

- racial identity revisited: An integrated conceptualization. *Child Development*, 85(1), 21–39.
- Verkuyten, M. (2016). Further conceptualizing ethnic and racial identity research: The social identity approach and its dynamic model. *Child Development*, 87(6), 1796–1812.
- Way, N., & Rogers, O. (2015). “[T]hey say Black men won’t make it, but I know I’m gonna make it”: Ethnic and racial identity development in the context of cultural stereotypes. In K. C. McLean & M. Syed (Eds.), *The Oxford handbook of identity development*. New York, NY: Oxford University Press.
- Yip, T., & Douglass, S. (2013). The application of experience sampling approaches to the study of ethnic identity: New developmental insights and directions. *Child Development Perspectives*, 7(4), 211–214.

## 2 Expansive Interdisciplinarity and the Moral Self

Javier Gomez-Lavin, Jesse Prinz, Nina Strohminger, and Shawn Nichols

### Introduction

Our investigation of the moral self has been a long journey, with various twists and turns. We knew from the start that the questions we wanted to answer—those at the center of our understanding of personhood, identity, and their interrelations with morality—would require a multi-pronged interdisciplinary approach, but we anticipated neither the range of methods that we would want to deploy, nor the lessons of those efforts. It is still a journey in progress, and we’ve been learning every step of the way. We’ve faced challenges that have required the acquisition of new skills and new research collaborators and conversation partners. Our vision of where we would go has not aligned perfectly with where we ended up, but each change in course was directed by something encountered along the way, and each re-rolling has deepened our understanding of the phenomenon we are investigating. We still have a long way to go, but there is also much to share in this progress report. Here, we take a moment to share a bit of what we’ve learned.

Central to this story is the process of doing interdisciplinary work, and frequent realizations that required us to expand our skill set. We’ve come, through this process, to think that the study of morality and the self is indeed one of the most demanding domains we know in the human sciences, one that requires moves beyond our primary fields of training: philosophy, cognitive neuroscience, and psychology. This recapitulation of our journey is also a plea for a more expansive interdisciplinary framework. We have come to think that “cognitive science”—a phrase that once labeled our conception of what it means to combine methods—is too restrictive.

### Theoretical and Empirical Background and Team Formation

Our journey began with a simple hypothesis: moral values are one of the facets of human psychology that people consider important to personal identity. Indeed, we thought moral values are more important than many other facets that have been the focus of philosophical work on personal identity, including memory, narrative, agency, cognitive faculties, and personality, among others. We call the conjunction of these two claims the “moral self hypothesis”. morals matter for identity, and they matter more than many other things. All members of our team had been exploring this hypothesis when we decided to join forces.

One of us (Jesse Prinz) had been exploring the moral self hypothesis in theoretical work for some time. He gave a theoretical paper on these ideas on March 14, 2007, in Delmenhorst, Germany, and wrote an (unpublished) article at the time. Another team member (Shawn Nichols) had begun working on the self around the same time, and was using experimental philosophy to investigate intuitions about personal identity, including an investigation of ideas stemming from the philosopher Bernard Williams. On March 21, 2010, Prinz presented his theoretical work on the moral self at a meeting of the Moral Psychology Research Group in Princeton. During this meeting, Prinz and Nichols joined forces and began to design experiments, which were eventually published in the *Routledge Handbook of Philosophy and the Social Mind* (Prinz & Nichols, 2016). Over the next few years, those experiments were refined and conducted by Nichols at the University of Arizona. Around this time, Nichols began a collaboration with psychologist Nina Strohminger, which resulted in a 2014 paper entitled, “The Essential Moral Self.” In 2015, Prinz began collaborating with Javier Gomez-Lavin, a doctoral student in philosophy who had been trained in cognitive neuroscience. The four of us decided to form a research team, and this project is the result of those efforts. We already had an overlapping history of collaboration and were unified by three things: shared faith in our hypothesis—this idea that morality and identity are deeply interconnected—a conviction that there were more questions to be answered, and confidence that a mix of philosophy and experimental psychology were the ideal tools for answering those questions. It was one of those rare and exciting moments of convergence, and we were eager and excited to embark on the next steps together. Along the way, some of this has changed. The hypothesis has undergone

some transformations—some subtle, some more significant, and some unresolved. Our sense of the most pressing questions has also evolved, and we have followed leads that we didn’t anticipate. We have also come to see that philosophy and psychology are only two tools in telling a story that actually requires many more.

