

The curious case of the self-refuting straw man: Trafimow and Earp's response to Klein

Theory & Psychology
2016, Vol. 26(4) 549–556
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0959354316648230
tap.sagepub.com



Stanley B. Klein

University of California, USA

Abstract

In their critique of Klein (2014b), Trafimow and Earp present two theses. First, they argue that, contra Klein, a well-specified theory is not a necessary condition for successful replication. Second, they contend that even when there is a well-specified theory, replication depends more on auxiliary assumptions than on theory proper. I take issue with both claims, arguing that (a) their first thesis confuses a material conditional (what I said) with a modal claim (Trafimow and Earp's misreading of what I said) and (b) their second thesis has the unfortunate consequence of refuting their first thesis.

Keywords

history of science, nature of theory, philosophy of science

In their critique of Klein (2014b), Trafimow and Earp (2016) identify two points of concern:

First, as we have shown, having a well-specified theory is not a *prerequisite* [emphasis added] for having replicable findings; hence the blame for apparent replication failures should not be placed upon ill-specified theories. And second, when there is a *relevant* [emphasis added] theory, experimental predictions depend much more strongly than Klein (2014) seems to appreciate on auxiliary assumptions, as opposed to on the theory proper. (Trafimow & Earp, 2016, p. 545).

Corresponding author:

Stanley B. Klein, Department of Psychological and Brain Sciences, 551 Ucen Road, UCSB, Santa Barbara, CA 93106, USA.

Email: Klein@psych.ucsb.edu

These are curious concerns. Well-formed theories specify the conditions of their falsification as well as criteria for determining the empirical relevance of many—but not all (see Note 6)—auxiliary assumptions. Thus, it is a truism that, absent such a theory, an investigator cannot be sure of what is to be replicated (e.g., Newell, 1973).

Rather than restate my thesis (Klein, 2014b), I focus herein on Trafimow and Earp's two points of concern. I argue that the first (see first section) confuses a material conditional (what I said) with a modal claim (Trafimow and Earp's misreading of what I said). The second (see second section), has the unfortunate consequence of refuting their first thesis. I discuss each in turn.

Erecting and tearing down a straw man

Klein (2014b) took the position that psychology generally, and social psychology in particular, lack well-specified scientific theories. One consequence of this, I argued, is that in the absence of well-specified scientific theory, replication attempts can be seriously compromised.

Since words matter when philosophical argument takes the form of verbal statements (as opposed to mathematical or logical formalisms), I took care to state what I meant by a "well-specified theory." Drawing largely on work by Margenau (1950), Fodor (1968), and Torgerson (1958), I defined "well-specified theory" as consisting in a set of propositions "capable of clearly linking physical observation to a well-formulated, conceptually sophisticated, and rationally integrated set of abstract constructs—thereby enabling computationally rigorous predictions (as well as conceptually satisfying explanations)." Continuing, "Absent such a *guide*, we have no way of knowing whether earlier studies are commensurate with, or antithetical to, whatever studies are presently under examination" (Klein, 2014b, p. 332).

I then discussed the paucity of theories in the social priming literature (the major battleground in the current psychological "replication wars"—although I made clear that my argument was not restricted to this particular domain) that meet these criteria. For example, since most social priming theories only permit deduction of the binary outcome "effect present/effect absent," they fail to provide the predictive precision expected from a well-specified scientific theory.

Despite my care to avoid claims about necessity (or sufficiency), Trafimow and Earp (2016) saddle my definition of well-specified theory with an implicational structure that goes well beyond what I stated. Specifically, they assert that according to Klein (2014b) such theory is a *necessary* condition for the proper conduct of a replication effort: "a well-specified theory is not a *prerequisite* [emphasis added] for having replicable findings" (Trafimow & Earp, 2016, p. 545).

Having effected this conceptual makeover (i.e., insertion of an "if and only if" condition), they take aim and refute their self-generated straw man by citing examples of successful replications unaccompanied by well-specified theory (but see the subsection "Are Trafimow and Earp's examples counterexamples?"): "as we have shown, having a well-specified theory is not a prerequisite for having replicable findings; hence the blame for apparent replication failures should not be placed upon ill-specified theories." (Trafimow & Earp, 2016, p. 545).

