

Psychological Science as a Complex System: Report Card

Abstract. Psychological science is increasingly diverse in the tools available for research and the questions it is able to ask. But this potential is seriously limited by a lack of diversity in study populations, situations and contexts explored, and who is conducting the research. It argues that the current situation is problematic and stubbornly so because of niche construction processes that favor the status quo. Systems level changes are needed to support a healthy psychological science.

Douglas L. Medin, Northwestern University

I have always loved data, even as a freshman at Moorhead State College in Moorhead, Minnesota. Thanks to the then Psychology Department Chair, Robert Solso, a friend (who will go nameless since he flunked out of school) and I were able to get access to the laboratory late at night to run a study on approach-avoidance conflict learning in rats based on our (the two of us, not Solso) very naïve reading of O. Hobart Mowrer's (1960) *Learning Theory and Behavior* book. Generally speaking, Solso let us do whatever we wished which led to bad experiments but good learning opportunities.

The idea of being able to do something new, something no one else had ever done was also of great appeal. My high school science teacher told me that I had a “gift for chemistry”, but I remember asking myself just how long it would take to get to the frontiers of chemistry research as I performed a lab “experiment” that probably hundreds of thousands of other students had done or were also doing. Contrast that with psychology where the first study I designed in my experimental psychology lab course was mediocre, but novel.

The lure of the frontier and data never wore off, through three years of graduate school at the University of South Dakota where I worked in Roger Davis' primate lab, ten years at Rockefeller University as part of the W.K. Estes mathematical psychology lab (also offering the opportunity to mingle with stellar scholars like George Miller, Neal Miller and Michael Cole), eleven years on the faculty at the University of Illinois (where other faculty were a wonderful, inexhaustible resource), a brief three year stint at the University of Michigan (where the highlight was the joint lab meetings with Ed Smith's group, aka “the Ed and Doug show”), and finally multiple decades at Northwestern University (coinciding with a shift towards including field studies that took me out of the lab and towards cultural and developmental research). My Northwestern experience has taught me to value diversity as a critical source of novelty, data, ways of looking at the world and ways of asking questions about it. This forms the basis of my comments on our field and its progress that I will present in the form of a report card.

The question of whether psychology is headed in the right direction contains interesting presuppositions in its very formulation. Among them are the potentially contentious ideas that psychology is an identifiable entity, that it is capable of movement, and that the movement has directionality with respect to some reference point or goal. A second order presupposition is that if psychology is not headed in the right direction we can do

something about it. To see just how bold the latter is, consider whether you think climate change is headed in the right direction (obviously not) and whether we can do something about it (not so obvious).

By now you're probably wondering when I'm going to stop dancing around this question and write something constructive; but presuppositions count. First, I don't think that psychology is an identifiable entity (but this may be a good thing). It seems to follow that *it* cannot move, rendering the questions of directionality and veridicality moot.

Imagine, as an alternative, that psychology is a complex system, interacting in various ways within and between other complex systems. For better or worse (better, I think), the research conducted under the label, psychological sciences, is diverse in method and content and in its interfacing with other disciplines such as neuroscience, economics, educational sciences, linguistics and anthropology. We have a broad range of theories (from framework theories that are judged for their usefulness to fine-grained models that may be able to predict reaction time distributions), methods, and theories about those same methods. That's the good news. The not so good news is that psychology is not very diverse in other critically important ways.

Complex system or not, questions about relevance, current practices and progress cannot be dismissed on grounds that our field is too dynamic to be evaluated at any one point in time. Congress won't buy this, NSF and NIH won't and we shouldn't either. Although psychology is not monolithic, it does have identifiable trends and its various areas have sets of consensual practices. And with respect to trends and practices that currently are in place, I would give our field a mixed review based on undeniable progress, paired with prominent points of consternation. Here are my grades in a few areas of particular consternation, along with some commentary.

1. Diversity of study populations (grade = D). I take our field as having the goal of identifying and understanding the range of human potential in forms of interaction with physical, biological and social environments. The Henrich et al. (2010) critique of our field's focus on samples from WEIRD (Western, Educated Industrialized, Rich, and Democratic) societies noted that WEIRD participants are particularly unrepresentative of the world at large. This paper is widely cited, its implications acknowledged, but there is little if any evidence that psychological science has changed in response to this acknowledgement (not to mention earlier critiques which have been met with the same apparent indifference). One could take encouragement from the fact that Henrich et al were able to find sufficient cross-cultural studies to make a strong argument, but I see it as contrastive with strong default norms. Why has change been so elusive? One guess is the same forces (e.g. needing more and more publications for tenure, etc.) that have led to the replicability crises.

