Student subject pools can provide an invaluable benchmark for investigating generalizability across different social groups or cultures.

In their excellent article, Henrich et al. rightly caution us to be careful when we draw general conclusions from WEIRD subject pools, of which undergraduates are the most frequently used one, also in economics. My main comment is that the right choice of subject pool is intimately linked to the research question. Since the different behavioral sciences also have different research questions, the right choice of subject pool will also often be different across disciplines. In my own discipline, economics, students are actually often the *best* subject pool for quite a few (fundamental) research questions. Here is why I believe so.

Economic theories normally do not come with assumptions (or even caveats) about the restricted validity to only a specific group of people; that is, they (implicitly) assume "generality." Like the assumption of selfishness, "generality" is a good assumption in the absence of rigorous data. The tools of experimental economics have been deployed to investigate the empirical relevance of the selfishness assumption (see, e.g., Fehr et al. 2002) and are now also used to probe the "generality assumption," that is, the importance of variations of behavior across population subgroups within a given society (e.g., Bellemare et al. 2008) or across societies (e.g., Herrmann et al. 2008).

However, my main point is this: The "right choice" of subject pool depends on the research question. If the researcher is interested in understanding behavioral variation between particular groups of people, then the right choice is running experiments with these people. The landmark study by Henrich et al. (2005) is a shining example. Yet, at least in economics, substantial effort is also devoted to test formal theories or to detect interesting behavioral regularities (Bardsley et al. 2010; Croson & Gächter 2010; Smith 2010). Because economic theories normally assume generality, *any* subject pool is in principle informative about whether theoretical predictions or assumptions contain behavioral validity. At that stage, generalizability to other subject pools is not (yet) an issue. Among the universe of potential subject pools to test a theory, students are often the *perfect* one: on average, students are educated, intelligent, and used to learning. These are very valuable characteristics because, in addition to the main aspect of a theory of interest to the researcher, economic theories often assume cognitive sophistication. It therefore makes sense to control for sophistication also by choice of subject pool (in addition to clear instructions), in order to minimize chances of confounding genuine behavioral reactions to the treatment of interest with lack of understanding of the basic decision situation.

Take recent theories of social preferences (as surveyed, e.g., in Fehr & Schmidt 2006) as an example. In addition to other-regarding preferences, these theories all assume cognitive sophistication. When testing these theories, the main point of interest is not to find out whether people are as cognitively sophisticated as the theories (maybe wrongly) assume, but to see to what extent other-regarding motives *exist*, holding everything else constant. Because students are typically above average with regard to cognitive sophistication, they are often a perfect subject pool for *first* tests of a theory. Moreover, students, unlike most other subject pools, are readily available (and cost effective). Experiments can therefore also easily be replicated, which is important to establish empirical regularity and hard to achieve with any other subject pool.

Of course, strictly speaking, observed results hold only for the subject pool from which evidence is collected. Generalizability is a generic issue in any empirical research (Falk & Heckman 2009). However, once a clear benchmark result is established, we can proceed by testing, for example, how age and life experience matter (e.g., Sutter & Kocher 2007b), or how results extend to more representative subject pools (e.g., Bellemare et al. 2008; Carpenter et al. 2008). Along the way, researchers often establish whether and how students differ from the general population.

As Henrich et al. point out, understanding the potential influence of cross-societal (or cultural) differences in (economic) behavior is a particularly interesting direction for investigating generalizability. But it poses further challenges, in particular if socio-demographic factors matter (as some of the above-cited research suggests). The reason is that socio-demographic influences might be confounded with genuine societal or cultural differences. The problem is exacerbated the more subject pools are actually being compared. Again, to ensure that confounds are minimized, student subject pools are often the best available choice (Bohnet et al. 2008; Herrmann et al. 2008) to establish a clean benchmark result on how people from different societal/cultural backgrounds behave in the exact same decision situation – a fundamental question from the generality perspective of economics. The benchmark can – and should(!) – then be taken as a starting point for investigating generalizability to other social groups.

