

Debates about the Scientific Status of Psychology: Looking at the Bright Side

Lucas B. Mazur^{1,2} · Meike Watzlawik¹

Published online: 18 June 2016

© Springer Science+Business Media New York 2016

Psychology Comes to Halt as Weary Researchers Say the Mind cannot possibly Study Itself. This headline appeared on the pages of the satirical newspaper, *The Onion*. What makes this headline humorous are the multiple layers of analysis it contains, and the clarity and succinctness with which it presents that which is supposedly unintelligible. In order to see the humor in this headline, the reader has to know about the apparent conflict between objective science and human subjectivity, but crucially, they also have to grasp both at the same time and to be able to flip back and forth between them. Debates within psychology regarding the field's scientific status often more closely resemble the content of *The Onion* article (with the categorical impossibility of a “science of the mind”), than the broader awareness required of *The Onion's* readers (containing both an appreciation of, and healthy distance to, those challenges). Academic discussions of psychology's scientific status are often artificial and divorced from the more complex picture that practitioners, researchers, and laypersons alike have of the field. Similar to how overuse in academia has led to the notion of a “moot point” shifting from meaning “that which is deserving of debate” to “that which is *not* worth debating,” many of the ostensive clashing dichotomies within the science of psychology have become more of an intellectual sparring ground than actual battlefield.

In the pages that follow it will be argued that these debates are helpful for better understanding human psychology, but that, like the readers of headline mentioned above, we should rise above them. These debates are important, but only in as far as they shed light on the true object of our concern, human psychological functioning. To this end, we should be cautious so as to not overvalue the methodological and

✉ Meike Watzlawik
meike.watzlawik@sfu-berlin.de

Lucas B. Mazur
lucas.mazur@sfu-berlin.de

¹ Sigmund Freud University, Berlin, Germany

² Jagiellonian University, Krakow, Poland

theoretical tools at our disposal. When we do so, we can unintentionally give credence to the pejorative meaning of the adjective “academic” – as in “divorced from practical reality” or “having no practical importance” (see “*academic*” in the *online dictionary of Merriam-Webster Inc.*, 2015). By rising above the fray, we can learn from all sides, and in the process, we can keep academic debates on psychological science both relevant for our lives and from following the historic path of the “moot point.”

How can the Mind Study Itself?

Regardless of how far back one traces the inception of the study of psychology people have pondered whether the mind is able to study itself and they have struggled to understand what the ability to even pose that question means for the debate (most famously, Descartes). How can we reconcile objective knowledge with subjective experience? This question, among others, has inspired some of the most well-known writings in the philosophy of science, such those of Thomas Kuhn and Karl Popper. Until today, scholars continue to discuss such issues, and do not agree on singular answers. The discomfort caused by such dichotomies – what Bernstein (1983) called *Cartesian Anxiety* – remains particularly acute within psychology.

To address this “anxiety,” it is worth exploring some of the dichotomies that often accompany debates surrounding the status of psychology as an empirical science. Despite being important and useful, these dichotomies usually are simplifications of a more complex reality. What is more, we believe that when confronting such dichotomies, while it is important to advocate the position one finds the most convincing in the given circumstances, it is even more important to step back so as to learn from the dynamics of the exchange itself. The “give and take” between the poles of such debates is itself incredibly valuable, arguably more so than either side by itself.

In order to limit the scope of these broad issues and to keep the article’s length to a minimum, the discussion will be largely framed around some of the points made by Smedslund (2016), who argues that psychology is not, and cannot be, an empirical science. Smedslund has identified four key issues that he believes separate psychology from the natural sciences, and on the basis of these issues he concludes that psychology cannot be an empirical science. The reasons for this are that within psychology:

- (1) the object of study is influenced by social interactions, including the experimental process itself;
- (2) the data gathered are more pseudo-empirical than empirical;
- (3) the processes studied are irreversible in nature (while those studies by the natural sciences are reversible) and can thus never be replicated, and;
- (4) the number of independent variables affecting the phenomena under study is uniquely large, too large in fact for psychologists to draw meaningful causal inferences.

While we generally agree with Smedslund’s (2016) critical approach and with many of his particular observations and arguments, we believe such a unilateral rejection of psychology as an empirical science to be founded on too rigid a commitment to the various dichotomies supposedly separating psychology from science. We believe the position taken by Smedslund on the four issues mentioned above might be loosely placed

under three classic dichotomies within psychology: (1) objectivity versus subjectivity, (2) the individual versus the aggregate, and (3) description versus prediction. We will now briefly revisit these three classic dichotomies in turn and argue that while useful for mobilizing critical thinking and intellectual debate, psychology does not neatly fit into one side over the other. Rather, psychology can and does learn from both sides.

