

**DIVERSITY FOR TRUTH:  
REPLY TO JUSSIM, STANOVICH, AND STROEBE**

Bernhard Hommel

Department of Psychology, Shandong Normal University, Jinan, China  
Faculty of Psychology, TU Dresden, Dresden, Germany

Corresponding Address:  
Bernhard Hommel  
Shandong Normal University  
88 Wenhua Donglu Road  
Ji'nan, Shandong Province 250014  
China  
[bh@bhommel.onmicrosoft.com](mailto:bh@bhommel.onmicrosoft.com)

Hommel (this issue) has discussed and expressed his worries about the increasing intrusion of activist thinking into science. Based on the article of Roberts, Bareket-Shavit, Dollins, Goldie, and Mortenson (2020) on “Racial inequality in psychological research”, he highlighted three hallmarks of this intrusion: (1) a blindness to the multidimensional nature of diversity; (2) the failure to distinguish psychological mechanisms from the impact of moderators; and (3) a blindness to agency (in contrast to circumstances or “the system”) as a factor to explain psychological phenomena. Hommel argued that the uncritical acceptance and the introduction of political activist arguments into science would be likely to damage scientific freedom and independence. Jussim (this issue), Stanovich (this issue), and Stroebe (this issue) have commented on these claims. All three commentaries agree that intermixing political activism and science is not a good thing and likely to damage the reputation of science and its advocates, which would further reduce the impact of scientific insights on public decision-making. All three commentaries also agree with the three problematic hallmarks that Hommel (this issue) discusses, they even provide further evidence for them and add further implications. Given this general agreement, it makes little sense to defend or repeat my claims, nor would the reader benefit from comments of a European working in Germany and China on the various US-American political issues and implications that the commentators are raising. Accordingly, I will focus on common ground that my target article and the three commentaries, and to some degree even the article of Roberts et al., reveal and try to extract three general lessons with respect to how our discipline, or science as a whole, might constructively deal with diversity in the future.

### **Science for power or truth?**

The systems-theorist Luhmann (1996) has characterized social systems by means of their currency or guiding value. According to his approach, the currency of science is truth, and it is this value that activities in the system try to maximize, while the currency of politics is power.

This distinction relates to the two definitions of diversity provided by the Oxford dictionary that Jussim (this issue) is citing. Maximizing truth in our scientific theorizing requires the consideration of diversity in the first, variety sense. If our conclusions and theoretical claims can be demonstrated to not hold for a particular subset of individuals, if we thus fail to capture the whole diversity humankind can offer in our theorizing, our approach is at least incomplete, if not entirely wrong. Considering this kind of diversity is essential for finding truth, and so it is this kind of diversity that is important for science as a truth-seeking activity. The second definition refers to the active inclusion or involvement of particular subsets of individuals. Demands that relate to this definition, as put forward by Roberts et al., commonly use arguments that refer to either representativeness or power, or both (as in the case of Roberts et al.).

Hommel (this issue) has argued that, given the number of possibly relevant variables, truly representative designs are simply unrealistic and Stroebe (this issue) has explained that and why true representativeness is actually of no theoretical relevance anyway. Hence, representativeness is unrelated to truth, at least if it comes to the investigated participants. How about representativeness on the side of the researchers? As the commentators and Roberts et al. have emphasized, finding the truth is likely to be facilitated by a diversity of opinions on the researcher side. While race may not be the most obvious factor generating or indicating opinion diversity (Stanovich, this issue), the commentators have discussed various findings showing that more diverse groups are likely to find solutions faster or more likely, or solutions that are more creative or interesting. Less clear is, however, how this kind of diversity can be achieved. Mixing as many personal features as possible might be an option, but I am not aware of convincing empirical evidence suggesting that worldviews, perspectives, intellectual preferences, and other relevant features are strongly correlated with any of the features discussed by political activists or considered by affirmative action, such as race, gender, socioeconomic background, or sexual preference. As long as we do not have a clear vision with

respect to how intellectual diversity can be effectively maximized, no selection (which implies trust in the randomness of the underlying distribution) seems more useful than strong selection based on cues with unknown validity. However, formulations such as “one might expect White journal editors—whose gatekeeping function positions them to govern what is worthy of publication—to be less likely than journal editors of color to publish research that highlights the role of race in human psychology” (Roberts et al., 2000, p. 1296) suggest that it is not quest for truth that is driving this argument but a struggle for power. Striving for power is not illegitimate in a democratic society, but it falls into the realm of political activism, rather than science.