And as we all know, WEIRD isn't the half of it. If you're an infant or young child and your parents are middle class and live within 10 kilometers of a major research university, your chances of being recruited for a developmental study are far better than if either condition is not satisfied. If you are not a college student at a major research university, your chances of being recruited for a study are very substantially diminished. Community college students need not apply. Even if you are attending a major research

Medin, D.L. (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*, 12(4) 669–674. [doi:10.1177/1745691616687746](https://doi.org/10.1177/1745691616687746)

university, your odds of being in a psychology experiment vary with your chosen field of study, with a psychology major a far better guarantee than an engineering or English major. On the rare occasions where fields of undergraduate study have been examined in psychological studies the results suggest that majors matter (e.g. Lehman & Nisbett, 1990; Frank, Gilovich & Regan, 1993, 1996; Wang, Malhotra & Murnighan, 2011). Overall, our discipline often acts as if it has an aversion to variability rather than seeing variability as its *raison d'être*. It may not be our explicit goal to maximize nondiversity but our cultural practices often have that consequence.

The increasing use of internet studies (e.g with Amazon's Mechanical Turk) is a potential palliative (this led to my giving us a D rather than a D- grade), but again internet studies have been conducted primarily for convenience and only rarely as a tool to explore sample variation. Narrow sampling in medical research is transparently immoral and inexcusable---is psychological science any different?

2. Contextual diversity (grade = C-/D+). I use contextual diversity as a broad category of variation that includes research procedures, environments and situations, and the use of multiple converging measures. Although there have been periodic critiques of psychological science's narrow choice of study populations, to my knowledge there have not been corresponding concerns expressed about diversity in procedures, measures and settings (but see Ceci et al, 2010 for one example). These forms of variation are critical to establishing robustness of findings as well as revealing patterned variation as a function of contexts and forms of assessment.

Multiple generations of social scientists have been brought up on Donald Campbell and Julian Stanley's "Experimental and quasi-experimental designs for research" first published in 1963, but still important enough to have a year 2015 edition (see also Campbell & Fiske, 1959). Among its many nuggets of wisdom is the discussion of the importance of convergent (and sometimes divergent) measures. It is ironic that the current so-called "replicability crisis" favors exact replications over the kind of robustness that can only be reinforced by converging measures.

Ideally, technology should facilitate our ability to instantiate methodological diversity. The advent of the personal computer dramatically increased the range of materials and stimulus presentation methods that could be employed and the ease with which this could be done, but it also led to greater methodological conformity to a participant interacting on a computer in a cubicle as the modal "social context." This is an extremely narrow sample of social contexts and situations (see Baumeister, Vohs & Funder, 2007, for a critique of psychology as the science largely limited to finger movements and self-reports).

A positive example of employing technology to assess external validity is the Hofmann, Wisneski, Brandt & Skitka (2014) study using smartphones to text participants randomly five times a day (between 9 a.m. and 9 p.m.) to probe whether they had witnessed, heard (or read) about or experienced a moral or immoral incident within the prior hour. Among their findings was evidence for moral self-licensing (engaging in a moral act made a subsequent immoral act more likely and a subsequent moral act less likely). This parallels

Medin, D.L. (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*, 12(4) 669–674. [doi:10.1177/1745691616687746](https://doi.org/10.1177/1745691616687746)

lab findings and suggests that licensing effects have considerable generality. Hofmann et al. also found that being the target of a moral act made it more likely that the person themselves would engage in a later moral act. Just a few years ago this type of study would have been either technologically impossible or prohibitively expensive to conduct (see also Hofmann & Patel, 2015). We need more positive examples.

Few, if any, social scientists would doubt that behavior in a given context is partly a function of the actor and partly a function of the situation. One of the delights of being at the University of Illinois was lunches with many fellow faculty. I remember vividly one such lunch with Phil Teitelbaum, who was conducting ground-breaking research on motivational systems and the brain. Phil told me the story of how he had done studies on rats pressing bars for food in a Skinner box and almost had concluded that the brain area he was studying using brain stimulation was the site of food motivation. Almost, but not quite. I forget just why he did it, but he put nesting materials in the box and then found that stimulation to the same area triggered nest building (and introducing another rat elicited mating behaviors). The clear lesson was that it is important to vary context.