The Objective and Subjective

While the notion of *objectivity* is surprisingly new within science, stretching only as far back as the nineteenth century (Daston and Galison 2007), the broader debates around the possibility of actually knowing the world, *as it really is not just as it seems to us to be*, are obviously much older. This distinction continues to color discussions within psychology, particularly when it comes to the kind and quality of data we collect. In short, do the data speak to the world *out there*, as it *really* is, or are they mere symbolic representations of the way we subjectively perceive it to be? Michell (2004) argues that the quantification of psychological phenomena we use (generally based on Likert scales) is built more on conjecture, politics, and precedent, and less on scientific evidence. Similarly, in the 1930's Ichheiser (1935) argued that what we usually take to be psychological facts are often what he called *as-if facts*, as they depend more upon the situation and the perceptions of the researchers than on the people they are actually researching. This is like the claim that snow *is* white, and only appears to be violet at sunset – something that is based on our own normative position (not to mention being dependent on the nature of sunlight, the atmosphere, etc.). Similar arguments are often encountered within the philosophy of science, where the “human element” repeatedly pops up as influencing theory, method and data. For example, the classification of illness and even the cause of death are subject to the influence of subjective, distally related and even arbitrary factors (Bowker and Star 1999). The role of subjectivity in ostensibly pure, objective science has been debated for decades, if not centuries (Gaukroger 2001; Reis and Sprenger 2014). It is tempting to think that while a person is not the same as another person, an atom of a particular element is always the same. The difference is not, however, as obvious or necessary as it would first appear, as people are in many meaningful ways interchangeable (e.g. various psychological functions are shared by various people, or absent in various people, with considerable regularity), and atoms of the same substance are not always actually the same (e.g. depending upon what “state” they are in, the speed at which they are traveling, the nature of their nuclei, etc.).

On this point, Smedslund (2016) nevertheless makes a powerful argument and a practical suggestion for the field of psychology in particular – although it would also seem to be of value to other fields as well. In psychology, it is often implicitly assumed that “hypotheses that make sense are true, and hypotheses that do not make sense are false”, something that undermines the very notion of empirical testing. Thus, “in cases of research where the independent and the dependent variable are logically connected, the charge of pseudo-empiricality holds up”. He continues with a very practical piece of advice: “A simple way to test for pseudo-empiricality is to consider if a *negation* of the hypothesis is possible and acceptable. If a negation is unacceptable (absurd, senseless), the hypothesis is not empirical since it expresses a necessity and could have been stated in advance”. This criticism is very serious, and leads to multiple related issues. For example, the burden of proof, especially how

“convincing” findings are based on the p -value obtained in statistical tests, is often influenced by the logic of the argument being made. For example, if the statistical results are in line with what we already know or believe, it is more readily accepted than if it runs counter to those positions, in which case a greater burden of proof often materializes (e.g. Goldfried 1959). This is indeed a serious challenge for the field, and it is also closely related to issues of funding, publishing, hiring, etc. (for a classic discussion of these issues, see Merton 1968).

The hard sciences are not, however, immune to charges of pseudo-science. Nevertheless, as with Smedslund’s other arguments, this is indeed a serious challenge to empirically based psychology, and the field would be well served by following his practical suggestion. However, once again we do not believe this should lead us to draw a clear line in the sand as Smedslund does: “the goal of finding some hitherto unknown genuinely empirical regularity is unlikely to be reached”. While the large number of pseudo-empirical hypotheses out there is indeed a serious problem for the field, that need not lead us to disregard the large body of good data that has been gathered. Well-formulated, empirical research has led to clear and counter-intuitive patterns – such as various, independent kinds of memory (short-term, long-term, episodic, semantic, motor, etc. e.g. Tulving 1985), the plasticity of memory (e.g. changing in the face of new conditions, e.g. Neisser and Fivush 1994; as well as physiology changes, e.g. Lamprecht and LeDoux 2004) and of the brain (e.g. shifting regional specializations, Huttenlocher 2002), or the power of collective identities (including the tendency towards ingroup bias, e.g. Spears 2007), to name but a few.