The lesson one can take from these considerations is probably well captured by the distinction between *representativeness* and *representation*. With respect to investigated participants, it seems more important to make sure that one’s theoretical assumptions and predictions hold for all kinds of people, which suggests that the representation of diversity in samples makes it more likely to find the truth, than to strive for true representativeness (according to whatever criterion one might favor). With respect to the investigators, it seems more important to make sure that as many ideas and views as possible are represented in a research team than to strive for representativeness of ideas and views according to some criterion. For instance, it is easy to imagine that the theorizing of a research group becomes deeper, more balanced, and (thus) more realistic if the team members vary in their background and theoretical conviction or perspective (such as environmentalists and geneticists in a team of developmental researchers), but it is hard to see why a truly representative distribution of corresponding positions in the team would be required. This suggests that representation is important for truth, but representativeness is not, whereas the opposite holds for politics: Given that the political currency is power, and power hinges on the numbers of supporters (e.g., majority), representation is nothing, but representativeness is all. Accordingly, it makes perfect

sense that political activists are interested in representativeness, while scientists are, or at least should be, interested in representation.

### **Grounding doubt in statistics or theory?**

The scientific quest for truth ideally culminates in theories that capture our insights into the functioning of our study objects or subjects, as in the case of psychology. Theories are thus the eventual goal of science and statistics are one of the many tools we are using to test and to build them. These roles of theories and statistics as ends and means, respectively, have changed and even tended to reverse in the recent years. In the course of the so-called replication crisis in psychology (e.g., Open Science Collaboration, 2015), statistics often dominate the discussions, and they seem to represent the actual goals of many studies. One example is the overly restricted interpretation of failures to replicate previous findings, which are often not taken as a theoretically interesting hint to a not-yet-fully understood moderator, context condition, or sample characteristic, but rather as evidence that the investigated phenomenon does not exist (e.g., Wagenmakers et al., 2016). If we are interested in promoting more interest in, and more empirical activities with respect to diversity, it would be useful to reconsider this unfortunate shift of emphasis from theory to statistics in two ways.

First, various authors, including the three commentators and Roberts et al. (2020), have rightly pointed out that researchers should actively engage in testing the generality of their theoretical ideas, but the narrow interpretation of replications failures as undermining the theory under test does not provide an optimal reward schedule for theorists to participate in such testing. Even though repeated failures to replicate should certainly be taken serious and at some point motivate dropping the theoretical approach, a realistic interpretation of at least some replication failures would be that they point to the existence of important moderators that the theory has yet to consider (Stroebe, this issue). According to this interpretation, failures to replicate would actually be welcome because they help guiding theorists to develop even better

versions of their theory. De-escalating the ongoing methodological debate and emphasizing this very valuable function of systematically extending theory testing to new populations would help a great deal to convince researchers of the use and promise of embracing diversity in science.

Second, the commentators have emphasized that jumping from statistics to theoretical conclusions is commonly more complicated than activist arguments imply. One example is the disparity fallacy mentioned by Stanovich (this issue), according to which non-representative distributions necessarily imply societal injustice—thereby neglecting various possible other factors that may just as well explain the distribution (Sowell, 2019). Another is the conclusion that theory testing with non-representative samples is necessarily problematic, a line of reasoning that Stroebe (this issue) forcefully argues to be incorrect. Hence, statistical observations as such commonly do not tell us much without further theoretical assumptions. Accordingly, the fact that the Stroop effect, say, has only been tested in particular parts of the world population is insufficient to argue that theories about the Stroop effect are necessarily incorrect. What is needed to bolster such an argument is a theoretical scenario that could explain why a particular not-yet-investigated subpopulation may indeed not show the effect, and why that would theoretically matter. Hence, even if the Stroop effect was not yet tested in Catholic nuns, say, this fact alone is insufficient to challenge a particular Stroop theory. But if a realistic theoretical scenario could motivate the idea that strong religious or Catholic beliefs somehow restructure the cognitive system in such a way that processing Stroop stimuli might indeed be different in nuns, including this particular subpopulation would become important and worthwhile. In short, scientific doubt needs to be grounded in theoretical reasoning, but not (just) in statistics.