### Philosophical Feedback and Conceptual Clarification

While working on our collaboration, we’ve had many opportunities to reflect on our original hypothesis. We have been led to revise and rethink various things, and some questions have opened up that remain unresolved. Some of this revision process has been conceptually driven and prompted by feedback on article submissions and from audiences at talks. We want to mention two examples here, because they underscore one key aspect of the interdisciplinary project. Though three of us are philosophers, we have tried to provide empirical evidence for the claims we are making. This requires operationalizing philosophical theories. But, once operationalized, some of our interpretations of our findings have met with philosophical resistance both from within the team and from those who learn about our work. That back and forth from philosophy to experiment, and then back to philosophy has been one of the most rewarding aspects of this endeavor.

Here we illustrate with two examples. The first example is the seminal discussion of personal identity in Western philosophy that occurs in John Locke’s *Essay Concerning Human Understanding* (Locke, 1690/1975). Locke is interested in diachronic identity: what makes a person count as the same as she or he undergoes changes over time. He offers the somewhat cryptic suggestion that temporal stages of a person are united by having the same consciousness, and, more helpfully, this gets worked out in his examples in terms of memory connections. I am linked to my past life stages through my capacity to call up past experiences in my mind. Our earlier research, before our participation in the Self, Motivation, and Virtue (SMV) grant project, had looked extensively at this memory idea and contrasted it with the idea that continuity of values might be even more important for survival over time. But, in pursuing that work, we’d left out a second major contribution of Locke’s influential discussion. He suggests there that personal identity is a “forensic” concept, meaning we need it to assign praise and blame for past events. This suggestion, which has been

widely endorsed, links personal identity to responsibility: I can be held responsible for the actions of a past self, if that self can be shown to be me, according to some criterion of diachronic identity. If intuitions about responsibility do not relate in some way to an alleged criterion of identity, that would be a reason to doubt the criterion of identity. In one of our first studies for this grant project, we tried to explore the relationship between moral continuity and responsibility.

We came up with a study design that used vignettes about parole hearings. In brief, we asked participants to imagine an incarcerated individual who had either changed moral values while in prison, or, in contrast, who had merely decided that he would refrain from future crimes because he wanted to avoid another prison term. That is, incarceration can lead to moral transformation or simply exert a deterrent effect. We then operationalized responsibility as participants' support for parole. We reasoned that changes in identity would result in a situation where the person who was coming up for parole wasn't exactly the same as the person who has committed the crime and should therefore be let out early. That is, if moral change is more related to identity than mere behavioral change (i.e., the deterrence vignette), then participants should favor parole in the former case more than the latter case even when controlling for expectations about recidivism. This is what we found. We also found that people are more likely to say the morally changed person is a different person, and this answer significantly mediates their parole decision.

In presenting this study in writing and lectures, we've met with some resistance. Some people think the finding is simply too weak to support the strong conclusion that we need: that a morally changed person is really a different individual and therefore shouldn't be held responsible for past deeds. At the same time, we have obtained survey responses that are strong and explicit. Strictly speaking we didn't predict that, since we recognize that morality is multi-faceted, and the contribution of moral continuity, though significant, does not encompass all aspects of identity. The prisoner is somewhat different as a person after a moral transformation, but not an entirely new person. Still, there is a deflationary interpretation on which he is not really a new person at all. Perhaps he is just a *better* person and should be rewarded for his progress. This response suggests that participants are answering the "new person" probe somewhat metaphorically—an interpretation that

we have tried to rule out in other work. We are not persuaded by the objection, but we do see that the study is really just the start of what must ultimately be a more thoroughgoing exploration of responsibility and the moral self. We want to pursue experimental designs that do not rely on such an indirect measure of responsibility, and we want to devise other scenarios to measure this mitigation effect. We also want to come up with additional ways to tease apart the idea that blame is reduced because of a change in identity as opposed to the competing hypothesis that it is just a reward for moral improvement. We continue to think our parole studies are suggestive, but feedback has helped us see that they can only be a first step toward answering the questions about responsibility presented in the philosophical literature.