A constructive example of attention to context comes from research on Attention Deficit Hyperactivity Disorder (ADHD), which could have been a field of research that focused exclusively on treating individuals (e.g. with drugs like Ritalin). But there is an accumulating body of work, often from school settings, looking for interactions with learning contexts. For example, there is evidence that children with ADHD spend more time on task with a small group arrangement than for either individual or larger group work (Hart, Massetti & Fabiano, 2010; Imeraj, Antrop, Sonagre-Barke, Deboutte, Deschepper, Bal & Rogers, 2013, see Harrison, Bunford, Evans & Owens, 2013; Van der Kolk, Agthoven, Buitelaar, & Hakkaart-vanReijnen, 2015 for reviews). Unfortunately the ADHD example of attention to situations is the exception to the rule. We need more such exceptions.

3. Researcher diversity, niche construction and biased defaults (Grade = solid D). Given that psychological science is a complex system interacting with other complex systems, it almost surely has evolved in a way that is adaptive for its practitioners. When our field's history includes an over-representation of (middle class) white males, these sorts of niche construction processes work to reinforce white male intuitions, white male values and white male practices. [These are big boxes---in addition to social class, political diversity may be important (Duarte et al, 2015).] Each of these factors may be a source of discord for would-be researchers who do not fall into this privileged (including privileged by precedent) group. Privilege includes who gets to decide what research questions are important (early research on primate reproductive behaviors focused on males and saw female primates as bystanders; this imbalance was countered when women entered the field) as well as who gets to become a research participant. It may be that the quickest route to study population diversity is researcher diversity.

In my opinion progress in our field is linked to researcher diversity, especially if diverse perspectives are encouraged and appreciated. In the absence of diversity we may be hindered in realizing that default assumptions are often cultural assumptions associated with one particular culture. These default assumptions may lead to limited procedural

exploration, a focus on WEIRD study populations, and deficit thinking on those occasions when non-WEIRD populations are studied. I'll offer one example of the latter.

The “word gap.” There is a considerable and somewhat controversial literature on the relation between language input conditions and children’s vocabulary development. Much of this was stimulated and perhaps sensationalized by a paper by Hart and Risley (1995) estimating that the children of middle class parents were exposed to thirty million more words than children of poor or working parents by the time they enter school (the so called “thirty million word gap”). These differences in exposure are correlated with SES differences in children’s vocabulary. There is also evidence that “vocabulary interventions” affect young children’s vocabulary learning (Marulis & Neuman, 2013; Roberts & Kaiser, 2011; (see also Goldin-Meadow et al., 2014), so the data linking language input and vocabulary size are not merely correlational.

What makes this literature controversial? If you’re interested in language learning you could hardly be faulted for studying variables related to language learning. The problems come when intervention effects become policy prescriptions. The word gap observations could be seen as suggesting that the reason that lower SES children are not as successful in school as their middle class counterparts lies with their parents who are not providing a rich language environment. This is a strong conclusion given that the original Hart and Risley study found no correlation between rate of vocabulary growth and children’s third grade scores in reading, writing, spelling or arithmetic (Hart and Risley, 1995, page 161). But the issues are deeper than that.

A cultural perspective on the word gap. One facet of culture is shared norms, and it is easy to slip from norms into notions about what is normal or even “natural.” Middle-class parents in the US talk to their infants from the moment they are born (sometimes even before birth). It doesn’t occur to them to not talk to their babies, even though babies don’t speak in response. In other cultures talking to infants is less frequent (Bornstein, et al., 2015; Sheffelin & Ochs, 1986; Ward, 1971). From the perspective of some cultures, Americans talk incessantly and often needlessly and in some of these cultural communities children “are to be seen and not heard.” Navajo adults, for example, may view a talkative child as self-centered and discourteous (Freedman, 1979).

It is important to recognize that a child-focused, steady stream of speech is just one of many approaches to child-rearing. There may be many other values that are important in different cultural communities rather than a single, “one best way.” For example, an important concept not addressed by most word gap research is “opportunity costs,” the observation that any one choice leads to the potential loss associated with not exercising alternative choices. For example, there is some evidence that adult-directed showing preschoolers the function of a toy limits children’s exploratory play and thus the limits the opportunity for discovery (Bonawitz, et al., 2011; see also Gweon, Pelton, Konopka, & Schultz, 2014).