On the most basic level, the charge of pseudo-empiricality is disturbing because, in general, want to know *the facts*. This is what makes horoscopes and pop-psychology so appealing. However, the goal of objectivity is a tall one, not just in psychology but in all fields, and it is more of an ideal than an actualized quality of research. This goal pushes us to reduce the kinds of biases in our hypotheses and methods in line with Smedslund’s suggestion, but it need not speak directly to the unattainability of *acceptably* objective data (given the current limits of our theory and method, among other things). Such statements as “the fact that it is raining...” need not be a priori qualitatively different in its degree of objectivity from such statements as “the fact that you feel that way...” It can stop raining before the sentence was completed (Or was it really hailing?), just as the person can stop feeling that way (Or maybe it was really X that they were feeling?). The claim to objectivity often depends less on the kind of data *out there*, and more on the qualities of the questions we ask, the methods we use, our hypotheses, and crucially, our operationalizations. Rather than a sharp line between objective and subjective data, what we see is a *spectrum of acceptable objectivity*. To give a slightly different, but related example, no one would claim that *pain* does not exist, but we have yet to develop any purely objective understanding of what pain, or levels of pain, actually mean (e.g. Williams et al. 2000). For this reason, psychologists are explicitly turning to introspection and other subjective approaches to the study of such phenomena (Price and Aydede 2005). Similarly, people do not doubt that depression exists; even though we do not have any singular definition of this psychological phenomenon that would satisfy all. What is more, there are certainly better and worse ways to assess depression, something that is determined not by how true the assessment is to the ideal of objectivity or to subjective experience. The real issue is what we make of the data, what conclusions to we draw from them, and what we then do with them.

At times, Smedslund (2016) seems to support this milder position, when he states that “instead of making a sharp distinction between empirical and pseudo-empirical one may also envisage a gradual distinction based, not on formal logic, but on the amount of change in world-view required by a given finding”. A similar position has been taken by scholars in areas other than psychology. For example, when speaking of risk assessment within medicine, it has been argued that there are no objective probabilities that could once and for all embody all possibilities (Spiegelhalter 2008). However, Spiegelhalter asserts that while probabilities may not be *objectively measurable*, they can be *personally quantifiable*, in that acceptably objective data can become meaningful for individuals and that it can come to meaningfully influence their worldview (as Smedslund 2016, also suggests).

The Individual and the Aggregate

Of what value are aggregate data for predicting the lives of the individual, or what is more, of what value are they for making professional, psychological prescriptions regarding such an individual life? Here, the notion of *value* has, of course, a double meaning. It can refer to the relative utility of aggregate data for the purpose of individual prediction. It can also refer to the relative importance that aggregate data *should* have in the lives of individuals; that is, to the relative weight one should give such data when living one’s own life (or when making suggestions to others regarding how they should live theirs).

It is clear that psychology has been enriched by both ideographic and nomothetic approaches. We have learned a tremendous amount about memory from studying individuals, for example famously the patient “H.M.” (Squire 2009), but we have also learned a lot about memory via aggregate data, such as Loftus’ (1980) famous studies about the plasticity of memory. We have learned a great deal about political psychology from the study of both individual leaders, but also trends in the masses, as exemplified in the notion of the *Authoritarian Personality*, and more recently, in *Social Dominance Orientation* and *System Justification Theory*. Similarly, we have learned a great deal about sexuality and sexual identity from nomothetic studies of populations (e.g. Kinsey et al. 1948), but also of detailed studies of individuals – always taking into account that findings represent a certain point in time and that they are influenced by the specifics of the surrounding social and cultural settings.

For the point being made here, it is important to note is that in all of these cases the given set of ideographic or nomothetic studies is also of value to the other side. The notion of the plasticity of memory has a lot to say about the working of our own unique memory, just as the working of H.M.’s brain can tell us a lot about the working of memory in general. Individual leaders tell us much about the populations they lead, just as general trends in the population tell us much about individual leaders. In this spirit, the likes of Gustaw Ichheiser (1936) and Michael Schwarz (2009) advocated for psychological research to focus on outliers and exceptions, not simply because of their uniqueness, but because of what they could also tell us about statistically and normatively “normal” psychological functioning. Similarly, aggregate level research, such as that on various indices of psychological disorders (e.g. depression), are generally valued in as far as they are suggestive for individual diagnosis and helpful for individual treatment. What

is more, individual variation and uniqueness are in fact highlighted by an awareness of what averages actually mean. Stated more broadly, the distinction between nomothetic and ideographic studies is not based so much on an absolute distinction existing between individuals and groups, as in its utility for helping us make sense out of a complex world by setting practical, functional limits on the scope of our research. It is then important for psychologists to assess what this research has to tell us about what psychology actually studies – human psychological functioning – something that includes *both* individuals and collectives. In short, when tracing the history of most psychological constructs, one generally finds them weaving back and forth between individual-level and aggregate-level data.