For the sake of argument, let's accept Trafimow and Earp's claim that Klein (2014b) argues a well-specified scientific theory is necessary for successful replication.¹ A necessary condition is a prerequisite either as a formal or as an informal axiom. The former case yields tautologies, as in the axiom of equality ($X = X$). The latter is axiomatic by way of self-evident implicature: theoretical condition X implies observation Y . Trafimow and Earp thus commit the elementary mistake of confusing the replication of an effect with the replication of the theory-driven method whereby the effect is brought about.

But, this is moot, since Klein (2014b) made no claims about necessity (or sufficiency). I argued only that theoretical specification provides *guidance* (see the above quote from Klein, 2014b, p. 332) in identifying conditions essential for a replication effort—not that well-specified theory was essential to identifying those conditions. Specifically, “A well-conducted replication requires, at a minimum, that the *essential* conditions of the study match those of the to-be-replicated study as closely as possible” (Klein, 2014b, p. 328).² I continued: “the class of essential conditions required for a successful, quantifiably predictable test of a scientific theory is [i.e., can be – not “only can be”] specified by abstract principles embodied in theory” (Klein, 2014b, p. 328).

Thus, while I accord theoretical specificity an important role in identifying conditions essential for a properly conducted replication effort, I never claim that theory plays an essential role. A well-specified theory is a potent—but not the sole—means by which one can re-instate the essential conditions for a replication attempt.

More formally, Trafimow and Earp (2016) maintain that their counterexamples show that $\neg P$ (poorly-specified theory) can be accompanied by Q (successful replication). From this they infer that P (well-specified theory) is not necessary for Q . But, a defective theory implies nothing!³ Moreover, the truth or falsity of their conclusion is *irrelevant to my stated position*.

Nor does Trafimow and Earp's creative reformulation warrant inferences about logical sufficiency. A reasonable reading of their argument is that counterexamples demonstrate that $\neg P$ (poorly-specified theory) is not sufficient for $\neg Q$ (unsuccessful replication). From this they infer that $\neg P$ cannot cause $\neg Q$. But, this follows only if they also assume: if $\neg P$ is not sufficient for $\neg Q$, then $\neg P$ cannot cause $\neg Q$. And it's clear why we should reject this assumption: it is an instance of the general, but fallacious, principle that one can infer that $\neg P$ doesn't cause $\neg Q$ from the fact that $\neg P$ is not sufficient for $\neg Q$ (e.g., an event often has multiple causes, none of which are sufficient on their own to bring about the event).

Indeed, it is easy to fashion cases in which X causes Y without X being *either* a necessary or sufficient condition for Y . For example, the finding that some heavy smokers (X) enjoy healthy lives (Y) does not sanction the conclusion that X can have no causal role in one's health (Y). Or, borrowing an example from Laudan (1990), because surgery (X) is not always necessary or sufficient to cure gall stones (Y), it does not follow that surgery is never a useful means for treating gall stones.

In summary, Trafimow and Earp (2016) first change my argument from an if-then claim (a material conditional) to a claim about necessity (a modal claim), and then set out to defeat their re-formulation. But Trafimow and Earp (2016) and Klein (2014b) are asserting different things.

Are Trafimow and Earp's examples counterexamples?

To support their argument, Trafimow and Earp (2016) turn to the history of science for instances of poorly specified theories associated with replicable outcomes (e.g., phlogiston and aether theory). Such “counterexamples,” they claim, demonstrate that “important and replicable findings” (Trafimow & Earp, 2016, p. 541) can be obtained even with poorly-specified theory.

While there are a number of unaddressed issues surrounding Trafimow and Earp's choice of “counterexamples” (e.g., naturally occurring regularities versus regularities based on theoretical deduction, differences between theory in the physical versus social sciences,⁴ and so forth), discussion of these problems would take us far afield (and push my commentary well over its word count limit). Accordingly, I restrict discussion to whether phlogiston theory—the “counterexample” to which Trafimow and Earp (2016) devote their most sustained attention—supports their claim that poorly specified theory can be associated with replicable outcomes.

Trafimow and Earp state that the theory of phlogiston is

a blatantly wrong and ill specified theory—at least from the perspective of hindsight ... — which nevertheless dominated the field from approximately the late 17th century to the late 18th century ... this theory held that the fire-like element of phlogiston was responsible for combustion, although the specific nature of this relationship was never precisely articulated. Nevertheless, despite this lack of specification, researchers were able to demonstrate—and replicate—the existence of oxygen (wrongly considered to be “dephlogisticated” air), nitrogen (“phlogisticated” air), and other major elements. (2016, p. 541)

As Trafimow and Earp see it, phlogiston theory is a prime example of a poorly specified theory that enabled successful replication.