Applying the concept of opportunity cost to the word gap domain, note that children are not just sitting around passively waiting for adults to speak with them. They may be engaged in creative play on their own, interacting with peers, and acquiring culture norms

Medin, D.L. (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*, 12(4) 669–674. [doi:10.1177/1745691616687746](https://doi.org/10.1177/1745691616687746)

through observing and engaging in family and community activities (e.g. Rogoff, 2003, Lareau, 2011), each of which may lead to the development of other important skills. To see the narrowness of focusing on vocabulary even within the domain of language, ask a friend whether they think that it's better to be able to speak using a large vocabulary or to be able to formulate an effective argument or tell a good story.

Just as there are distinct types of music, there are culturally distinct narrative practices. Some cultural communities favor “getting to the point” quickly and narratives organized as linear sequences. But other cultural communities focus more on creating a rich context for listeners and employing more non-linear or cyclical narratives. For example, when we asked rural Wisconsin Native American and European American adults to tell us about their last encounter with fish, the median number of words before fish were mentioned was the 27th word for European Americans and the 83rd word for Native Americans (Medin & Bang, 2014). As has been noted so often, differences are not deficits, and we need diverse perspectives to remind us of this. “Diverse perspectives” might well also include power sharing with and the perspectives of the “researched” in the form of partnerships with communities.

Summing up. This report card for psychological science is less than impressive. In his APA Presidential address George Miller (1969) said, “In my opinion, scientific psychology is potentially one of the most revolutionary intellectual enterprises ever conceived by the mind of man.” I think George Miller was and is correct but perhaps he should have italicized “*potentially*.”

It will take systems level changes to encourage and support study sample diversity, methodological diversity, and perhaps most importantly, researcher diversity. In complex systems, nonlinear processes are common, and sometimes small interventions have surprisingly large effects (e.g. Stephens, et al, 2012; Stephens, Hamadani & Destin, 2014). Therefore, a failing report card need not lead to complete pessimism about the future. Of course achieving these various forms of diversity likely would require training psychological scientists to understand culture and contexts that cultural groups find meaningful (and not just the lab). But if you like data, what could be more fun?

Footnote: Daniel Hruschka, Sandra Waxman, bethany ojalehto, Natalie Gallagher, Sylvia Perry, Kalonji Nzinga and Alissa Baker-Oglesbee provided constructive suggestions that improved this paper.

References

Baumeister, R.F., Vohs, K.D., & Funder, D.C. (2007). Psychology as the science of self-reports and finger-movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, 2, 396-403.

- Medin, D.L. (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*, 12(4) 669–674. doi:10.1177/1745691616687746
- Bonawitz, E., Shafto, P., Gweon, H., Goodman, N. D., Spelke, E., & Schulz, L. (2011). The Double-Edged Sword of Pedagogy: Instruction Limits Spontaneous Exploration and Discovery. *Cognition*, 120(3), 322-330.
- Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. D. (2015). Mother-Infant Contingent Vocalizations in 11 Countries. *Psychological Science*, 26(8), 1272-1284. doi: 10.1177/0956797615586796
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56(2), 81.
- Campbell, D. T., & Stanley, J. C. (2015). *Experimental and Quasi-Experimental Designs for Research*: Ravenio Books.
- Ceci, S.J., Kahan, D.M. & Braman, D. (2010). The WEIRD are even weirder than you think: Diversifying contexts is as important as diversifying samples. *Behavioral and Brain Sciences*, 33, 87-88.
- Duarte, J.L., Crawford, J.T., Stern, C., Haidt, J. Jussim, L, & Tetlock, P.E. (2015). Political diversity will improve social psychological science. *Behavioral and Brain Sciences*, 38, 1-58.
- Frank, R. H., Gilovich, T., & Regan, D. T. (1993). Does Studying Economics Inhibit Cooperation? *The Journal of Economic Perspectives*, 7(2), 159-171.
- Frank, R. H., Gilovich, T. D., & Regan, D. T. (1996). Do Economists Make Bad Citizens? *The Journal of economic perspectives*, 10(1), 187-192.
- Freedman, D. G., & DeBoer, M. M. (1979). Biological and Cultural Differences in Early Child Development. *Annual Review of Anthropology*, 579-600.
- Goldin-Meadow, S., Levine, S. C., Hedges, L. V., Huttenlocher, J., Raudenbush, S. W., & Small, S. L. (2014). New Evidence About Language and Cognitive Development Based on a Longitudinal Study: Hypotheses for Intervention. *American Psychologist*, 69(6), 588.
- Gweon, H., Pelton, H., Konopka, J. A., & Schulz, L. E. (2014). Sins of Omission: Children Selectively Explore When Teachers Are under-Informative. *Cognition*, 132(3), 335-341.
- Harrison, J. R., Bunford, N., Evans, S. W., & Owens, J. S. (2013). Educational Accommodations for Students with Behavioral Challenges: A Systematic Review of the Literature. *Review of Educational Research*, 83(4), 551-597.
- Hart, B., & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: Paul H Brookes Publishing.
- Hart, K. C., Massetti, G. M., Fabiano, G. A., Pariseau, M. E., & Pelham, W. E. (2010). Impact of Group Size on Classroom on-Task Behavior and Work Productivity in Children with ADHD. *Journal of Emotional and Behavioral Disorders*, 19(1), 55-64.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most People Are Not Weird. *Nature*, 466(7302), 29.
- Hofmann, W., & Patel, P. V. (2015). Surveysignal a Convenient Solution for Experience Sampling Research Using Participants' Own Smartphones. *Social Science Computer Review*, 33(2), 235-253.
- Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in Everyday Life. *Science*, 345(6202), 1340-1343.