The second example illustrating the "back and forth" between theory and experiment concerns another nagging worry that comes up when we present this material. We have conducted many vignette studies showing that moral change is interpreted as impacting identity, and that people think that a person who undergoes moral change is literally, to some degree, not the same person. But these findings can be interpreted in terms of two notions of identity: qualitative and numerical. Qualitative identity refers to attributes that affect the classification of an object (Geach, 1973). When a caterpillar becomes a butterfly, it is a different kind of animal, qualitatively speaking. Numerical identity refers to being the exact same individual. So, a butterfly may be numerically identical to the caterpillar from whence it came. The philosophical topic of personal identity is usually framed as a question about numerical identity. What makes a person the same exact individual over time? In principle, two numerically distinct people (perfect duplicates) could be qualitatively identical, and one qualitatively changing individual could be numerically the same across those changes (like the butterfly case). As of yet, our moral self-effect does not settle between these options. If moral transformations are merely qualitative, questions will arise about whether the moral self hypothesis really competes with philosophical theories of personal identity.

This is a tricky question from a methods perspective, but it has put our group into dialogue with Vilnius Dranseika, a researcher who is conducting experimental philosophy studies in Lithuanian, a language that contains different adjectives for qualitative and numerical identity. This provides a way to ask questions about the

different kinds of identity in a way that is very natural. Preliminary evidence suggests that moral changes may be merely qualitative (Berniñas & Dranseika, 2016)! This may require a revision of our hypothesis. But the jury is still out. For one thing, we are not sure about the premise that other philosophical theories really concern numerical identity. There are famous puzzles (explored by Derek Parfit in *Reasons and Persons*, 1984) about applying concepts of numerical identity to selves. Moreover, philosophical theories of personal identity, such as the Lockean memory approach, may be best interpreted as qualitative. It would be useful to obtain cross-linguistic evidence about what kind of identity is impacted by memory loss. We predict that these changes would most greatly impact qualitative identity. If so, such findings may require a clarification not just of our view, but of personal identity theories more generally. That is work in progress.

### Empirical Revisions of Hypothesis

Revision of our hypothesis has also been prompted by experimental findings. This too illustrates the value of interdisciplinary work, since the changes in question would not have been noticed without attending to empirical findings. Here we illustrate with an example that required mastery of methods that were new to some members of our team.

In our original formulation, we had proposed that moral values are important to identity. But what about other kinds of values? And what about moral norms promoted by a society that aren't valued by a given participant in our studies? We hadn't set out to answer this question, but we did sometimes include questions that used the word "moral" and others that used variants of "value." For instance, a prompt may require that a participant answer, "How morally bad would lacking *x* trait make you?" or, conversely, "How much do you value *x* trait?" where *x* would be substituted with one of many differentially valenced traits (e.g., honesty, curiosity, ebullience). In one study where we used both "moral" and "value" prompts, we noticed that responses were not identical. This ultimately required that we learn some statistical techniques and software that allow us to qualify which of these two constructs has a bigger impact on identity, including the use of advanced causal modeling methods, mediation and moderation analyses, and multidimensional scaling techniques (Hayes, 2013;

Hout, Papesh, & Goldinger, 2012). To our surprise, valuing a given trait seems to be more important than participants' ratings of how moralized a given trait is. This initial finding led us to do follow-up work, which seems to confirm that valuing drives our effects, not morality as such. We are still doing follow-up work to tease apart some inconsistencies in our findings. Here, the data will ultimately settle the hypothesis.