But what sanctions their assertion that phlogiston theory was “ill-specified”? The opposite case can (and often has) been made. Ladyman (2011, p. 98; see also Chang, 2010), for example, provides a *partial* summary of the observational regularities subsumed by phlogiston theory:

- Metal + heat (in air) → calx [metal oxide] + phlogiston [de-oxygenated air]
- Calx + charcoal (source of phlogiston) → metal (+fixed air [carbon dioxide])
- Metal + water = calx + inflammable air
- Water = inflammable air [hydrogen] = dephlogisticated air [oxygen]

These empirical phenomena can be explained by phlogiston theory. A few examples:

- Metal = calx + phlogiston (explaining what it is that metals have in common)
- Charcoal = fixed air + phlogiston

Phlogiston theory even made novel predictions, such as the existence of new acids (e.g., formic; Scheele, 1931).

Thus, contra Trafimow and Earp's assertion, phlogiston theory is a decidedly questionable instance of the type of counterexample Trafimow and Earp (2016) require to

support their first thesis. If anything, the history of phlogiston is a case study in which a well-specified relation between phenomena (theory and empirical outcomes) is preserved in subsequent science even though aspects of the ontology of the theory are not (i.e., the replacement of phlogiston by oxygen). In fact, phlogiston and oxygen theory make virtually identical predictions if one assumes phlogiston is negative oxygen (e.g., Chang, 2010; Wisniak, 2004). Negative oxygen, however, requires negative mass, a property that makes the phlogiston harder to accept.

In summary, despite Trafimow and Earp's claim that "important and replicable findings" can be obtained despite poorly specified theory, the theories in which phlogiston (or aether) were explanatory were, in fact, well-specified. This is why failure of their predictions led to their abandonment. For example, objects taking on phlogiston should show a change in weight, but don't. Trafimow and Earp thus confuse a "wrong" theory with a "poorly specified" theory. Ladyman (2011) is adamant on this point: "Phlogiston theory identified a number of real patterns in nature and it correctly described aspects of the causal/nomological structure of the world as expressed in the unification of reactions into phlogiston and dephlogistication" (p. 100). A similar conclusion is reached by Chang (2010):

Some people think that phlogiston theory deserved to be consigned to the dustbin of history because phlogiston was just an imaginary entity, not based on anything empirical. This is a basic misconception, as phlogiston had some *detailed* [emphasis added] links with observed phenomena and with very concrete practical operations. And Lavoisier's theory relied essentially on caloric, the material fluid of heat, which was just as unobservable or hypothetical as phlogiston ... Lavoisier's theory won because it was inherently simpler than phlogiston theory. (p. 57)

The take away message is that oxygen theory replaced phlogiston theory *not* because the latter was poorly specified, but because it entailed some controversial assumptions (e.g., negative mass) and made some predictions that were shown to be false. Phlogiston also fell victim to Occam's razor (i.e., it unnecessarily complicated things).

Thus Trafimow and Earp make the error here (and in other "counterexamples"⁵) of confusing poorly specified theory with well-specified, but incorrect theory. Phlogiston theory (like aether theory) played the part of all well-specified scientific theories: it specified the conditions of its refutation.

Auxiliary assumptions and the self-refuting thesis

Drawing on what appears a variant of the Duhem–Quine thesis—that is, the proposal that most scientific hypotheses only make testable predictions relative to the background assumptions (or auxiliary hypotheses) that tie them to the evidence—Trafimow and Earp (2016) argue that "even in the case where there is a clear theory to draw upon, it is important to remember that empirical predictions come from the combination of a theory and auxiliary assumptions rather than from a theory alone" (p. 542). They further assert that auxiliaries actually are *more* important to replication efforts than is theory proper: "experimental predictions depend much more strongly than Klein (2014) seems to

appreciate on auxiliary assumptions, as opposed to on the theory proper” (Trafimow & Earp, 2016, p. 545)—although their reasons for this preferential ranking are never made clear.

