trustworthiness. My defense of the way I proceeded in the end is that I thought it important to distinguish the normative, what is worth accepting as true, from the success in obtaining truth in an overall account of knowledge. The possibility of invincible deception precludes the identification. The principle that connects our being trustworthy with success in obtaining truth can explain our success in terms of our proceeding in a trustworthy way.

What David has discerned as the change of my views is that I have come to accept that a theory of trustworthiness must contain, as a part, a theory of how our being worthy of our trust explains the contingent connection between our being trustworthy in what we accept and the successful truth connectedness of that trustworthiness. We aim at proceeding in a way that is worthy of our trust in what we accept and also accept that proceeding in that way is successfully truth connected. If we are right in accepting the latter, our being worthy of our trust is successfully truth connected in a way required for obtaining knowledge. If we are wrong, the prize of knowledge eludes us. Are we aiming at truth in accepting what is true in a way that is worthy of our trust, or are we aiming at worth? We are aiming at a systematic and complex objective. The goal, newly formulated here, is to accept that p just in case p is true in a way that makes accepting it worthy of my trust. The goal is an organic whole that does not decompose into the goals of truth and worth.

The point of bringing in our being worthy of our trust is to indicate that systematic objectives, most saliently, being able to answer objections to accepting that p, rather than simple success in being right about p, even reliable success, is required to distinguish good fortune in arriving at truth, no matter how frequent, from knowledge. Personal justification in accepting p, a normative component, does not suffice for knowledge, even if p is true, as Gettier taught us. You need the right match between personal justification and successful truth connectedness to obtain undefeated justification and knowledge.

2.2 Reply to Pascal Engel

My reply to Engel, who admirably and fairly describes my views, concerns the interpretation of acceptance and trust. First of all I do have a modest disagreement with Engel when he says,

I shall then try to apply the lessons of the classification to his account of trust as the foundation of knowledge.

My disagreement is that I did not intend trust or self-trust to be a foundation of knowledge. I took self-trust and even the principle of being worthy of that trust, being trustworthy, to be required in a coherent system of justification. But, being a coherence theorist, I held that the principle of trustworthiness is supported by the acceptance of other things that justify it, that yield the conclusion that it is reasonable, just as the acceptance of it supports the acceptance of other things and justifies them. Knowledge neither has nor needs a foundation.

This is of some importance because Engel attributes to me the view that the principle of trustworthiness is a bootstrapping principle as well. He says,

There is bootstrapping in that to accept that P is to take oneself to be trustworthy in one's belief that P. Lehrer's idea is to use acceptance as a minimal epistemic attitude which can serve as a lever to obtain knowledge, but without taking it to involve knowledge, for otherwise the strategy would be circular.

Finally, noting correctly that I hold that acceptance of the kind that concerned me is a condition of knowledge, he says in conclusion,

If so, it must contain an evidentialist and a reliabilist component. Keith Lehrer admits that acceptance is based on a capacity to accept. What else can this capacity be, if not a disposition to believe that I am trustworthy? Hence my acceptance of my own trustworthiness must be based on my believing myself to be reliable in my acceptances. And this kind of belief cannot bootstrap itself into the status of a keystone of reason.

Acceptance, as I see it, aims at truth, but it is not a condition of acceptance that it be reliable. I acknowledge that unreliable acceptance will not take us to knowledge. Neither will reliable acceptance. What takes us to knowledge is the systematic justification of acceptance that is undefeated by errors in the background system used to defend the claim to knowledge. I agree that the acceptance of the principle of trustworthiness is necessary for personal justification of what one accepts, and the truth of the principle is a necessary condition of the conversion of personal justification to undefeated justification and knowledge. But the principle is no bootstrap lifting itself up miraculously and filling in for a foundation. We do not need a foundation for knowledge or a bootstrap for acceptance or justification. We need systematic defense of what we accept, aiming at accepting that p if and only if p in a way that is worthy of our trust. Engel is right to see that acceptance has the aim of accepting something in a way that is trustworthy. But there is no bootstrap or foundation in this role of acceptance. Our aim is to be right in a way that is worthy of our trust, not by luck, because that is needed to convert acceptance to knowledge. So acceptance has an epistemic goal, one that is not reducible to reliability, as the Cohen problem, the new evil genius problem illustrates.

Our acceptance must be trustworthy to lead to justification that converts to knowledge, undefeated justification. Such justification, however, requires an evaluation system, a background system used in defense of what one accepts. The acceptance of one's trustworthiness in what one accepts is a step on the path to knowledge, that is what makes it epistemic, and the truth of it is required to reach knowledge. The acceptance of the principle of trustworthiness is supported by other things one accepts, while at the same time supporting them. The mutual dependence and support shows why I prefer the metaphor of a keystone. Other things we accept support the acceptance of the principle of trustworthiness as the acceptance of it supports them. It has a special role in a system that yields knowledge from the right match between coherence and truth in what one accepts. It explains why we are reasonable in accepting it, provided it is true, at the same time that it explains the reasonableness of accepting other things we accept. It is the coherence of the system

that gives the principle the role of a keystone adequate for the arch of knowledge. Engel is right in seeking a connection between acceptance and being worthy of what one accepts. The connection is not, however, that one be trustworthy in what one accepts but rather this. The goal of acceptance is to accept that p if and only if p in a way that is worthy of one's trust. So being worthy of one's trust in the pursuit of truth is a goal of acceptance. The reason is that epistemic acceptance aims at knowledge, that is what makes it epistemic, but acceptance does not entail that you are trustworthy or that your trustworthiness is successfully truth connected. These are things we accept in the quest for knowledge, but the step of acceptance, though necessary, does not presuppose success.