We came across the same issue in other ways as well. We were trying to run a cross-cultural test on one of our moral self-effects in Germany. With U.S. participants we had found that political party affiliation could serve as a proxy for moral values. When asked about a hypothetical individual who changed political parties, Americans tended to say that this was a change in identity. This pattern of responses contrasted with other kinds of changes that we had coded as non-moral, such as changes in musical taste. We wondered whether the effect would hold up in Germany because their partisanship—at least before the most recent parliamentary elections in 2017—was less fixed and polarized than in the U.S. We predicted (and obtained) a substantial weakening of the effect. Party changes are not seen as identity changes in Germany to the same degree. But we also obtained a surprising result: changes in musical taste were regarded as having a sizable impact on identity! This has led to some follow-up work, replicating and extending the finding. Some individuals seem to have "aesthetic selves," such that they see matters of taste as contributing to identity. This adds support to the idea that values outside of morality may matter to our sense of self. Once again, an unexpected empirical finding has led to new or revised hypotheses.

We first considered the very real possibility that this unexpected effect of musical taste on identity in our German sample was a fluke, or a product of an error in our experimental design. Our original study had provided the following prompt to make explicit the sense of transformation amongst our study questions: for a change in music we asked participants to "imagine that you had only liked listening to classical music, but now you find yourself enjoying pop music," and then asking them to rate how much of an impact on identity—both perceived from their own, first-person perspective and from a third-person perspective—this change would have. Of course, once we found this effect in our German sample, we immediately thought that this may have been driven by the intuitive association between classical music and a kind of

German national pride; after all, many of the “greats” of classical music share a German or Austrian heritage. Thus, perhaps our choice of example implicated political or other strongly valenced values—maybe even moralized values associated implicitly with German nationalism, and this in fact drove our effect. However, to ascertain whether this was the case, we needed to get a sense as to how musical genres were related to one another and whether this overall network of relations differed between German and American populations.

This led us to a technique, popular in psychology, known as *multidimensional scaling* or “MDS” (Hout et al., 2012). In its most simple guise, MDS requires that participants rank the similarity of a class of stimuli against one another—so for instance, if you wanted to rank the similarity of bird species, you would pick your species to be studied, and for each one, say “cardinal,” you would ask participants to give a rating, “How similar is a cardinal to *x*?” where *x* would be serially filled-in by the other species you were interested in (e.g., ostriches and blue jays). MDS has also recently been used successfully in other experimental philosophy paradigms, particularly in understanding folk perceptions of collective action (Tollefsen, Kreuz, & Dale, 2014). Using a few algorithms—including PROXSCAL<sup>1</sup>—allows us to then transform these responses into a map-like space of similarity, where similar stimuli—as rated by the participants—are grouped together. Thus, using MDS we were able to create maps of the similarity space of ten common musical genres that had been previously examined in the context of personality psychology (e.g., which Big Five personality traits correlate with common musical genres—see Renfrow & Gosling, 2003). One example map of “common space” from our German sample is reproduced in Figure 2.1.

As you can see from the map of similarity scores, pop music is centrally located, and as would be expected, classical and jazz music are grouped in one quadrant, while folk and country are closely grouped in another. Importantly, the absolute location of genres on the map is not informative; only their relative distances from one another are. Just as we can produce a “common map” that averages across all the participants in a given sample, we can also produce maps of group differences, for instance, taking the similarity judgments only of participants who scored highly on another measure, such as an authoritarianism scale, and then comparing these maps against one another.