Smedslund (2016) asserts that “one can make no valid inferences from averages to individuals, and, after all, individuals are what psychology is all about”. We would argue that such rejections of aggregate level data are made, at least in part, on the basis of a misreading of what aggregate level data mean for the individual. Aggregate level data speak not only to averages, but also to the individual variation out of which they are built. This point is argued in a very personal piece written by Steven Jay Gould (1985). After being diagnosed with abdominal mesothelioma, Gould was terrified to find out that his life expectancy was put at eight months. After looking deeper into the literature, he found that this average hid a fairly wide range of life expectancies. Gould uses this very moving personal example to illustrate not only how we often misunderstand what aggregate data mean, but also how we continue to have a choice of what it means *for us as individuals*, even after we do understand it. Aggregate data have much to offer, however, we still have to make sense of them from the perspective of our own individual lives, since as Gould’s title reminds us, *The Mean isn’t the Message* (for a similar argument regarding *aleatory* and *epistemological* uncertainty, see Spiegelhalter 2008). Thus, in the face of aggregate data, the individual is more like a car boxed in by painted lanes on the road, than a train rolling down fixed tracks. One may not always remain in the middle, and in fact, many do not. However, if you want to learn about driving (e.g. when traveling between England and mainland Europe), such averages are nevertheless an important piece of data to have. The same goes for any area of psychological research, such as language learning, head injuries, emotions, or semiotics.

Description versus Prediction

As a field, psychology is remarkably poor at predicting human behavior (Sutton 1998). Thankfully. Not only is this the result of the tremendous complexity of human psychology and the tremendous limitations of our theories and methods, but crucially, also because of the tremendously high bar that we continue to set in the field. Despite all of the criticisms of the artificiality of the traditional psychological experiment (e.g. experimental control), it pales in comparison to that of other fields. For example, the predictive power of the natural sciences is in large part due to their relatively strong commitment to functional operationalizations (e.g. In physics “...in general, we mean by a concept nothing more than a set of operations; the concept is synonymous with the corresponding sets of operations”, Bridgman 1927, p. 5), just as psychology’s poor predictive power is linked to the field’s relatively weak commitment to them (e.g. the ever-changing definitions of *well-being*, *power*, or *motivation*).

Something similar is seen in the distinction that Smedslund (2016) draws between how psychology and the natural sciences might describe the raising of an arm and extension of a finger (presumably referring to the index finger and not the thumb or middle finger). He asserts that while the number of possible psychological descriptions is infinite, the number of physical descriptions is limited. This is however, a simplification of the true complexity of the physical sciences, as such an assertion is based on a great many “controls” – such as the assumption of an average or ideal type (a finger is just a finger, regardless of whose/which actual arm and finger we are talking about), the physical location in the “here-and-now context” (which assumes a static and singular physical space and relation between a limited number of physical bodies and forces in which to place the extended arm and finger), and agreement within these fields regarding theory, method and data. This simplified image of the natural sciences is appealing and widespread (see Irwin and Wynne 1996). Equally appealing and widespread is the simplified image of the (relative) complexity of human psychology. While we may not truly be intuitively aware of how much of our body is composed of water or empty space, as it appears solid enough to us, we are acutely aware of the gaps in our understanding of our own and other’s behavior and psychological functioning.

This is also the case when we draw a sharp distinction between the apparent reversibility of physical processes, but not psychological processes; as it depends on the restrictions placed on the operationalizations in question. While psychological processes (including those induced by experimentation) influence our subsequent psychological functioning in some ways, provided we are clear in our operationalizations, we are certainly able to make claims of reversibility – as in saying that one can become happy or sad *again*, or that one can *recultivate* an old friendship. On the flip side, while the physical world appears to be more easily reversible, as in raising one’s arm *again* or returning a gas to its prior state, one can challenge the assumptions on which such assertions of reversibility are made. A physicist explained claims of reversibility in physics as follows:

“They usually relate to something changing from an orderly state into an unordered one. Think about opening a gas bottle filled with helium outdoors. All helium atoms will appreciate the new won freedom and will wander off in all directions. How do you want to reverse that process? Let’s say, you make it easier for yourself and you open that bottle in a closed, completely empty room. Then (1) the helium atoms are still trapped somewhere, and (2) all atoms in the room are helium atoms. With a (very, very good) compressor, you could force them back into the bottle. BUT: to run the compressor, you would need electricity, and to generate that you would need to accept disorderliness somewhere else [...]. No matter what you come up with: If you think it through consequently, things will never be as orderly as they were before opening the bottle” (M. Kiefer, personal communication, February 2nd, 2016)

Looking at this example, one can also pose the question: When raising one’s arm again, is its positioning in the world and the functioning of its parts *exactly* as they were before, as if nothing internal or external to the body parts in question had changed? Is its meaning *exactly* the same to observers or to the owner of the arm himself? Or are we making a statement resembling the return of an emotional state or of a friendship –

that is, the return of a state that (“acceptably”) fits our operationalization of a higher-order construct?