2.3 Reply to Hannah Tierney and Nicholas Smith

Tierney and Smith have done a brilliant job of assembling my views concerning justification and the basing relation. I agree with their discussion and with their conclusion. Moreover, I like their revision of my views about evidence that pulls justification away from the basing relation, which they formulate as follows:

Lehrer seems to suppose that his requirement is a version of the basing relation, but it is really another relation altogether—we may think of it as the "justification–giving relation." This relation is embodied in an epistemic agent's own metacognitive assessment of his or her justification, which is entirely independent of what qualifies as the actual basis of the belief itself. To satisfy Lehrer's justification–giving relation, the agent must count the evidence that justifies his or her holding the belief in question *as actually justifying* his or her holding that belief (which is why the agent would appeal to that evidence when called upon to justify his or her belief).

I was inclined to offer the account of the justification–giving relation, as they describe it, as a version of a basing relation because there is some ambiguity in ordinary usage in the use of the expression "based on". It seems to me that Raco, for example, could claim that his medical view of the malady is based on his medical evidence, not on his racism, even if he must acknowledge the causal role of the latter in causing and causally sustaining his belief about the malady. However, I am now inclined to agree that greater clarity is achieved by representing my view as the view that the evidence that justifies a person accepting that p and knowing that p does not require that her belief that p is causally based on the evidence. So I appreciate their proposal.

I conclude my comments with two reflections. Tierney and Smith speak of being conscious of the reasons that constitute the evidence. That gives me a small qualm. I think that the person must accept the reasons that are the evidence of the agent and be ready to call them to consciousness in defense of the claim they justify, but she need not be consciously reflecting on the evidence for it to justify what she accepts. It suffices that she accepts it and would appeal to it to defend what she accepts against objections.

The other consideration concerns causality and the distinction between acceptance and belief and may explain the clash of intuitions my examples elicit. I would be inclined to say that Raco and the gypsy lawyer *believe* what they do because of their racism or superstition, but that is not the correct causal account of why they professionally *accept* what they do. They accept what they do because they follow cogent canons of reason and evidence. However, they must be trustworthy in the way that they accept what they do for the justification to succeed. Some of the conflicting intuitions may be due to the fact that people think that a person whose beliefs are formed and sustained because of racism and superstition cannot be trustworthy in what they accept according to the canons of reason. Racism and superstition, they think, undermine reason. I think that the way Tierney and Smith present and elaborate the examples of Raco and the gypsy lawyer show that this need not be the case.

2.4 Reply to Glenn Ross

The essay by Ross is well reasoned and, in fact, articulates a view that I would now accept concerning the relationship between acceptance of the negative lottery hypotheses, knowledge, and, therefore, the kind of justification required for knowledge. I would add a qualification concerning acceptance, which I distinguish from belief, in treating acceptance as being relative to an objective or goal the subject understands. I have focused, as does Ross in an admirable manner, on a kind of acceptance aimed at knowledge. Ross is right in noting that accepting that a ticket will lose in a lottery is not the kind of acceptance that will lead to knowledge. My explanation for this conclusion is that there is an objection to the claim one's ticket will not win that one cannot answer, namely, that that the winning ticket has the same probability, and, for all you know, your ticket may be that winning ticket. Objections to what one accepts that one cannot answer show one does not know.

However, there are other kinds of acceptance with other objectives and purposes, for example, accepting a hypothesis for testing. A kind of acceptance that is more interesting in terms of its goal is that accepting something with the purpose of accepting something that is more reasonable to accept than not to accept. There are things it is reasonable to accept, I propose, even if there are objections one cannot answer to what one accepts. One example is accepting that there is, has been or will be intelligent life elsewhere in the universe. When one reflects on the spatial–temporal vastness of the universe, it seems extraordinary that intelligence should have come into existence only on earth in an astronomically brief period of time. Scientifically considered it would be astronomically improbable. There is, however, an objection to accepting that there is intelligent life elsewhere in the universe, life elsewhere in the universe, life elsewhere in the universe.

namely, that we have no direct evidence of such existence. I conclude that it is reasonable for me to accept the claim with the goal of being reasonable, but it is not something I know, even if it is true. So what distinguishes this case from accepting that one's lottery ticket is a loser? One might suggest the size of the lottery, but I do not think that is right. I think, on the contrary, that one will not be led into inconsistency following principles of impartiality and consistency as principles of reason in accepting the intelligent life hypothesis as one will following those principles in accepting that one's lottery ticket is a loser. The life of reason requires consideration of the systematic results of what one accepts following the principles of reason however subject to revision those principles may be. Ross has done us a service in exploring the consequences of such principles in the life of reason.

2.5 Reply to Guido Melchior

There is much in Melchior's argument that I find admirable and insightful. However, there is also a misconstrual of my position. Melchior notes that I hold that we can give no proof that the skeptic is wrong because any attempt to do so, will beg the question against the skeptic. That position is consistent with offering a correct *explanation* of how and why we can know common sense claims to be true and skeptical hypotheses to be false even when we can offer no proof that the skeptic is wrong. Proof and explanation separate here. Moore, unlike me, claims to have offered a proof that the skeptic is wrong. I think, contrary to Melchior, that the proof is fallacious as an argument against the skeptic because it begs the question. A robust literature, which Melchior sites, has developed around the Moore argument. So the question remains concerning the merits of Moore's alleged proof.

There is a question as to what is meant by a proof in this context. I think that when Moore says he is offering a *proof*, he means that he is offering a sound argument that a skeptic, who doubts the conclusion, *ought* to accept and relinquish his doubt. I do not see that anything Moore says has the result that some skeptic who doubts that he knows *ought* to relinquish his doubt that he knows and accept that he knows. Moore has offered no proof, no argument that ought to convince the complete skeptic, because what Moore says, however right he may be, begs the question of whether we know.