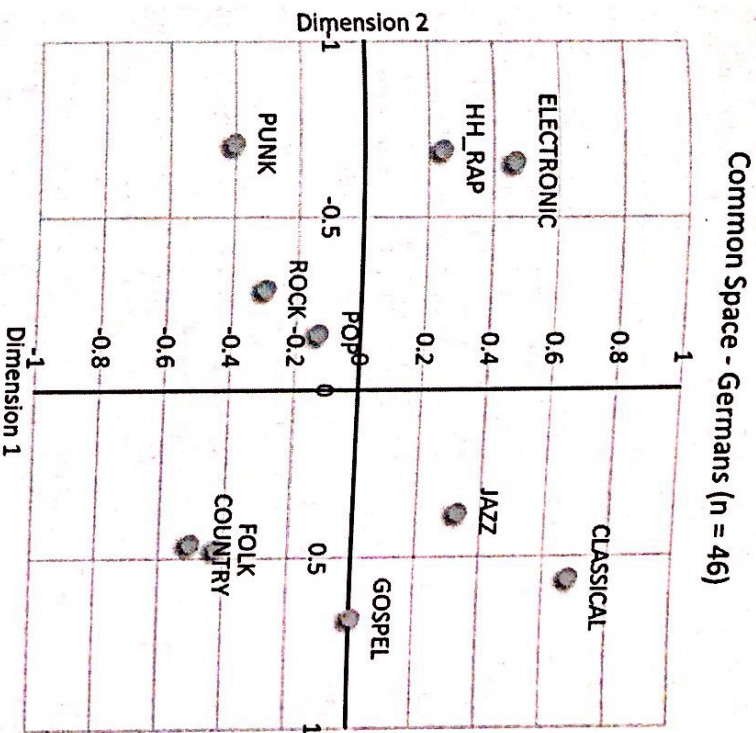


Figure 2.1 Averaged Similarity Map of Musical Genres Generated From Ratings of Our 46 German Participants Produced With the PROXSCAL Algorithm

Currently we are still sifting through the large amount of data collected, but we can state that—in general—the maps of common space between our American and German samples are similar (with the only significant difference being in the relative location of “hip-hop/rap,” coded as “HH\_RAP” in Figure 2.1). Crucially for our original question, the relative locations of pop and classical genres in both our American and German sample are very similar—hence, the relative perception of musical genres is not driving our effect in Germany. Further exploratory work conducted in conjunction with Joerg Fingerhut and others in the Berlin School of Mind and Brain has revealed that this effect—the impact of aesthetics on

identity—generalizes. We are in the process of developing more measures to isolate this effect of “aesthetic selves” and its unique mechanisms.

### Cross-Cultural Psychology

Our studies in Germany—along with later work in a similar vein in Taipei—have required acquisition of new methodological competencies and new partnerships. Some team members had never done cross-cultural studies, and we needed to learn some basic things: how to work with an internal, university ethics review board abroad; how to recruit participants; how to create and test a translation; and so on. We also needed to learn about other cultures to generate hypotheses, and we needed to reflect on the pros and cons of cross-cultural work.

Beginning with the last of these issues, we had come to think that cross-cultural work tended to be vulnerable to two opposing risks: a kind of Scylla and Charibdis for cross-cultural psychology. Some of this work is too universalizing, brushing over significant cultural differences in order to establish that a phenomenon is pan-cultural. In testing the moral self hypothesis overseas, we didn't want to assume its truth or construct designs that would overlook differences. We wanted test instruments that were sufficiently sensitive to pick up on subtle cultural differences in the phenomenon. A related risk is that some cross-cultural research is too simplistic in its analyses of difference. A surprising volume of work in this area uses the East/West contrast and distinguishes these using a single coarse dimension, such as collectivism versus individualism. Such work can be informative, but it blurs over differences that exist between Eastern populations and Western populations, and it tends to promote harmful, essentializing stereotypes such as “the Asian mind” (Markus & Kitayama, 1991; Nisbett, 2003).

To overcome these risks and hurdles, our challenge was to formulate hypotheses that did not rely on coarse stereotypes and distinctions that blur important differences. Our aforementioned choice to investigate partisanship and identity in Germany was part of this effort. We want to show that not all Western nations respond in the same way. Likewise, when testing our vignettes and prompts in Taipei, we focused on perceived differences amongst major religions and their affiliated practices and cultures in Taiwanese society. Morality may be important to identity in both

cultures, but it gets grounded in different social-group affiliations. We needed to find informants who could tell us about political attitudes in Germany, subtle religious differences in Taiwan, and their relations to morality. We learned many things. For example, Germans tend to think it is wrong to publicly discuss who they voted for, and class structure is often interwoven with religious background in Taiwan.