While the philosophical puzzles such statements pose are many, we merely wish to argue that the apparent dichotomy between psychology and the natural sciences in this regard is not as crisp and clean as we may often think, but rather, that it is based on a certain degree of caricaturing both “sides.” How crisp and clean something may appear, also depends on the operationalization of certain phenomena. In the past, for example, gender or gender identity often had higher predictive power when still operationalized as only two options (male and female), and their determinants to be dichotomous female or male external genitalia. Our understanding has become more nuanced, and while coming at the cost of decreased predictive power, it has increased our understanding of individual variability and hopefully our acceptance of difference. Thus, our growing awareness of the complexities of human psychology generally preclude us from committing to operationalizations that would allow us to even entertain reaching the kind of predictive power achieved in the other sciences – and again, thankfully so! Any definition of gender identity or life satisfaction one might put forward would contain within itself the seeds of its undoing (although we have certainly seen improvements over the years). We simply care about the exceptions and variation within psychology in a way that is not as readily found in other fields (but is certainly there, if one looks close enough).

Description and prediction are intimately intertwined. Psychology does not, and cannot, ignore the matter of prediction, and prediction is impossible without some degree of prior description. The importance is to recognize the inherent limits of all attempts at prediction, while also recognizing that we describe in order to predict.

The Utility and Limitations of Dichotomies

Smedslund (2016) beautifully argues how interpersonal prediction is achieved not by accurate reading of the averages or individuals, but by active engagement between people, and thereby via the establishment of common rules and mutual trust (and of course by selectively processing and favorably interpreting subsequent behavior so as to make it appear in line with those predictions). Similarly, the client-patient relationship in clinical psychology, to give one example, is not and cannot be based solely on ostensibly objective data or ostensibly subjective opinions, nor on aggregate data alone or individual-level data, nor on perfect prediction (that does not exist in the first place) nor on pure description (as if we did not care about prediction). These divisions of the world are just that, divisions of something larger – and that something larger is precisely what we actually care about.

We can learn a great deal about the world by breaking it down into different units that are intelligible and more readily accessible to us. It is precisely this ability to meaningfully and constructively ask questions that lies at the bottom of science. In this spirit, psychological research can illuminate elements of our psychological functioning about which we have no explicit awareness – similar to our blindness to our own blind spot. In other words, it need not be “limited to explication and analysis of what is already implicitly familiar” as Smedslund (2016) suggests. Importantly, the history of psychology has seen considerable contributions from all sides of the dichotomies discussed above. The way that we break down the world, as well as the divisions between academic fields, are useful in that they allow us to focus our research efforts.

At the same time, no single approach has a monopoly on understanding our actual lives, something Goethe already stated in 1798 (as cited in Jardine 2000, p. 40):

“[a phenomenon] can never be isolated, appearing as it does in a constant succession of forms. In order to describe it, the human intellect determines the empirically variable, excludes the accidental, separates the impure, unravels the tangled, and even discovers the unknown.”

Calls for interdisciplinary research reflect an awareness of this fact, if often little else. The problems arise, however, when we confuse such methodological or theoretical divisions for divisions in the object of our study itself. Each “side” is inherently limited in ways that the phenomena we are studying are not (a claim that of course assumes some sort of awareness of this something greater). It is important that we appreciate these limits, but also the fact that we can learn from them. We should be critical but not dismissive, and importantly, we should learn from as many approaches as possible. While we should be skeptical of claims to theoretical or methodological hegemony, we should fear neither theoretical or methodological orthodoxy nor novelty, as the lessons we stand to learn from them in no way inherently betray the larger object in which we are interested.

However, by learning from these various positions, we will inevitably encounter moments in which the abstract conflict becomes acute in very personal ways. This is the tension that Gould (1985) confronted with his illness, and that Spiegelhalter (2008) explores within the information exchange of doctor-patient relationships. In such moments, it is important not to panic and rush headlong to one side over the other. Similarly, in such moments, so as to regain the certainty we have lost, we often try to merge both sides of the dichotomies so as to somehow overcome the limitations of each. This however, often undercuts the power of both, while simultaneously failing to improve on either.

Such confusion arguably finds expression in a quantitative questionnaire recently issued by the British Ministry of Education (Doughty 2016; Walker 2016). This questionnaire, distributed to 13 to 18 year olds, was reported to present 25 category options for the self-identification of participant gender identity. The debate that ensued did not revolve so much around the questionnaire’s ability to capture individual-level variation in gender identity (for which it was arguably very well designed), but rather, around the confusion that would result on the part of the overwhelming majority of participants accustomed (either by personal identification or socialization) to but two, or perhaps three, options. This study is interesting in that it is an attempt to gather quantitative data on a population, while simultaneously amplifying, in an attempt to mirror ideographic data (or at least detailed, qualitative studies), the potential variation in one area of identity.