Melchior argues that Moore has an advantage over my theory, which I am not sure I understand. First of all he says that the knowledge claims to which Moore appeals are immediate and can be known by a young child. I do not see this as an advantage. The issue of whether young children, who lack the conception of truth and do not understand what it means to be trustworthy in pursuing truth, know what Moore says is true, is one that is much discussed and is not worth more ink here. I distinguished between primitive knowledge, which children have, and discursive knowledge that requires that the knower to be in a position to distinguish truth from error to defend and justify a claim to knowledge. There are two kinds of knowledge. Discursive knowledge requires the activation of faculties the young child lacks. Adding up primitive knowledge will not take you to discursive knowledge, which I consider our distinctly human achievement and the basis of science. You need a different ability, a different understanding, and one that my form of the coherence theory was intended to explain.

I turn to Melchior's claims about knowledge, assuming they are ones about a kind of knowledge that requires justification and defense. This is a fair assumption, I think, because of Melchior's reference to the problem of easy knowledge. He concedes that this is a problem for Moore, and says it is a worse problem for me. It is not a problem for me. Melchior focuses on the principle for trustworthiness in my theory in a way that distorts the account of knowledge and justification I developed suggesting that all you need for justification is the principle of trustworthiness. My basic idea of justification, which Melchior presents, is that justification requires the ability to meet *all* objections, to answer or neutralize them. The objection that I am not trustworthy in what I accept is *one* objection that must be met for me to be justified, and accepting that I am trustworthy in what I accept and being right in this meets that one objection. My accepting that I am trustworthy in what I accept and being right in this is, therefore, a necessary condition of knowledge. It is not sufficient on my account, unlike the versions of reliabilism that create the easy knowledge problem. For there are other objections I must be in a position to meet on my account, and that avoids the problem of easy knowledge. The objection for Roxanne is that the gas gauge may be inaccurate and rereading it does not show it to be accurate, or, to put the matter as Melchior does, the objection is that she lacks any independent evidence the gauge is reliable. That is an objection Roxanne cannot meet and, as a result, she lacks undefeated justification and knowledge on my theory.

The role of trustworthiness is not a bootstrap for the justification of other claims. It is a necessary condition, a keystone condition, because it has a special role. It is not a foundation, and it is not bootstrap. It cannot support a structure of justification by itself, and it cannot bootstrap acceptance into justification. It is a keystone necessary condition in a system that enables one to meet objections. If there is the right systematic connection with truth, then the justification resulting from meeting those objection is undefeated, and one knows. The truth-connected system, not any single claim within it, is what gives you knowledge. The support for knowledge in the single case results from coherence within a system, from the way in which the system enables the subject to meet objections. The particular instance and general trustworthiness, a keystone in the system, stand or fall together without a foundation and without a bootstrap. You need a system to know a particular.

I would note that there is a kind of immediacy compatible with the account I have offered. The particular claim, though it must be defensible, need not be inferred from any general claim. My knowledge that I see a hand coheres with a system that justifies my accepting that I see a hand without inferring it from anything else I accept. Coherence is a capacity to meet objections within a system rather than an inference from what I accept therein. There is more to be said, as there always is, in answer to the interesting questions Melchior raises. I end with an expression of my appreciation to Melchior for challenging me to clarify my position with his perceptive argument. My conclusion is that the coherence theory avoids the problem of easy knowledge because you need coherence with a system to know—a system

that enables you to know that you are worthy of your trust and that your trustworthiness is successfully truth connected.

3 Practical knowledge

3.1 Reply to Nenad Miscevic

I find Miscevic conducting himself in a manner here that explains the esteem in which he is held by those who have the good fortune to enjoy his philosophical company. Not satisfied with the theory so far articulated, he asks probing questions and advances bold historically and analytically informed proposals. In his remarks here, I find that he has gotten hold of the account of wisdom I advanced. But there is an emphasis in my account, and the one I articulated with Smith, that may distort the outcome. It concerns the role of considering objections to one's preferences and having the capacity to meet those objections in terms of one's background evaluation system. Some objections will, of course, be one's concerning the consequences of preferences, what the results will be of acting on them.

I was using preference in a way that connected it with choice, though I may not have been explicit enough about that. If a person prefers A to B, then given a choice between A and B now, he must choose A. The *akratic* person may think he ought to prefer A to B when he chooses B, but on my account the person does not prefer A to B when he chooses B. Here there is a simple gap between thought and preference. So a person preferring A to B will not choose B so long as he has the preference. However, that point being made, there is a difficulty, Miscevic alludes to it, between evaluation and preference. The evaluation that A has more merit than B has to convert to preference for A over B, or else the person is not wise. So, in fairness to Miscevic, having closed the gap between choice and preference, there is one created between positive evaluation and preference. That seems to me the right place for the gap. A person may be full of insightful evaluations of what has more merit than another, but if those evaluations do not carry over into the formation of preference, the person may be smart but not wise.

This issue is rather central to the emphasis that Miscevic places on first order desires and beliefs. He is looking for an account that joins evaluation, no matter how correct, with choice in a way that leads to wisdom. A good talker, even a very good talker, may strike us as smart but not wise. Moreover, Miscevic has a point in suggesting that first order beliefs connect us with thought and action in a way that is important. I did not mean to deny this as I placed emphasis on higher-level evaluation of both desires and beliefs. They give us useful input information. However, they are unreliable in a way that anyone who experiences illusions, the bent stick is a favorite, the changing size of the moon is another, will acknowledge.

I want to conclude my remarks on Miscevic with a couple of brief remarks admitting that to answer all his questions would lead me to write at much greater length than is appropriate here. Firstly, the utility of first order beliefs and desires results from the way in which they are causally connected with truth and value. The causal impact of truth is acknowledged, as is the imperfect influence of it on belief. The same, I suggest, this is more controversial, should be acknowledged for value. Value, a kind of generic value, seems to be part of causal order. I first became convinced of the causal impact of value thinking about epistemic value, the evident character of something. I construct an argument to make something evident to another. If I succeed, the conclusion will be so evident, I hope, that it will cause the other to believe. I cause something to become evident, and being evident, it is part of the causal nexus. Similarly in morals, does anyone doubt that wickedness is without causal consequences or that one person cannot cause another to be wicked? The character of my friends causes my life to be full of value, and the value of my life causes me to seek to give value to the life of others. So the role of first order beliefs and desires is what information they give us, however, imperfect about what is true and what has value.