## It's not WEIRD, it's WRONG: When Researchers Overlook uNderlying Genotypes, they will not detect universal processes

doi:10.1017/S0140525X10000105

Lowell Gaertner,[a] Constantine Sedikides,[b] Huajian Cai,[c] and Jonathon D. Brown[d]

[a]*Department of Psychology, University of Tennessee, Knoxville, TN 37996-0900;* [b]*Center for Research on Self and Identity, School of Psychology, University of Southampton, Southampton SO17 1BJ, United Kingdom;* [c]*Institute of Psychology, Chinese Academy of Sciences, Chaoyang District, Beijing 100101, China;* [d]*Department of Psychology, University of Washington, Seattle, WA 98195-1525.*

**gaertner@utk.edu**
**cs2@soton.ac.uk**
**huajian.cai@gmail.com**
**jdb@uw.edu**

**Abstract:** We dispute Henrich et al.'s analysis of cultural differences at the level of a narrow behavioral-expression for assessing a universalist argument. When Researchers Overlook uNderlying Genotypes (WRONG), they fail to detect universal processes that generate observed differences in expression. We reify this position with our own cross-cultural research on self-enhancement and self-esteem.

We dispute the level of analysis Henrich et al. have employed to conclude that members of Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies are non-representative of the human species and to determine, more generally, whether an observed cultural difference contradicts a universalist argument. Borrowing from the biological lexicon, our position is as follows: Analysis of difference at the level of a narrow phenotypic behavioral-expression precludes detection of human universals that operate at the level of an abstract genotypic process. Stated otherwise, When Researchers Overlook uNderlying Genotypes (our acronym WRONG), they will fail to detect universal processes that generate observed differences in expression (Kobayashi & Brown 2003). We first frame our position with an example and then reify our position with our own cross-cultural research on self-enhancement and self-esteem – phenomena from which Henrich et al. derived their WEIRD conclusion.

The human diet exemplifies our position (Sedikides & Gregg 2008). When considered at the narrow level of observed behavior, human societies appear extraordinarily different in regard to what they eat (e.g., a Kosher diet precludes pork; a Jain diet is vegetarian). When considered more broadly, however, the diverse diets are connected and assimilated by a universal need for sustenance. It would be faulty indeed to conclude that the need for sustenance is less pronounced, if not absent, in one

society because it consumes less, if not any, of the foods consumed by another society – what is consumed depends on factors such as climate and custom. Our point, of course, is that human universals operate at the abstract level of process-and-function, and the expression of the universal emerges in conjunction with contextual considerations (Schlenker 1974).

To be clear, we do not oppose the study of concrete behavior. Cataloguing behavioral differences across societies certainly contributes to understanding the human condition. However, the presence of a behavioral difference per se is not evidence contrary to a universalist argument. The necessary consideration is whether the observed difference is produced by a process or function common across societies.

Henrich et al. suggest that WEIRD societies are peculiar, in part, because they uniquely possess positive self-views. Such a conclusion, however, is a consequence of the WRONG strategy. Our own cross-cultural programs of research on the *self-enhancement* motive (i.e., need to maintain a positive sense of self) and *self-esteem* (i.e., an affective self-evaluation) indicate that a positive self-view is a human tendency.

Our primary studies and meta-analytic syntheses indicate that both Westerners and Easterners self-enhance, but they do so on different attribute dimensions.[1] Westerners self-enhance (i.e., deem self as superior to peers) on attributes relevant to individualism, and Easterners self-enhance on attributes relevant to collectivism. This is because Westerners deem individualism, and Easterners deem collectivism, as important. Here a common process (self-enhancing on important attributes) is differentially expressed (individualism vs. collectivism), because culture affects the expression, not the presence, of the enhancement motive (Brown & Kobayashi 2002; Sedikides et al. 2003; Sedikides et al. 2005; 2007a; 2007b). Furthermore, that common process has the same functional association with psychological adjustment in both cultures: Self-enhancing on important attributes promotes better adjustment (e.g., greater well-being, less depression,) among Easterners and Westerners (Gaertner et al. 2008; Kobayashi & Brown 2003; O'Mara et al. 2009). Therefore, when assessed at the abstract level of process and function, members of WEIRD societies are quite normal in their striving for a positive self-view.