As only a very small portion of the general population identifies as some form of transgender (Gates 2011), this particular study raises numerous questions regarding the relationship between individual uniqueness and aggregate data. As we know from decades of debates regarding racial categories (e.g. Bowker and Star 1999, Chapter 6), the selection of possible answer options is never neutral, and the attempt to pre-identify all possible individual-level variation is idealistic and unrealistic at best, but more often simply confusing and misleading (not to mention often hurtful and even harmful). There are certainly many other ways these categories can be categorically broken down – as there are a great many ways in which people can understand even the two most common options, male and female. People’s unique stories and identities cannot be

captured in a questionnaire, and adding more options, while often informative and even laudable, will *never* overcome the inherent limitations of this kind of quantitative research. This is precisely why we talk about *mixed* methods approaches, not *merged* methods approaches. Something will always be missing from the data, whatever approach we use, and even our best attempts to fill the ever-present gap by combining them will fall short precisely because our tools are by definition simplifications, or due to the uniqueness of the cases we select. However, this does not mean that we learn nothing from those attempts. We actually learn a great deal. The science of psychology may be flawed and disappointing on many accounts, but it also has a lot to offer. This is all the more true provided that in moments of uncertainty we can take stock of what its various approaches have to offer, while remaining aware that the object of our concern will remain just out of reach or will have changed in the moment of investigation.

Conclusion

What we have tried to argue is that (1) the classic dichotomies are tremendously powerful and useful for generating debate and new research, but (2) they are simplifications of the wider reality, and therefore, (3) while useful, their strict division should not be overstated. In other words, the notion of *psychological science* is not an oxymoron, but neither should it be taken for granted. These various dichotomies, as well as the purported division between psychology and science, are illusions that make the world seem intelligible, albeit in different ways. The natural sciences are not free of psychology, just as psychology is not a world apart from science. In fact, both grow in the same soil – that of our individual and social lives.

During the era of the classic *Methodenstreit* between historicism and positivism, Max Weber argued that we ought not to cling too tightly to either of the two, suggesting rather that we look for “adequate causality” (see Ringer 2004, Chapter III). This pragmatic approach was an attempt to keep the social science moving forward during a time when researchers were in effect arguing whether the left foot or the right foot is more important for locomotion (Weber 1906/1949, p. 115):

“Indeed, just as the person who attempted to govern his mode of walking continuously by anatomical knowledge would be in danger of stumbling, so the professional scholar who attempted to determine the aims of his own research extrinsically on the basis of methodological reflections would be in danger of falling into the same difficulties.”

In many ways his approach can be understood as suggesting the power of the middle way (“*in medio stat virtus*”). However, Weber’s position was not that simple. He clearly argued for a transparent commitment to both the causal relations we discover in the world (the ethic of *responsibility*) and our own values (the ethic of *conviction*) – values that can neither justify, nor be justified by, those causal relations. These two elements of his ethics are not two sides of a single spectrum, but two independent elements of our lives. The challenge he laid before us is not to find balance between opposing poles, but rather to skillfully use, but not overuse, the various tools at our disposal. In meaningfully using such tools as *ideal types*, Weber suggested that we both

embrace, and be responsive to, the reality “out there” in the world, while also being aware of, and committed to, our own personal perspectives (*Einseitigkeit*). In a similar way, we need not understand more recent calls to reinvigorate other, forgotten elements of the social sciences, such as *imagination* (Cornejo 2016), as rejections of the scientific elements of psychology. Thus, we believe that many of Smedslund’s insights (2016) need not necessarily be understood as necessitating the rejection of psychological science.

Smedslund (2016) rightly identifies what might be thought of as the field’s enchantment with science, in that we are starting to believe that psychological research can be guided by science alone. In light of such mystification, it is indeed important to point out that psychological research is driven by both data and values, and that it will never be solely data-driven. This awareness underlies, at least in part, recent calls for greater transparency regarding psychological findings, such as calls for open access data files and the submission of raw datasets when publishing. Recent discussions around the replicability of psychological findings are intimately linked with an awareness of the fact that non-scientific elements affect our scientific findings (Achenbach 2015; John et al. 2012). While such efforts will reduce certain kinds of bias, they will not and cannot, remove them entirely. Psychological science simply does not work that way, and even if it did, it would not be, by definition, an undertaking we would care as much about.

There is nevertheless something very comforting in accepting or rejecting fixed categories, as seen above in the case of gender identity for example. The push and pull of categorization is an inevitable part of who we are, and it is a tremendously powerful tool that can help us to live meaningful lives in a complex world – and to do meaningful research. Psychologists have found ways to point out the complexity and flux of human development by different approaches. A good example is the literature on intersectionality; e.g., Meyer (2012). Similarly, Goldstein (1995, p. 27) stresses that “[a]s soon as we attempt to grasp [living organisms] scientifically, we must take them apart.”