We conducted studies in Taiwan, a far Eastern setting, for two principle reasons. We were curious whether moral self-effects could be found outside of the West, and we also wanted to exemplify an East/West comparison that did not rely on the usual reductive assumptions about individualism and collectivism. Having looked at political party affiliations as a proxy for morality in the United States and Germany, we wanted to do the same in Asia. We considered Japan, but its biggest parties are very closely related, and China, but it is really a one-party system. We chose Taiwan, which has significant, popular religious and political divisions, and we set to work trying to understand these. This required a lot of research. We read about the history of Taiwan, its religions, its political parties, and its elections. We also identified collaborators at Taipei Medical University and asked them about local politics and values. This was a very new way to do research for us, and it was both challenging and exciting. We made many discoveries. The most interesting, perhaps, is that the liberal/conservative distinction, which is so divisive in the United States, is not viewed as the major political fault line there. Instead, parties and people divide on attitudes toward China: are the Han people of Taiwan the true Chinese who will eventually reunite with the mainland, or are they Taiwanese—a truly separate cultural and political entity? We also learned much about religion in Taiwan, for example, that Buddhism is considered more refined than Taoism, and that Mandarin has no word for Taoism. We worked closely with psychologists, philosophers, and neuroscientists in Taipei to develop a survey instrument. We confirmed the moral self effect but in a way that reflected social identities specific to Taiwan.

Our cross-cultural work has been a steep learning curve, and our international collaborations will continue. We wanted to explore both universality and difference—a pan-cultural tendency to link values and identity, but in culturally specific ways. This is what we have found, and we are excited to be doing East/West comparisons that move beyond familiar stereotypes.

**Shifting Priorities: Real-World Cases**

Along the way, there has also been a shift in our priorities, and that shift has played a guiding role in our interdisciplinary collaboration and even in our idea of what interdisciplinary work should look like.

With backgrounds in philosophy, psychology, and neuroscience, we began with an interest in a psychological effect—namely, our robust, repeatable moral self-effect—and the mechanisms that underlie it. The effect had already been established in our earlier work, and we thought we'd spend a lot of time looking at mechanisms in our collaboration. In thinking about mechanisms, we were really focused on internal mechanisms—more specifically on psychological constructs such as emotion and executive function. In our prior work, we had also mostly explored the moral self using vignettes that describe morality in a very generic way: “imagine a person whose moral values change . . .”.

As our work began, these priorities guided our efforts. We developed an elaborate reaction time study to look at the role of executive function, and team members also carried out a study looking at the impact of personality variables, such as psychopathy. These efforts were worthwhile, but we were increasingly convinced that the most interesting thing about the moral self hypothesis lies elsewhere. We began to think that our vignettes were, like many thought experiments in philosophy, too hypothetical. We wondered, do moral transformations happen in the real world? Two team members (Nichols and Strohlinger) had published a study that used both hypothetical vignettes (e.g., a brain transplant—Strohlinger & Nichols, 2014) and scenarios that could occur in real life (e.g., cases of dementia; see in particular, Strohlinger & Nichols, 2015). We began to think more broadly about real-world cases, and we came up with many examples. This, we now think, is the most interesting, important, and pressing direction for our research—that is, to look back on the plethora of examples of moral transformation and identity change provided by history itself.

Our study of parole intuitions was a move in that direction. Prison reform is one real context where values can change. We plan to continue that work in a more applied way, looking at actual parole decisions and other archival work. Our studies of partisanship also fall into this category, and we have conducted several studies looking at attitudes toward religious conversion (as well as some theoretical work on “cults” or new religious movements).

To take another example, we became interested in the European immigration crisis. We conducted a series of studies in Germany asking people about their attitudes toward immigrants. We found that when Germans consider immigrants who come from an Islamic country, they regard it as positive when they assimilate in their values, not just their behavior. They also consider value assimilation to involve a change in identity. In effect, they want immigrants to become new people.

An additional example concerns something called “moral injury.” The term refers to a kind of post-traumatic stress experienced by soldiers who are deployed in combat zones. Witnessing or participating in acts of war can make people feel like they have confronted behavior that violates their moral norms. This can make reimmersion in civilian society difficult. In some exploratory studies we showed that people regard such changes as moral in nature and as changes in personal identity. We hope to follow up on this work with military populations with collaborators from other SMV project teams. In particular, we want to examine military personnel before and after engaging in recruit, or “boot camp” training, to measure whether moral values are significantly affected by this process. Additionally, partnering with medical professionals who work primarily with veterans may allow us to empirically test whether moral injury plays an important role in veterans’ mental health, including in pathological cases such as post-traumatic stress disorders.