Secondly, however, belief and desire require evaluation to correct the manifest and multiple errors and imperfections of them. Desires and beliefs, when evaluated positively reflecting the right causal connection with truth and value, become preferences and acceptances. If the evaluation is sound and not based on error, then the positively evaluated desires will be the ones one prefers to satisfy and chooses accordingly. Moreover, such evaluations and preferences give rise to rules of them that act as guides concerning what desires to prefer and beliefs to accept without further higher order reflection as endorsed by the higher order system. The *phronesis* Miscevic considers important at the first level is a combination of the causal influence of truth and value on belief and desire and the higher order evaluation of beliefs and desires forming preference and acceptance. The background evaluation system yields principled responses, guided by rules of the system, leading us to prefer the satisfactions of some desires, the acceptance of some beliefs. When error does not defeat us, we will be wise and knowing.

To conclude, let me refer to the role of feeling. Evaluation without feeling and emotion may lead to a form of life that is defective. Here I concede a lack in my account. Feelings and emotions, like desires and beliefs, are important, even to the wise and knowing, and a more complete account, suggested in this volume by Borgwald and elsewhere by Konzelmann, would include an account of feelings and emotions, evaluated like beliefs and desires, leading to distinguish those that are unreasonable from those that are not. The evaluation will endorse the role of rational feelings and emotions in thought and action. I believe this accords well with what Miscevic concludes.

3.2 Reply to Kristin Borgwald

The essay by Borgwald is one that I much admire and wish I had written myself, though I would not, I suspect, have produced such a valuable essay as she has. She is good enough to conclude,

While the goal of my project is different from other Lehrer-influenced philosophical pursuits, the implications of his work on self-trust to issues of feminism and sentimentalism reflect some of his own concerns in his forthcoming work. Self-trust, he argues, is a condition of how we represent ourselves in the world and, we might add, how we, and especially, perhaps, we women, can reconfigure how we represent our emotions, our aspirations and our place in a reconfigured world of our own making.

I find the concept of epistemic personhood and the loss thereof under some conditions to be an excellent development of the role of self-trust. I am in such complete agreement with what she says that I have little to add.

The priority of epistemic personhood over moral personhood in action seems an important insight of hers. If a person does not consider herself worthy of her own trust in what she accepts, then she will not consider herself worthy of her own trust in what she accepts about moral matters, and she will be lost. I also agree that selftrust and the acceptance of one's trustworthiness is of crucial importance in the evaluation of emotions, which are an important source of information about ourselves and our relations to others. Moreover, just as acceptance of one's trustworthiness is a basis of reasonable evaluation of what one believes and desires, so it is the basis of reasonable evaluation of one's sentiments and emotions. So, I end my comments with a simple expression of appreciation of the extension and application of my views.

3.3 Reply to Danilo Suster

There is a brilliant lucidity to the account that Suster offers of my work and that of others. I think that his account is fair in what he ascribes to myself and Lewis with the possible exception that I had much earlier, in reply to Ginet, a position like the denial of the agglomeration principle. That is, it follows from my choosing otherwise that the conjunction of the laws and antecedent conditions would not have been true, but it does not follow that either if I had chosen otherwise the law statement would not have been true or that if I had chosen otherwise the statements of antecedent conditions would not have been true. All that follows is that if I had chosen otherwise the conjunction of the laws and antecedent conditions would not be true. Something would have had to be otherwise if I had chosen otherwise, but what would have chosen otherwise is open. However, that logical point being noted, we do need some account that explains, as Suster notes, the strength of our intuitions that the laws and antecedent must be as they are and, therefore, that the idea that we could have chosen otherwise must be an illusion. We need more than a logical trick to explain the source of the illusion and the intuition based on it.

My response was to think about the nature of laws. We ascribe a kind of necessity to them. So how could they have been otherwise? My answer is that a misconception of scientific laws is the source of the illusion. The function of scientific laws is to validate inference, including counterfactual inference, and that is what distinguishes a law from an accidental generalization. Now it is important in finding generalizations that warrant counterfactual inference to find ones that do not depend on human choice. As a result, paradigmatic examples of laws telling us that under conditions C you get result R are ones that do not depend on human choice. Laws of motion and chemistry become our paradigms of scientific laws such that the truth of the law does not depend on any human choice. But when we confront the case of scientific laws about human choice, the constraint that the truth of law does not depend on human choice becomes implausible. How can it be true that a result R of conditions C does not depend on human choice when R is a choice? So, if a person had chosen otherwise, then some law about what the person would choose would also be otherwise.

So, Suster might ask, why does the intuition that creates the paradox remain? I admit that in the two cases he describes, skepticism and determinism, I can shift my attention in a way that yields the skeptical or the hard determinist result. So why? I noted that there is a distinction between belief and acceptance, the latter being the evaluation of reason in terms of a background system of evaluation, the former persisting in the face of negative evaluation of reasoning. Both systems have their use and function, both their hazards and defects. One knows that if one had chosen otherwise, which one could have, then something else would have had to be otherwise, which it could have been. That is the reflection of reason aiming at giving a coherent and consistent account of ourselves in an imagined world of science. We recognize that theory construction and explanation is our doing, satisfies our aims, and one of those aims to construct an account of ourselves in the world of science that is coherent. For the sake of coherence, we need to give up some beliefs, even that human choices do not make any difference in laws about human choices.