However, in regards to intersectionality, Budde (2013, p. 248) comes to the conclusion that “one would need an infinite amount of categories [...] to describe the complexity of life situations and power relations in an accurate and differentiated manner. Out of numerous reasons—mainly methodological and research-pragmatic ones—this categorical expansion is limited.¹” Similarly, Goethe classically asserted the reality and value of the whole; something that is eternally richer than any and all of our scientific studies, which are based on simplifications. He stated that (as cited in Holdrege 2014, p. 13), when scientifically examining the various elements of our lives and the surrounding world, we should:

“... not stop and dwell on [isolated findings], cling to [them], and view [them] as existing in isolation. Instead we should look about in the whole of nature to find where there is something similar, something related. For only when related elements are drawn together will a whole gradually emerge that speaks for itself and requires no further explanation.”

Thus, as the former head of the Kinsey Institute, John Bancroft (2014), suggests, we should consciously work so as to shore up our tolerance of uncertainty. In debates about

¹ Original: „[...] man [bräuchte] unendlich viele Kategorien [...], um die Komplexität sowohl von Lebenslagen als auch von Machtverhältnissen angemessen und differenziert zu beschreiben. Aus zahlreichen – vor allem methodischen und forschungspragmatischen – Gründen ist diese Ausweitung aber nicht beliebig weit zu treiben.“

the science of psychology, this would suggest that we be wary of claims to certainty, not by dismissing our limited theoretical and methodological tools, but by keeping in mind the fact that our tools, no matter how good, will never fully capture the wonder of our lives.

Compliance with Ethical Standards

Conflict of Interest Both authors declare that they have no conflict of interest.

Ethical Approval This article also does not contain any studies with human participants or animals performed by any of the authors.

References

- Achenbach, J. (2015). No, science's reproducibility problem is not limited to psychology. *The Washington Post*. Retrieved from: <https://www.washingtonpost.com/news/speaking-of-science/wp/2015/08/28/no-sciences-reproducibility-problem-is-not-limited-to-psychology/>. Accessed June 2016.
- Bancroft, J. (2014). *Tolerance of uncertainty*. Bloomington, IN: Author House.
- Bernstein, R. (1983). *Beyond objectivism and relativism. Science, hermeneutics, and Praxis*. Philadelphia, PA: University of Pennsylvania Press.
- Bowker, G. C. & Star, S. L. (1999). *Sorting things out. Classification and its consequences*. Cambridge, MA: MIT Press.
- Bridgman, P. (1927). *The logic of modern physics*. New York, NY: Macmillan.
- Budde, J. (2013). Intersektionalität als Herausforderung für eine erziehungswissenschaftliche soziale Ungleichheitsforschung [intersectionality as challenge for educational and social inequality research]. In S. Siebholz, E. Schneider, A. Schippling, S. Busse, & S. Sandring (Eds.), *Prozesse sozialer Ungleichheit [processes of social inequality]* (pp. 245–257). Wiesbaden: VS Verlag.
- Cornejo, C. (2016). From fantasy to imagination: A cultural history and a moral for cultural psychology. In B. Wagoner, I. Bresco, & S. H. Awad (Eds.) *The Psychology of Imagination: Social and Cultural Perspectives*. (in press).
- Daston, L., Galison, P. (2007). *Objectivity*. Brooklyn, NY: Zone Books.
- Doughty, S. (2016). Children as young as 13 to be asked whether they are 'gender fluid', 'demi-girl' or 'intersex': Official survey asks pupils to pick from a list of twenty-five genders. *Daily Mail*. Retrieved from: <http://www.dailymail.co.uk/news/article-3420203/Are-gender-fluid-demi-girl-intersex.html>. Accessed March 2016.
- Gates, G. J. (2011). *How many people are lesbian, gay, bisexual, and transgender?* Report issues by the Williams Institute. Retrieved from: <http://williamsinstitute.law.ucla.edu/wp-content/uploads/Gates-How-Many-People-LGBT-Apr-2011.pdf>. Accessed March 2016.
- Gaukroger, S. (2001). History of objectivity. In N.J. Smelser & P.B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 10785–10789). Philadelphia, PA: Elsevier.
- Goldfried, M. R. (1959). One-tailed tests and "unexpected" results. *Psychological Review*, 66(1), 79–80.
- Goldstein, K. (1995). *The organism*. New York: Zone Books.
- Gould, S. J. (1985). The median isn't the message. *Discover*, 6(6), 40–42.
- Holdrege, C. (2014). Goethe and the evolution of science. *Context*, 31, 10–23.
- Huttenlocher, P. R. (2002). *Neural plasticity*. Cambridge, MA: Harvard University Press.
- Ichheiser, G. (1935a). Psychiczne konflikty i ich znaczenie dla psychologicznego zrozumienia jednostki [Psychological conflicts and their value for the psychological understanding of individuals]. *Przewodnik Pracy Społecznej*, 13(3), 85–90.
- Ichheiser, G. (1935/1936). *Wypadki przy pracy ze stanowiska psychologii [Workplace accidents from the perspective of psychology]*. Warszawa: Instytut Spraw Społecznych.