All these studies have come with their own challenges. We have had to learn about parole practices, international politics, and military psychiatry, among other areas. This has required a lot of reading in fields outside our areas of training. We have also opened dialogue and collaborations with people in other fields and other countries. One team member, Jesse Prinz has also been teaching a graduate seminar in political psychology with a number of political science students, to get up to date on studies of political attitudes. Learning these things has been one of our biggest challenges, and one of the more rewarding aspects of this collaboration.

We’ve also come to see that real-world cases require something beyond knowledge of current events; they require knowledge of history. We realize now that it would have been valuable to have a historian on the team (and perhaps a political scientist and a sociologist). We have been looking at historical events that involve moral change. Examples include: the emergence of liberal democracies, the rise of the Third Reich and subsequent de-Nazification

programs, and the Cultural Revolution in China. We want to understand how these changes occurred, and how they impacted identity. For example, supporters of Mao frequently talked about the construction of a "New Man," and re-education initiatives were described as transformative. We have been reading cultural and political history, and also scouring qualitative research for evidence of a link between values and identity. For example, Theodore Newcomb's classic study of political indoctrination in college (*Personality and Social Change: Attitude Formation in a Student Community*, 1943) contains many interviews with students. Reading these, we've been struck by the use of language that makes explicit references to identity transformations.

All of these efforts have required the pursuit of new skill sets. Reading history, looking through databases for primary sources (e.g., articles on de-Nazification efforts published just after the war), and scrutinizing qualitative research were not part of our standard repertoire before we started. We've opened up dialogues with new conversation partners, too, and we are seeing new avenues for research. The move to real-world cases has made it vividly clear to us that our previous fields of training—philosophy, cognitive neuroscience, and psychology—work at a level of abstraction that tends to block off many dimensions of human variation. These fields tend to treat all people as alike and ignore the impact of life experiences and history. There are exceptions of course, but efforts to include such dimensions of variation are often done without the hard toil of studying social and historical conditions in detail. We have realized how important this is, how hard it is, how much we have to learn, and how much we can benefit from the expertise of others.

### Conclusion: Toward an Expansive Moral Psychology of the Self

In conclusion, ours has been a journey of changing paths and discoveries. Our thinking has evolved, our priorities have changed, and our skill sets are slowly expanding.

Some of the challenges we met were expected. Combining philosophy and psychology has been a central goal for each of us individually, and, in collaboration, we have found a constant need to move back and forth between theory and experimental design. This has involved spelling out theoretical statements, operationalizing, looking at results, getting feedback, and revising theories

accordingly. A constant, mutually informative dialogue exists between theory and data.

There have also been some challenges that required new skills within this framework of psychologically informed philosophizing. Though our initial studies began with the research tools of experimental philosophy, with a focus on vignette paradigms and rather simpler statistical analyses, we quickly found ourselves learning and using more sophisticated statistical methods to model the relationship among our variables of interest. For instance, modeling the interactions of our dependent variables—that is, changes in the measures that we're interested in closely examining, such as changes to perceived identity—required that we use both structural equation and causal mediation models to help us determine the directionality of our findings.<sup>2</sup> This was particularly relevant when we were examining participants' judgments about responsibility in our parole case; after all, we wanted to determine that it was participants' perceptions of identity change that determined their parole decisions. These models allowed us to test our hypothesis that participants' attitudes about whether a character's identity has changed were the driving force for their judgments about the dependent variables of interest; for example, responsibility in our parole studies, and assimilation in our immigration studies. Importantly, use of these models allowed us to probe the possible causal relationship among the various independent and dependent variables of interest—for instance, seeing whether perceived changes in identity are a direct result of a moral change or if there is some mediating factor, such as the perceived value of the trait in question, that also influences participants' judgments of identity. In our research on the aesthetic self, one philosopher in our team (Gomez-Lavin) also learned multi-dimensional scaling; this technique allowed us to visualize how individual differences—say, participants' scores on a third-party scale of authoritarianism—affected how they conceptualized the similarity among aesthetic genres. For instance, we've found that high-scorers on authoritarianism scales tend to view hip-hop and rap music as the least similar to other genres, yielding significantly different similarity spaces than other groups. These similarity spaces are constructed by taking into account hundreds of participant ratings of the similarity of various, traditional musical genres (e.g., we would ask participants to rate how similar hip-hop is to punk music and so forth). These new skills will allow us to do things that go beyond what is found in most experimental philosophy.