Should we say that the impossibility of combining freedom with a scientific picture of the world is just an illusion? An illusion remains, like the illusion of the bent stick or the changing size of the moon, but there is a correction to the illusion in the use of reason that philosophers are strangely reluctant to accept. The correction is to realize that it remains open to us to explain the illusions away as we construct a coherent view of ourselves in our world, our world in ourselves. The illusion that laws do not depend on us is powerful, and, perhaps, comforting. The truth is that scientific laws and theories are our creations constrained by the test of experience, the ability to warrant counterfactual reasoning, and the power to explain our world and ourselves.

3.4 Reply to Konzelmann

Konzelmann offers an application of a model of consensus Wagner and I developed as providing an adequate account of institutional virtue while other models are problematic. She says,

Given that motivation for supererogatory action is neither inferred from statutory duties nor accommodates a right of reprobation, modeling institutional virtue on collective rationality or explaining it in terms of joint commitment both prove problematic. In a third step, I argue that LWC has the explanatory potential to account for institutional virtue. Due to its main features, iteration and evaluation, it provides a non-trivial analysis of continuity and thereby satisfies basic constraints on the notion of genuine institutional virtue.

She provides an illuminating account of discussions of institutional virtue others have offered which I do not have the space to pursue here. Given that individuals in

a group have assigned weights of respect, weights they are willing to apply to aggregate their position and that of others, individual and group virtue coincide in the aggregate combing individual virtue with group virtue as the individual preferences and group preferences converge in the process of iterated aggregation. It is an elegant application of the model.

I would like to offer an extension of the application of the model. First of all, it seems to me that if the individuals are to be worthy of the ascription of virtue the institution exhibits, they must be connected with other members of the group in the appropriate manner. I suggest that the model offers a model of connection, namely, the condition of connectedness that would yield convergence toward consensus in the iterated aggregation. The condition is quite minimal. It is that there be a vector of positive weights assigned by individuals, each individual assigning positive weight to the next person in the vector, which connects all the individuals in the group. So if the members of the group were arranged numerically in a sequence, and each person assigns positive weight to the next person in the sequence, as well as positive weight to herself, the vector of those weights will connect the members of the group. Note as a special case, exhibiting the potential role of a leader, that if all the members of the group assign positive weight to one person in the group, a leader, and that person assigns positive weight to each member of the group, that will suffice to connect the group assuming self respect. The importance of the connectedness condition is that it insures that iterated aggregation will lead to convergence if it is satisfied. So the condition has an important functional role. Moreover, such connectedness makes it possible to compute a vector of weights, which if applied to the original preferences, will find the point of convergence.

Thus, the model has this merit. Given connectedness, we can speak of an implicit consensus, which is a mathematical consequence of the weights members of the group assign. So even if the iteration does not occur, the initial assignments of positive weights might carry justification for the ascription of virtue toward the position that would constitute the point of convergence, especially if actions of members of the group reflect an implicit awareness of that point. I hope that these remarks may offer some further suggestions for research on this innovative application of the model of consensus to issues of institutional virtue.

4 Philosophy of language

4.1 Reply to Adrienne Lehrer

My reply to Adrienne Lehrer can be a brief expression of gratitude for her excellent summary of the work that we have done together. Her attempt to combine social and individual, synchronic and dynamic, formal and empirical components in a theory of word meaning with such brevity is extraordinary, and, I hope will provoke others to develop it. The model of consensus that is presupposed here took the informal suggestion of Ziff, whose work on understanding utterances was so innovative, concerning vectors to a level of formal sophistication that I think would have surprised him. I find myself greatly indebted to an experimental linguist, Adrienne

Lehrer, on one hand, and a philosophical mathematician, Carl Wagner, for the cooperative assembly of a powerful model of sense and reference in terms of vectors of weights of respect of individuals in a language community. I should like to add a comment that might clarify the difference between our model and other semantic models of word meaning. It is tempting to try to extract meaning from denotation across possible worlds. This might be called the standard model theoretic view of the matter. But it is empirically untenable in a way that the formal structure of our model of meaning in terms of vectors of sense (word-word relations) and reference (word-thing relations) illustrates. The vectors of sense are not determined by the vectors of reference in the actual world, as everyone concedes, or across possible worlds. First of all, there are semantic relationships, like antonomy, that are part of the sense of a word but not something implied by the consequence relation. So even if reference across possible worlds captured the consequence relation, that would not yield all semantic relationships. Moreover, and more importantly, as Daniel Andler noticed when we first presented our model in Paris, the vectors of reference and the vectors of sense might not yield consistent results, even for a single word with a single meaning. Consider a word, noted by Adrienne, as she studied container words; "coffin" has a strong implication of "container made of hard materials for burial of a corpse". However, someone witnessing a scene in which a group of people with feathers in large supply and hard materials in short supply wove together feathers to use as a burial container would not hesitate to say that the container was a coffin. This simple example illustrates our view that relations of sense and reference are indeterminate, a matter of degree, where the degrees of these relations in a communal language are determined by social consensus which results from weights individuals give to the usage of other individuals. The result is that there is a fixed point in the aggregation in terms of weights. Once the divergence between sense and reference is acknowledged as well as the indeterminacy of relations of sense and reference, the social information is accommodated in vectors of respect.

4.2 Reply to Alfred Schramm

Schramm's sympathetic presentation of the semantic model Adrienne and I have constructed and acute observations concerning it reflects his philosophical depth, which I have so been fortunate to enjoy over the years in Graz. He combines his approbation with some important critical remarks I wish to answer. For example, he says of the model,

it explains beautifully, but it doesn't predict—at least not predict facts of the same kind as the ones it explains.