- Medin, D.L. (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*, 12(4) 669–674. doi:10.1177/1745691616687746
- Imeraj, L., Antrop, I., Sonuga-Barke, E., Deboutte, D., Deschepper, E., Bal, S., & Roeyers, H. (2013). The Impact of Instructional Context on Classroom on-Task Behavior: A Matched Comparison of Children with ADHD and Non-ADHD Classmates. *Journal of School Psychology*, 51(4), 487-498.
- Lareau, A. (2011). *Unequal Childhoods: Class, Race, and Family Life*. Berkeley, CA: University of California Press.
- Lehman, D. R., & Nisbett, R. E. (1990). A Longitudinal Study of the Effects of Undergraduate Training on Reasoning. *Developmental Psychology*, 26(6), 952.
- Marulis, L. M., & Neuman, S. B. (2013). How Vocabulary Interventions Affect Young Children at Risk: A Meta-Analytic Review. *Journal of Research on Educational Effectiveness*, 6(3), 223-262.
- Medin, D. L., & Bang, M. (2014). *Who's Asking? Native Science, Western Science and Science Education*. Cambridge, MA: The MIT Press.
- Miller, G. A. (1969). Psychology as a Means of Promoting Human Welfare. *American Psychologist*, 24(12), 1063.
- Mowrer, O. H. (1960). *Learning Theory and Behavior*. Hoboken, NJ: John Wiley & Sons.
- Roberts, M. Y., & Kaiser, A. P. (2011). The Effectiveness of Parent-Implemented Language Interventions: A Meta-Analysis. *American Journal of Speech-Language Pathology*, 20(3), 180-199.
- Rogoff, B. (2003). *The Cultural Nature of Human Development*. New York, NY: Oxford University Press.
- Stephens, N.M., Hamedani, M.Y.G., & Destin, M. (2014). Closing the social-class achievement gap: A difference education intervention improves first generation students' academic performance and all students' college transition. *Psychological Science*, 25, 943-953.
- Schieffelin, B.B. & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15, 163-191.
- Stephens, N.M., Fryberg, S.A., Markus, H.M., Johnson, C.S., & Covarrubias, R. (2012). Unseen disadvantages: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology*, 102, 1178-1197.
- van der Kolk, A., Bouwmans, C. A. M., Schawo, S. J., Buitelaar, J. K., van Agthoven, M., & Hakkaart-van Roijen, L. (2015). Association between Societal Costs and Treatment Response in Children and Adolescents with ADHD and Their Parents. A Cross-Sectional Study in the Netherlands. *SpringerPlus*, 4(1), 1.
- Wang, L., Malhotra, D., & Murnighan, J. K. (2011). Economics Education and Greed. *Academy of Management Learning & Education*, 10(4), 643-660.
- Ward, M. C. (1971). *Them Children: A Study in Language Learning*. New York: Holt, Rinehart, Winston.

Medin, D.L. (2017). Psychological science as a complex system: Report card. *Perspectives on Psychological Science*,12(4) 669–674. [doi:10.1177/1745691616687746](https://doi.org/10.1177/1745691616687746)