More surprising was the realization that our hypothesis would benefit from methods that go beyond philosophy and experimental psychology. We have had to learn some political science, some psychiatry, some criminology, and some history. We have pursued these efforts with passion, though much work lies ahead. Some of us remember those early days when philosophers first started running psychological studies. The early studies were primitive and flawed, but also empowering. If we compare the skills of those of us in that first wave (Nichols and Prinz) with those in the new generation (Gomez-Lavin), it is very gratifying. Some philosophers are now doing work that meets the standards of psychology. These advances have been possible in part because of psychologists (e.g., Strohminger) who are interested in philosophical ideas and are willing to work with philosophers. Both philosophy and psychology have benefitted by those ongoing alliances.

We are now, perhaps, in a similar position with respect to fields such as history. Cognitive science has left such fields out, even though they have much to teach us about human behavior. We are in the foothills of a daunting learning curve but united in the recognition that we would benefit from moving beyond the form of interdisciplinarity that has been institutionalized as cognitive science. When dealing with constructs as rich as morality and identity, mental processes and mechanisms can only take us so far. The micro-level of analysis might be combined with the macro-level. Our journey has, in many ways, been a journey to that point of discovery. We've made progress on our original hypothesis, but that may be less important, ultimately, than the methodological lessons. More than anything, we emerge with a deep and humbling appreciation of the limitations of our training, the narrowness of our conversations, and the need for a more expansive approach to the study of the human mind.

## Note

1. PROXSCAL stands for PROXimity SCALing and enables the scaling of relative similarity, or dissimilarity, distance-like data from a higher dimensional to a lower dimensional space that can easily be visualized. For more information, please consult Commandeur and Heiser (1993).
2. These variables contrast with what psychologists term "independent variables," or the manipulations that we, as the experimenters, introduce into our in which our character undergoes a moral transformation, while in the other the character only undergoes a behavioral shift.

## Bibliography

- Berniñas, R., & Dransseika, V. (2016). Folk concepts of person and identity: A response to Nichols and Bruno. *Philosophical Psychology*, 29(1), 96–122.
- Commandeur, J. J. F., & Heiser, W. J. (1993). *Mathematical derivations in the proximity scaling (PROXSCAL) of symmetric data matrices* (Tech. Rep. No. RR-93-04). Leiden, The Netherlands: Leiden University, Department of Data Theory.
- Geach, P. (1973). Ontological relativity and relative identity. In M. K. Munitz (Ed.), *Logic and Ontology*. New York, NY: New York University Press.
- Hayes, A. E. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.
- Hout, M. C., Papeish, M. H., & Goldinger, S. D. (2012). Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(1), 93–103.
- Locke, J. (1690/1975). An essay concerning human understanding. In P. H. Niddich (Ed.), *The clarendon edition of the works of John Locke*. Oxford: Clarendon Press.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253.
- Newcomb, T. M. (1943). *Personality and social change: Attitude formation in a student community*. Fort Worth, TX: Dryden Press.
- Nisbett, R. E. (2003). *The geography of thought: How Asians and westerners think differently—and why*. New York, NY: Free Press.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Oxford University Press.
- Prinz, J., & Nichols, S. (2016). Diachronic identity and the moral self. In J. Kiverstein (Ed.), *The Routledge handbook of philosophy and the social mind*. Abingdon: Routledge.
- Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *The Journal of Personality and Social Psychology*, 84(6), 1236–1256.
- Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, 131(1), 158–171.
- Strohminger, N., & Nichols, S. (2015). Neurodegeneration and identity. *Psychological Science*, 26(9), 1469–1479.
- Tolfsen, D., Kreuz, R., & Dale, R. (2014). Flavors of "togetherness". In J. Knobe, T. Lombrozo, & S. Nichols (Eds.), *Oxford studies in experimental philosophy* (Vol. 1). Oxford: Oxford University Press.