- Irwin, B., & Wynne, B. (Eds.) (1996). *Misunderstanding science? The public reconstruction of science and technology*. Cambridge, UK: Cambridge University Press.
- Jardine, N. (2000). *The scenes of inquiry: on the reality of questions in the sciences*. Oxford: Clarendon Press.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), 524–532.
- Kinsey, A., Pomeroy, W. B., & Martin, C. E. (1948). *Sexual behavior in the human male*. Philadelphia, PA, USA: W. B. Saunders Company.
- Lamprecht, R., & LeDoux, J. (2004). Structural plasticity and memory. *Nature Reviews Neuroscience*, 5(1), 45–54.
- Loftus, E. F. (1980). *Memory*. Reading, MA: Addison-Wesley.
- Merriam-Webster, Incorporated (Ed.).(2015). *Academic* [Online Dictionary]. Retrieved from: <http://www.merriam-webster.com/dictionary/academic>. Accessed March 2016.
- Merton, R. K. (1968). *Social theory and social structure*. New York, NY: The Free Press.
- Meyer, D. (2012). An intersectional analysis of lesbian, gay, bisexual, and transgender (LGBT) People's evaluations of anti-queer violence. *Gender & Society*, 26(6), 849–873.
- Michell, J. (2004). *Measurement in psychology. Critical history of a methodological concept*. Cambridge, UK: Cambridge University Press.
- Neisser, U., & Fivush, R. (1994). *The remembering self: Construction and accuracy in the self-narrative*. Cambridge, UK: Cambridge University Press.
- Price, D.D., & Aydede, M. (2005). The experimental use of introspection in the scientific study of pain and its integration with third-person methodologies: the experiential-phenomenological approach. In M. Aydede (Ed.), *Pain: New essays on its nature and the methodology of its study* (pp. 243–273). Cambridge, MA: MIT University Press.
- Reis, J. & Sprenger, J. (2014). Scientific objectivity. In E.N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from <http://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=scientific-objectivity>. Accessed March 2016.
- Ringer, F. (2004). *Max Weber: An intellectual biography*. Chicago, IL, USA: University of Chicago Press.
- Schwarz, M. (2009). Is psychology based on a methodological error? *Integrative Psychological and Behavioral Science*, 43(2), 185–213.
- Smedslund, J. (2016). Why psychology cannot be an empirical science. *Integrative Psychological and Behavioral Science*, 50(2), 185–195.
- Spears, R. (2007). Ingroup-outgroup bias. In R.F. Baumeister & K.D. Vohls (Eds.), *Encyclopedia of social psychology* (pp. 483–485). London: Sage Publications.
- Spiegelhalter, D. J. (2008). Understanding uncertainty. *Annals of Family Medicine*, 6(3), 196–197.
- Squire, L. R. (2009). The legacy of patient H.M. For neuroscience. *Neuron*, 61(1), 6–9.
- Sutton, S. (1998). Predicting and explaining intentions and behavior: how well are we doing? *Journal of Applied Social Psychology*, 28(15), 1317–1338.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26(1), 1–12.
- Walker, E. (2016). Brighton school children asked to choose from list of 23 terms to describe their gender. *The Arhus*. Retrieved from: http://www.theargus.co.uk/news/14236195.Brighton_school_children_asked_to_choose_from_list_of_23_terms_to_describe_their_gender/
- Weber, M. (1906/1949). *The methodology of the social sciences*. Glencoe, IL, USA. The Free Press.
- Williams, A. C., Talfryn, H., Davies, O., & Chadury, Y. (2000). Simple pain rating scales hide complex idiosyncratic meanings. *Pain*, 85, 457–463.

Lucas B. Mazur received his PhD in social psychology from Clark University (Massachusetts, USA). He is currently Visiting Professor at Sigmund Freud University in Berlin, Germany and Research Fellow at the Jagiellonian University, in Krakow, Poland. He may be contacted at lucas.mazur@sfu-berlin.de.

Meike Watzlawik is a Professor at the Sigmund Freud University, Berlin, Germany. From 2011 to 2014, she was a visiting professor at the University of Osnabrück, Germany, Department of Development & Culture, and from 2008 to 2009, she was a visiting scholar (Feodor-Lynen Awardee of the Humboldt Foundation) at Clark University, Worcester, MA, working in the field of Theoretical & Cultural Psychology. Examining different aspects of identity has been the focus of her research (e.g., sexual identity development of adolescence, the impact of sibling relationships on identity development, couple identity, occupational identity).