### **Diversity as noise or explanandum?**

Lewin (1931) has pointed out that our present scientific practice in psychology is still reflecting the Aristotelian idea that sorting (e.g., phenomena into processing or system categories) implies understanding. Aristotelian (as contrasted to Galilean) science is based on the conviction that only non-random events should be considered in scientific research, which among other things is reflected in our current statistical reasoning, according to which truth is revealed by eliminating or at least controlling the noise in the data (Hommel, 2020). Diversity is a particularly potent generator of noise in data, which makes it a natural enemy of researchers interested in noise-free, clean data that make it easier to find significant effects. Accordingly, we either try to reduce the amount of inter-individual variability by testing homogeneous or highly selected groups of participants, or by testing large samples, in the hope that inter-individual differences cancel each other out. In contrast to these strategies, Lewin (1931) strongly advocates a transition to a Galilean mode, in which both mean differences and inter-individual variability around the particular means need to be explained by using the same theory. A Galilean theory of the Stroop effect would thus not only need to explain why people are faster and more accurate when naming the color of the word GREEN written in green than when naming the color of the word RED written in green, but it would also need to explain why some people show a more pronounced difference between these two conditions than others. Galilean psychology would thus be more challenging to theorists, but it would also generate more interesting, and eventually more useful theories and make much more efficient use of collected data than our traditional Aristotelian approach (Hommel & Colzato, 2017)—a kind of “from nose to tail” approach to empirical research. Adopting a Galilean approach in psychology would very naturally attract attention to, and generate interest in human diversity, because diversity would no longer be considered an annoying noise factor but a central explanandum and key target of all theorizing.

## Conclusions

The target paper of Hommel (this issue) and the three commentaries by Jussim (this issue), Stanovich (this issue), and Stroebe (this issue) asked the question how psychology should deal with human diversity. My impression is that the major conclusions from this exchange are threefold. First, all contributions share the major claim of Roberts et al. (2020) that diversity is, or at least should be an important ingredient of psychological research and theorizing. Yes, people are different and this needs to be reflected in our work. Second, science and political activism both have their rights, functions, and importance, but mixing them up is very likely to do a bad service to psychology in general and weaken its reputation and role in public discussions in particular. Science is interested in truth while political activism is interested in power, and both aims are unlikely to support each other. Hallmarks of the intrusion of politics into science are the reduction of the diversity concept to one or few features that are assumed to indicate power disparities that call for repair, the failure to distinguish between the mechanisms and the content of psychological representations, and a selective focus on circumstances, society or “the system” as the only factor in explaining human behavior. Third, relying on theoretical considerations, rather than real or apparent statistical disparities, and considering diversity as explanandum, rather than noise, may provide a good starting point for raising more and broader interest in human diversity and the important information it can provide.

## References

Hommel, B. (2020). Pseudo-mechanistic explanations in psychology and cognitive neuroscience. *Topics in Cognitive Science, 12*, 1294-1305.

Hommel, B. (this issue). Dealing with diversity in psychology: Science or ideology? *Perspectives on Psychological Science*.



Hommel, B., & Colzato, L.S. (2017). The grand challenge: Integrating nomothetic and ideographic approaches to human cognition. *Frontiers in Psychology*, 8:100.

Jussim, L. (this issue). Diversity is diverse: Social justice reparations and science. *Perspectives on Psychological Science*.

Lewin, K. (1931). The conflict between Aristotelian and Galilean modes of thought in contemporary psychology. *Journal of General Psychology*, 5, 141–176.

Luhmann, N. (1996). *Social systems*. Redwood City: Stanford University Press.

Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. DOI: 10.1126/science.aac4716

Roberts, S. O., Bareket-Shavit, C., Dollins, F. A., Goldie, P. D., & Mortenson, E. (2020). Racial inequality in psychological research: Trends of the past and recommendations for the future. *Perspectives on Psychological Science*, 15, 1295–1309. <https://doi.org/10.1177/1745691620927709>

Sowell, T. (2019). *Discrimination and disparities*. New York: Basic Books.

Stanovich, K.E. (this issue). Toward a psychology of ideas rather than demographics. *Perspectives on Psychological Science*.

Stroebe, W. (this issue). The myth of the need for diversity among subjects in theory-testing research. *Perspectives on Psychological Science*.

Wagenmakers, E. et al. (2016). Registered replication report: Strack, Martin, & Stepper (1988). *Perspectives on Psychological Science*, 11, 917-928. 10.1177/1745691616674458.

### Acknowledgements

The author was supported by a Double-100 Talent Grant of the Province of Shandong, China.