My reply is that starting from hypotheses about what individual vectors of reference are like, the test is the confidence level of applying the term, and vectors of sense of an individual are predictions of how probable it is that a person will draw inferences. The tests are the confidence levels. Adrienne has done a good deal of research along these lines, and her work convinced me the views of Quine about degrees of semantic connections are empirically confirmed. Schramm raises some doubts about how the empirical linguist can distinguish between sense and reference. The question is one about how to operationalize the distinction to study the two. One way to do so is to test response to sentences containing the two words, for example, the degree of acceptability of the sentence, "Coffins cannot be made of feathers" while the test of the degree of acceptability of referring would be tested by asking a person shown a container made of feathers how acceptable it is to call it a coffin. Our suggestion is that answers to questions of two sorts, or modifications thereof to refine the issue, might give different degrees of acceptability, illustrating the way in which vectors of sense and vectors of reference might fail to match.

On the point Schramm raises about analyticity, Quine has often remarked that there are extreme cases of sentences that we are extremely unlikely to give up. However, I think that Adrienne has found empirical evidence in favor of analyticity, though something short of a proof. Subject's judgments of extreme degrees of acceptability in extreme cases are more stable than judgments of intermediate degrees. This suggests that there is a psychological basis for judgments of analyticity, extreme acceptability, than other judgments of degrees. In conversation, Quine seemed to agree that the question of the scientific cogency of notions of analyticity should be made empirical. His objection, I believe, was to the traditional view that the distinction is based on a necessarily true a priori principle of demarcation. I prefer to leave the dispute between Quine and Carnap on the issue unresolved, for it appears to me that one was discussing the empirical tenability of analyticity in natural languages and the other the theoretical tenability in an artificial language useful for science.

To end on a positive note, I would agree that switching from idiolect to idiolect, from one language to another, is an important dynamic and allows for an adequate degree of communication, which may involve considerable overlooked miscommunication. However, there is a need to resolve misunderstanding, and the assumption of a shared language, a communal language, carries with it the means for resolution. Some people are given more authority than others in the use of parts of the vocabulary, and, of course, there is always the appeal to a dictionary. Moreover, the quickness of communication, of understanding what another says, suggests a shared communal language, even if there is a fictional component therein. The nonfictional, and empirically based ground of the fiction, is a similarity of sense and reference in idiolects to one another, and the basis for coping with problems arising from differences causing confusion is the respect we given to the usage of others that enables us to form consensus.

There is clearly something important in Schramm's idea of our ability to manage a good deal of idiolect and language shifting without confusion. We think some linguistic authority, even if only an empirically grounded mathematical average, is useful for the resolution of differences. If one embraced a dynamic view that—when it comes to meaning anything goes—we would be lost in confusion. I do not suggest that is what Schramm is proposing. We have attempted to construct a model of an empirically grounded fiction, like the average person, with the average weighted by the respect we assign to each other. We seem so much in agreement with Schramm and so greatly value his appreciation of our work, that I suspect what separates us may be the value I attach to formal models for explanation. The interest in maximizing explanation is not one I seek to establish. It is a motive for my philosophizing.

5 Philosophy of mind

5.1 Reply to Johann Marek

I greatly value and appreciate Marek's study of the similarity between my own views and those of Meinong. It reveals his combination of historical knowledge and analytic insight from which I have so long benefitted in my thought and writing. Marek is my hidden muse. I find his account of Meinong on consciousness very close to my own, and, having failed to notice that, I am glad to acknowledge his precedence.

There is a question about conscious states that are unexemplarized. I comment that reflection and introspection convince me that there are somewhat confused initial states of consciousness that precede normal cognitive functioning and, therefore, representation. Marek asks how I can know of the existence of such states if there is no representation of them at the time they occur. My answer is simple, by memory. There is an indistinct and confused memory of states occurring before I know what is going on or what they are like. I think that this is important and will occur again in the discussion of Fürst. There are states occurring in us that are conscious and can be used as vehicles of representation that are not immediately used in that way.

Marek, in a probing discourse about paying attention to unnoticed qualities, remarks:

In order to notice a conscious experience (or a feature of it), the conscious experience has to be exemplarized, and in order to be exemplarized, the conscious experience has to be noticed. I cannot exactly see how Lehrer's representational loop escapes this dilemmatic circle.

My reply is that paying attention will, in a normal state, consist of exemplarizing the particular quality. Exemplarizing is an activity of generalizing from the exemplar and marking a distinction between what is in the generalized plurality and what is not.

I suspect that the discussion leading up to his question in which he asks about conscious experience and properties is the source of the difficulty. What exists in conscious experience is an individual state with an individual quality. Attending to what general kind of thing that conscious state is, noticing what it is like, involves exemplarizing it to construct a general conception of what kind of thing it is. It is we, as Reid averred, who sort individuals and individual qualities into kinds and properties, not nature. We are directly acquainted with individuals from whom we form a general conception, including that of properties. We then mistakenly infer, as I believe no less a philosopher than Russell did, that since we are directly acquainted with individuals, we are directly acquainted with the properties we conceive in

terms of them. It is a natural enough mistake to make, confusing the awareness of the individuals with the way we conceive of them. It is a mistake, nonetheless. I do not attribute this to Marek. I mention it in defense of my view.

After proposing strong similarities between Meinong's theory of self-presenting states and my theory of exemplarized states he concludes,

Exemplarization, in the sense of exemplarizability seems to be built into the notion of a state's being conscious and can be seen, then, also as a mark of the (conscious) mental.

Flattered as I am to be associated so closely with such a giant of Austrian philosophy, I must end on a note of disagreement. I agree that we can exemplarize conscious states, that seems to me to be a feature of them, but when Marek proposes that

Conscious states are necessarily exemplarizing states and exemplarizing states are necessarily conscious states

pressing my view close to Meinong, I must decline the compliment of the association. I regard it as our special capacity to exemplarize to form general conceptions and what we exemplarize is not necessarily confined to conscious states.