As Henrich et al. suggest, Eastern samples typically provide lower explicit reports of self-esteem than do Western samples (Heine et al. 1999). Such explicit reports, however, are compromised by a pervasive modesty norm in Eastern cultures (Brown, in press; Kurman 2003). Indeed, the cultural differences occur in reports of cognitive self-evaluation, not affective self-regard, and these differences vanish when modesty is statistically controlled (Cai et al. 2007). Similarly, cultural differences in self-esteem vanish when self-esteem is assessed with implicit measures that circumvent modesty norms (e.g., Yamaguchi et al. 2007). Furthermore, self-esteem reveals the same functional patterns across cultures. Self-esteem predicts greater well-being and lower depression in the East (Cai et al. 2009), just as it does in the West (Taylor & Brown 1988). Likewise, self-esteem bolsters against threats to self-worth in both cultures such that failure feedback more strongly erodes immediate feelings of worth for low rather than high self-esteem persons (Brown et al. 2009). Hence, when assessed at the abstract level of process and function, members of WEIRD societies are quite normal in their possession of a positive-self view.

In summary, testing human universals at the level of narrow behavioral differences between societies is the WRONG strategy. Human universals operate at the abstract level of process and function, and such universals can generate observed differences. We conclude with application of our argument to the man-to-boy insemination rituals of New Guinea with which Henrich et al. began their article. At the narrow level of the observed behavior, the rituals seem bizarre in regard to WEIRD standards. When considered more abstractly in terms of process or function (i.e., a social practice marking a boy's passage to manhood), the rituals connect with coming-of-age rituals practiced in other societies, such as the Bar and Bat Mitzvah, Credo-baptism, Debutante Ball, and Sweet Sixteen. The observed behaviors certainly differ, but the underlying psychological process is the same.

NOTE

**1.** Here and in the following, we use the terms *Eastern* and *Western* for expedience in reference to samples from East Asia versus samples from the United States, Canada, and Western Europe.

## Wired but not WEIRD: The promise of the Internet in reaching more diverse samples

doi:10.1017/S0140525X10000300

Samuel D. Gosling,[a] Carson J. Sandy,[a] Oliver P. John,[b] and Jeff Potter[c]

[a]*Department of Psychology, University of Texas, Austin, TX 78712;* [b]*Department of Psychology, University of California, Berkeley, CA 94720;* [c]*Atof Inc., Cambridge, MA 02139.*

**samg@mail.utexas.edu www.samgosling.com**
**cjosandy@mail.utexas.edu www.carsonsandy.com**
**o_johnx5@berkeley.edu http://www.ocf.berkeley.edu/~johnlab/**
**research@outofservice.com http://research.outofservice.com/**

**Abstract:** Can the Internet reach beyond the U. S. college samples predominant in social science research? A sample of 564,502 participants completed a personality questionnaire online. We found that 19% were not from advanced economies; 20% were from non-Western societies; 35% of the Western-society sample were not from the United States; and 66% of the U. S. sample were not in the 18–22 (college) age group.

Henrich et al. show that the vast majority of research in the behavioral sciences continues to be based on populations the authors call WEIRD because they are unlikely to be representative of humankind. Even more alarmingly, much of the research published in top-tier journals is not even representative of the populations in WEIRD (Western, Educated, Industrialized, Rich, and Democratic) countries. For example, in the 510 samples published in the *Journal of Personality and Social Psychology (JPSP)* in 2002, 85% of them were student samples, 71% of the participants were female, more than 80% were white, and the mean age was 22.9 years (Gosling et al. 2004).

What should we do about this? Henrich et al. conclude their article by urging institutions to improve the infrastructure for collecting data from non-WEIRD samples as well as the incentives for studying them. However, Henrich et al. offer very little in the way of concrete practical suggestions for expanding the reach of research in the behavioral sciences.

We propose that the Internet holds great promise for broadening the participant base of research in the behavioral sciences (Gosling & Johnson 2010; Reis & Gosling 2010). Using the Internet, researchers can deliver to participants a broad range of graphics, photographs, and dynamic media (Krantz & Williams 2010); obtain informant reports (Vazire 2010); and administer surveys (Tuten 2010), questionnaires (Johnson 2010), ability tests (Schroeders et al. 2010), and experiments (Reips & Krantz 2010). Participants can be randomly assigned to experimental conditions, reaction times can be measured, and a broad range of incentives for participation can be offered (Göritz 2010).

Internet methods offer researchers many advantages over traditional methods in terms of improved efficiency, accuracy, cost effectiveness, and reach (Gosling & Johnson 2010; Gosling et al. 2004; Reis & Gosling 2010). But how do Internet samples fare regarding Henrich et al.'s concerns about generalizability and representativeness? We do know that Internet samples are