Scholarly discourse on the precise nature of auxiliary assumptions is somewhat disputatious (even Quine and Duhem did not hold identical views—for example, Duhem, unlike Quine, maintained that his arguments did *not* apply to the social sciences [see also Note 3]; for reviews see Klee, 1997; Laudan, 1990). Accordingly, it is important to understand how Trafimow and Earp (2016) conceptualize auxiliaries. The authors, though brief, are admirably clear: An auxiliary is a “logical assumption that is required to link the theory to an actual observation” (p. 542). In effect, they are saying that to qualify as an empirically relevant and theoretically defensible auxiliary assumption, the auxiliary must be grounded in a logical relation to the theory under investigation. This helps avoid recruitment of ad hoc stipulations (i.e., auxiliaries lacking clear logical connection with the theory being tested—e.g., number of cubicles used, type of timing device employed, number of age-relevant words presented, academic status of the individual conducting the study, and so forth; for discussion, see Klein, 2014b) to insulate the theory from empirical refutation.

If a theory is an explanation of the cause of a phenomenon, the theory will include either statements that impose theoretically mandated conditions on the observance of the phenomenon, or impose an implicit default condition (e.g., this phenomenon can always be observed under any conditions operating at the time). A well-specified theory identifies these conditions as part of the explanatory reasoning of the cause of the observed phenomenon. These conditions need to be justified as a logical consequence of the theoretical structure being invoked—as would any theory-claim—and their alleged effect(s) tested as one might any other claim within that theory.⁶ To avoid ad hoc stipulation, the “conditions-required-to-be-present-in-order-that-the-phenomenon-might-be-observed” need justification themselves, not simply “I assume such and such” (e.g., much as many in psychology say “I assume the variable is real-valued and continuous” with no explicit justification for why that should be so; Michell, 1999; Uttal, 2008).⁷ That is, we cannot simply assume an auxiliary is applicable to a particular domain of inquiry: its relevance to the theory figuring in the replication effort must be logically justified.⁸

To their credit, Trafimow and Earp include a non-ad hoc provision (2016, p. 543) that requires a theory be sufficiently well-specified (or, to use their terminology, “relevant”) in order that logical connections between theory and auxiliary can be forged. Unfortunately, this requirement has the consequence of refuting their initial thesis (i.e., that a well-specified theory is *not* necessary for a successful replication). Put differently, Trafimow and Earp present two theses: (a) well-specified theory *is not necessary for replication* and (b) well-specified theory *is necessary for replication* in virtue of its role in establishing the epistemic warrant of auxiliary assumptions. The falsehood of these two claims is guaranteed as the logical consequence of holding both to be true.

Acknowledgements

Special thanks go to Dan Robinson, Kirk Michaelian, Galen Strawson, Carl Craver, Tim Lane, Sven Bernecker, Byron Kaldis, Robert Klee, Charles Talieferro, Myra Schectman, Alba Papa-Grimaldi, and Paul Barrett for excellent comments and suggestions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. Interestingly, Trafimow and Earp never tell us what, in their view, a well-specified theory consists in. They simply label a theory “well-specified” if it is “relevant” or “clear.” But such designations tell us little. Relevant in what way? Relevant to what? What does the clarity of a theory consist in?
2. This, of course, is simply a tautology (i.e., essential conditions are essential).
3. At issue here is what is called “paradoxes of material implication.” For example, if P is false then it implies *every* Q (this is referred to as “explosion”). Equivalently, if P implies that either Q or its negation is true, then their disjunction is implied by every P.
4. A hallmark of scientific theory (at least since Galileo)—sophisticated mathematical formalisms enabling precise parametric prediction—is far more likely in the physical than in the behavioral sciences (e.g., Collingwood, 1946/2005; Fodor, 1968; Shapin, 1996; Uttal, 2008; for recent discussions of the difficulties of extending normative scientific theory to explorations of the mind, see Klein, 2014a, 2014b, 2015, in press). For example, following a careful review of the evidence, Shulman (2013) concludes: “In the realm of the personal, nothing can be assumed with the assurance associated with mathematics, or even the empirical laws of physics ... no generalizations ... from psychology can be taken to be the starting point from which reliable neuroscientific truths can be deduced” (p. 51). Fodor (1974) puts things more bluntly: it is not “required that the taxonomies which the special sciences [e.g., psychology] employ must themselves reduce to the taxonomy of physics. It is not required, and it is probably not true” (p. 114).
5. Here are a few examples. First, Trafimow and Earp (2016) rely on physics’ authority to assert that Newton never defined mass. But $F=MA$ is at least in the ballpark of a definition: Mass is F/A . To claim that Newton didn’t pin down his terms strikes me as incorrect (but I’m not a physics expert). Second, Galileo worked out the mathematics of falling bodies and inclined planes in considerable detail and was expressly testing Aristotle’s theory connecting speed of descent to weight. To claim this as an example of poorly specified theory strikes me as, at best, controversial.
6. While these conditions follow from Trafimow and Earp’s definition of auxiliary assumptions (2016, p. 543), not all auxiliaries derive justification from principles embodied in the theory being tested. Some are established by well-confirmed hypotheses from other scientific fields (e.g., the operating characteristics of various measuring devices used to conduct research).
7. I thank Paul Barrett for this example.
8. As discussed in Klein (2014b), the social priming literature is replete with auxiliary assumptions that have an ad hoc quality greatly exceeding any rational connection to theory (i.e., this phenomenon is found only if N cubicles are used; the timing instrument for conducting a successful replication effort must be of form Y, etc.).