Marek notes that I have said particulars other than conscious states can be used as samples and hence exemplarized. He thinks I should not hold this view because I should restrict exemplarization to a form of conceptualization that does not presuppose any antecedent conceptualization. I suppose an advantage of that view is that one could view such conceptualization as the empirical foundation. I denied the need for a foundation of justification, and for similar reasons, I would deny the need for a foundation of conception. You need a system even if, as I would claim, the reflexivity of exemplarization provides a special truth security within that system. He suggests that exemplarization is the mark of the mental. I am more cautious. Conscious states, for all we know, may be material particulars. I am not convinced that they are. I am agnostic on the issue of token-token identity of the conscious states with material states. I am not prepared to agree that exemplarized states are necessarily conscious states. I will rejoin the issue in discussing Fürst. I close by noting that Marek, following Meinong, wants to attach the capacity to be exemplarized to the conscious state, while I want to attach the capacity to exemplarize conscious states to a more general capacity we have to generalize particulars, sorting them into kinds.

5.2 Reply to Leopold Stubenberg

The analysis that Stubenberg gives of my account of exemplarization and consciousness is so precise and exact that I must begin by thanking him for his fine attention to the details of my account. He notes, correctly, that my view separates from standard functionalism, when he concludes.

Lehrer begins with a role occupant with a quite particular qualitative nature a pain, say, or a sensation of blue—and the (inferential) role that this item plays upon its exemplarization is determined by that item's intrinsic qualitative nature.

I agree with this with one caveat, namely, that the intrinsic qualitative nature is the individual quality of the state. That appears consistent with what he says.

He then comments in disagreement with my claim that, having explained how we know what our conscious states are like by the theory of exemplarization, the only problem left concerning consciousness is a scientific problem, by which I meant the scientific problem of explaining why they exist. Stubenberg asks,

That may well be so, but none of this addresses the hard question: how is it possible for pain qualia to arise from purely physical conditions?

What puzzles me is why he thinks this is not a scientific question. The answer is that he thinks it is a metaphysical theory, for he says concerning my theory of exemplarization:

The theory does not explain the place of qualitative consciousness in the physical world. This metaphysical question has survived Lehrer's onslaught unscathed.

However, he then goes on to propose rescuing my claims by appealing to a proposal made by Sellars and articulated by Galen Strawson to the effect that the particular qualities of consciousness might turn out to be qualities of matter. Since they obviously do exist, we know that they exist and, by exemplarization, something of what they are like, a materialist theory must accommodate such knowledge even at the expense of revising his theory of matter. I am inclined to agree. Moreover, I remarked in the earlier article, somewhat in debt to Dretske, that conscious states, including those of particular qualities, are one way that we receive information about the world and ourselves. I added that it is a scientific question why we receive information in this way on the analogy with the question of why we nourish ourselves the way we do rather than in some other way. I think the question of why we receive information from the particular qualities of consciousness rather than in some other way is a hard scientific question. However, I am inclined toward Quine's views to the effect that we do not have a sound principle of demarcation between the necessary and the contingent and, therefore, between the metaphysical and the scientific.

I have long had a suspicion that will become more salient as I proceed, that what convinces people that there is a hard metaphysical problem about consciousness over and above any scientific question is a metaphysics of the existence of properties and the claim that some such properties are necessary to the existence of conscious states, namely, phenomenal properties. This creates the problem of how such properties could be properties of matter. Richard Taylor had a reply like the Sellars–Strawson view, namely, that a materialist, Taylor claimed to be one, should not assume in advance of consulting experience what the properties of matter are. My line is closer to Quine and his nominalism. Qualia are as I experience them individual qualities. When we generalize over individuals, as we must in science, we should allow ourselves the freedom to generalize in ways that maximize explanatory coherence. We must respect the reality of individual qualities of consciousness in our experience, but we do not face the problem of accommodating some strange properties that generate incoherence. For properties do not exist, only individuals and individual qualities exist. As we generalize to form the general conceptions of science, the individual qualities of conscious states may puzzle us, they may confound us, but how we conceive of them as we generalize over them may respect their individual character at the same time that we form general conceptions to maximize explanatory coherence. They present us with a hard scientific problem that will be multiplied by the assumptions of the existence of properties and the necessary possession of them. Scientific explanation is difficult enough without that metaphysical complication. It is not a complication I ascribe to Stubenberg in his careful and judicious study.

5.3 Reply to Martina Fürst

Fürst, like Stubenberg and to some extent Marek, argues, in great detail and admirable precision, that my argument that we know what our conscious states are like by exemplarizing them, fails to explain how we can be deceived in thinking that zombies, our functional twins who lack consciousness, might be impossible and token–token physicalism correct. She writes correctly characterizing my view,

When we think of ourselves being, for example, in pain, we exemplarize a pain-state, whereas in the zombie-case we think of the very same state in physical-functional terms. Since phenomenal concepts gained by exemplarization are ineffable, no descriptive concept can express what is exhibited in the exemplar. Therefore, these two different ways of conceptualizing one and the same state explain why we can conceive of zombies even if, in fact, they are impossible.

But she objects, quoting me first to use my words against my conclusion,

Ostensive exemplarization of an exemplar of experience to stand for other experiences that are like it may show us in some special cases all we know about what the experience is like.

She concludes,

The situation then is the following: in conceiving of an alleged physical state in terms of exemplarization we get access to ineffable, but essential, properties of this state, but not to its fundamental nature. But in conceiving of the same physical state in terms of physical concepts we get insight into its fundamental nature, but conceive of it lacking those essential properties. This strikes me as a puzzling outcome that fails to support the claim that the relevant concepts involved are co-referential. In short: the key point of my argumentation is that conceiving of experiences in physical terms and in terms of exemplarization gives us access to different, but *essential*, properties of experiences which in the case of co-reference should be connected, but in fact are not. Therefore, I conclude that exemplarization cannot offer a satisfactory explanation of the conceivability of zombies which supports physicalism.

I quote her in detail to show the clarity of her challenge. My reply is essentially simple. What we experience are individual qualities of experience, which are fairly enough called phenomenal, but it is not properties we experience. It is individuals and individual qualities. So there are no essential properties of conscious experiences to create difficulties. Moreover, some individuals that we exemplarize are clearly physical, samples of paint, movements in dance, are examples. Both Marek and Fürst interpret what I have said about exemplarization to imply that only conscious experiences can be exemplarized. That is not my view, and I do not find it plausible. The exemplarization of conscious exemplars makes them a vehicle of representation that exhibits what the state is like. The exemplarization of other things, a sample of paint, a dance movement, makes them a vehicle of representation that exhibits what they are like. So the exemplarization of conscious states does not entail that they are not physical. I do not take this as an argument that physicalism is correct. My claim is that when we imagine zombies being in the same physical states that we are in, and we imagine that they do not exemplarize their states, we may be imagining something that, for all we know, is impossible. For exemplarizing may be a physical process, and, if we do it, and zombies are physically the same, then so do they exemplarize, contrary to what we imagine. If the individual qualities of conscious states are physical, and if exemplarizing is a physical process, then, contrary to what we imagine, the zombies imagined to be physically identical to us but not conscious might be impossible.

This reply to Fürst might be further clarified in terms of my reply to Stubenberg. The individual qualities of conscious states may be initially represented in terms of how we experience them, and our conception of them may not extend beyond individual qualities of experience. But such initial conceptions are not the end of conception. We form general conceptions of the world to maximize scientific explanation—to explain as much as we can and to leave as little unexplained as we must. That is my chosen perspective. The importance of the exemplarization of individual qualities of experience as a vehicle of representation into our representational system to insure an inseparable connection between experience and representation. Fürst would like to conclude from my remarks about the exemplarization of conscious states that the resulting conception is all there is to them. That does not follow, however hallowed the tradition.

We confront a world of individuals, individual qualities included, and sort those into kinds by generalizing. Exemplarizing in an ostensive manner so that the individual quality reflexively refers to itself as well as other things is one way of both generalizing and marking a distinction with individual qualities of consciousness. It does not preclude other ways of conceptualizing those individual qualities, for example, as being token–token identical to physical states. One purpose for such conceptualizing is to explain the connection between the individual qualities of experience and the physical world. Of course, as Stubenberg noted, that means that we shall conceive of the physical states as having the individual qualities, the felt qualities, of experience. Some think that is odd or mysterious. I do not. Our including those individual qualities of experience as exemplar representations provides the needed connection between experience and theory in science. I do not see how that path leads to the conclusion that the individual qualities of experience are physical states or that they are not. That we can imagine that they are not, which the zombie argument illustrates, does prove that they are not. We can imagine many things that do not exist and could not exist.

One advantage of the notion of exemplarization is that it serves to explain how we can imagine zombies. Even if individual conscious qualities are physical qualities, we can think about them by exemplarizing them ostensively without generalizing further to conceive of them in the class of physical qualities. The fact that we can think of them, in terms of ostensive exemplarized concepts, without thinking of them as physical, does not demonstrate that they are not physical.

This much should be admitted to Fürst and to Stubenberg and Marek. Including our ostensive exemplarizations in more general conceptions of the physical is an extension of our conceptions of the physical to include more than the conceptions of present physics. How to do that is the scientific problem yet to be solved. It is, however, simply the problem of how to include experience in our representation of theory. Theoretical conceptions of the world that include ostensive exemplarizations must explain the function of individual qualities of experience in representation. I conjecture that they provide a special way of processing information about the physical world including the individual qualities of experience as a part of the world, but explaining that is a hard problem, and an unsolved scientific problem. Does that mean that it is not a philosophical problem? It is the most interesting scientific-philosophical question of the present stage of inquiry.

5.4 Reply to Joseph Tolliver

I have profited from ongoing discussions with Tolliver for a long time and thank him for his illuminating and dogged attempts to make me improve my work and correct my errors. After a detailed and accurate account of exemplarization containing an enhancement of that I wrote about the special role of the exemplar as an exhibit of what it is used to represent in exemplar representation, Tolliver says,

Lehrer's suggestion seems to be that some things cannot be represented in a way that fully conveys what they are like until they are experienced.

He amplifies this interpretation of my view contrasting it with the view that Mary has a new mode of presentation of just the same physical state completely represented by her antecedent physical description:

The position expressed above is different. Mary's theory contains no representation of what her experience of red is like, at least none that represents it in its full particularity, none that represents it just so. Her experience is a physical/functional state, but no description can represent that state just so. So, Mary's prior theory is not so much false as incomplete, and necessarily so, for no descriptive theory, materialist nor immaterialist, will do.

He adds,

The subject needs no prior notion of what the exemplarized experience is like to become aware of what it is like by means of the exemplar concept of that experience. (...) In being directly referring, exemplar concepts display the first characteristic of the ineffable: They are introduced as new subject matter (and vehicle) of thought or talk by mechanisms that exhibit them.

An important characteristic is lucidity:

Features of an experience that S fails to notice cannot be part of the application conditions for an exemplar concept that includes that experience, i.e., the application conditions for an exemplar concept cannot outstrip a user's ability to exemplarize them.

With all this I agree and express my gratitude for the clarification of my position. This account of my position is an improvement I am happy to embrace. My only qualification is that I think that ostensive exemplar concepts, which are the directly referring and lucid ones, can become part of other concepts; they can add something to the representational character of the concepts of physical theory. I conclude that the addition is a representational looping conception that insures truth as it loops back onto itself, and, when added to the concepts of physical theory both elaborates the character of the physical in terms of individual qualities of experience and secures a truth connection of a representational vehicle of the physical theory.