References

- Chang, H. (2010). The hidden history of phlogiston. *HYLE – International Journal for the Philosophy of Chemistry*, 16, 47–79.
- Collingwood, R. G. (2005). *The idea of history*. Oxford, UK: Oxford University Press. (Original work published 1946)

- Fodor, J. A. (1968). *Psychological explanation: An introduction to the philosophy of psychology*. New York, NY: Random House.
- Fodor, J. A. (1974). Special sciences (Or: The disunity of science as a working hypothesis). *Synthese*, 28, 97–115.
- Klee, R. (1997). *Introduction to the philosophy of science: Cutting nature at its seams*. New York, NY: Oxford University Press.
- Klein, S. B. (2014a). *The two selves: Their metaphysical commitments and functional independence*. New York, NY: Oxford University Press.
- Klein, S. B. (2014b). What can recent replication failures tell us about the theoretical commitments of psychology? *Theory & Psychology*, 24, 326–338.
- Klein, S. B. (2015). A defense of experiential realism: The need to take phenomenological reality on its own terms in the study of the mind. *Psychology of Consciousness: Theory, Research, and Practice*, 2, 41–56.
- Klein, S. B. (in press). The unplanned obsolescence of psychological science and an argument for its revival. *Psychology of Consciousness: Theory, Research, and Practice*.
- Ladyman, J. (2011). Structural realism versus standard scientific realism: The case of phlogiston and dephlogisticated air. *Synthese*, 180, 87–101.
- Laudan, L. (1990). Demystifying underdetermination. In C. Wade Savage (Ed.), *Scientific theories* (pp. 267–297). Minneapolis: University of Minnesota Press.
- Margenau, H. (1950). *The nature of physical reality*. New York, NY: McGraw Hill.
- Michell, J. (1999). *Measurement in psychology: Critical history of a methodological concept*. Cambridge, UK: Cambridge University Press.
- Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283–308). San Francisco, CA: Academic Press.
- Scheele, C. W. (1931). *The collected papers of Carl William Scheele* (L. Dobbin, Trans.). London, UK: G. Bell & Sons.
- Shapin, S. (1996). *The scientific revolution*. Chicago, IL: The University of Chicago Press.
- Shulman, R. G. (2013). *Brain imaging: What it can (and cannot) tell us about consciousness*. Oxford, UK: Oxford University Press.
- Torgerson, W. S. (1958). *Theory and method of scaling*. New York, NY: John Wiley & Sons.
- Trafimow, D., & Earp, B. D. (2016). Badly specified theories are not responsible for the replication crisis in social psychology: Comment on Klein. *Theory & Psychology*, 26, 540–548.
- Uttal, W. R. (2008). *Time, space, and number in physics and psychology*. Cornwall-on-Hudson, NY: Sloan.
- Wisniak, J. (2004). Phlogiston: The rise and fall of a theory. *Indian Journal of Chemical Technology*, 11, 732–743.

Author biography

Stanley B. Klein is a Professor in the Department of Psychological and Brain Sciences at UCSB and also has an appointment in the Department of Philosophy. His primary research interests are memory, the self, consciousness and the philosophy of science.