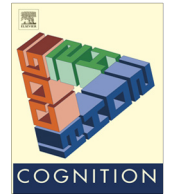




ELSEVIER

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

## Cognition

journal homepage: [www.elsevier.com/locate/COGNIT](http://www.elsevier.com/locate/COGNIT)

## The essential moral self

Nina Strohminger<sup>a,\*</sup>, Shaun Nichols<sup>b</sup><sup>a</sup> Kenan Institute for Ethics, Duke University, United States<sup>b</sup> Department of Philosophy, University of Arizona, United States

## ARTICLE INFO

## Article history:

Received 28 May 2013

Revised 14 December 2013

Accepted 17 December 2013

## Keywords:

Personal identity

Numerical identity

Self

Morality

## ABSTRACT

It has often been suggested that the mind is central to personal identity. But do all parts of the mind contribute equally? Across five experiments, we demonstrate that moral traits—more than any other mental faculty—are considered the most essential part of identity, the self, and the soul. Memory, especially emotional and autobiographical memory, is also fairly important. Lower-level cognition and perception have the most tenuous connection to identity, rivaling that of purely physical traits. These findings suggest that folk notions of personal identity are largely informed by the mental faculties affecting social relationships, with a particularly keen focus on moral traits.

© 2013 Elsevier B.V. All rights reserved.

In 1848, a 13-pound tamping iron shot through the skull of a 25-year-old man named Phineas Gage, taking a chunk of his brain with it. Formerly mild-mannered and responsible, Gage emerged from the accident impulsive and foul-tempered. His character changed so markedly that those who knew him said he was “no longer Gage” (Macmillan, 2000). Likewise, in describing the radical memory loss experienced in the wake of Korsokov’s syndrome, Oliver Sacks (Sacks, 1985) wonders whether his patient has become “de-souled”.

When someone undergoes dramatic mental change, their numerical identity—whether they’re the same person as they were before—can seem to become disrupted.<sup>1</sup> While the philosophical literature has focused on metaphysical claims about personal identity, the scope of this paper is descriptive: what drives lay theories of numerical identity

across different sorts of mental transformation? What do people consider the most essential parts of the self?

Some philosophical accounts of personal identity have advocated the importance of physical continuity (Ayer, 1936; Williams, 1973; Thomson, 1997), but most current discussions revolve around psychological continuity (Locke, 1690/2009; Parfit, 1971; Shoemaker & Swinburne, 1984; Unger, 1990; Olson, 2003). Folk intuitions largely accord with the psychological view. When children are asked to imagine what would happen to a hamster put in a “duplication” device, they predict the duplicate will inherit the original hamster’s physical traits, but not its memories (Hood, Gjersoe, & Bloom, 2012). Young children place more weight on behavioral changes than physical changes in determining identity continuity, even when that physical change involves a brain transplant (Johnson, 1990). It is only after children learn the brain is the seat of the mind that brain transplants are seen as altering identity (Gottfried, Gelman, & Schultz, 1999). (As Dennett (1978) points out, the brain is the only organ of the body where it is preferable to be the donor than the recipient.) The intuition that the mind is an identity-conferring and body-independent entity persists into adulthood (Corriveau, Pasquini, & Harris, 2005; Blok, Newman, & Rips, 2005).

\* Corresponding author.

E-mail address: [ns141@duke.edu](mailto:ns141@duke.edu) (N. Strohminger).

<sup>1</sup> Some philosophers claim that personal identity is absolute: you’re either Nina or you’re not (Butler, 1736/2008; Reid, 1785/1850). Empirical research, however, finds that people judge a partially modified self to be capable of surviving partially (Libby & Eibach, 2007; Bartels & Urminsky, 2011; Bartels, Kvaran, & Nichols, 2013), consistent with the idea that identity can be graded and relative (Hume, 2000; Parfit, 1971).

The mind may be central to notions of the self, but are all mental faculties equally important? This question has received scant attention, so here we lay out some possibilities. Perhaps identity persistence depends only on the magnitude of mental change. In that case, the mental faculties relating to overall functionality and consciousness—such as perception and low-level cognition—might be the most potent determinants of identity. In spite of their broad-ranging functional impact however, it is far from clear that basic cognitive faculties are foundational to identity: they are not terribly unique from person to person, and have little relation to why we would care about an individual. Indeed, this may be a domain of the mind that contributes only minimally to personal identity.

A long philosophical tradition links identity to memory, particularly autobiographical memory (Locke, 1690/2009; Shoemaker, 1959; Williams, 1970; Parfit, 1971; Perry, 2002). Autobiographical memories provide a continuous inner narrative which may be required for a unified sense of self. Social scientists would also predict memory to be important, though on different grounds. A number of theories propose that what matters most to social and group identity is distinctiveness (McGuire & Padawer-Singer, 1976; Nelson & Miller, 1995; Vignoles, Chryssochoou, & Breakwell, 2000; Blanton & Christie, 2003). To the extent that one's collection of memories represents a uniquely identifying set, memory ought to play a central role in the construction of identity. Indeed, recent empirical research on numerical identity makes precisely this point, that episodic memory is paramount because of its novelty (Hood et al., 2012). Previous work has found that memory loss disrupts identity judgments (Blok et al., 2005; Nichols & Bruno, 2010), though these studies do not compare memory to other mental traits.

Memory is not the only part of the mind that can be distinctive. Desires and preferences—such as liking archeology, being afraid of snakes, and enjoying adventure—exhibit high heterogeneity between individuals, as do personality traits. Dispositional traits have long been considered at the heart of person perception (Allport, 1937), and people report that personality traits and preferences are an essential part of who a person is (Haslam, Bastian, and Bissett (2004), Gelman, Heyman, and Legare (2007)). Whether these traits have direct relevance to numerical identity, and how they might compare with other types of mental content, remains untested.

While memory often receives headlining status in discussions of identity, some philosophers have suggested that morality is an important (Parfit, 1984) or perhaps even the most important (Prinz, *in press*; Prinz & Nichols, *in press*) part of personal identity. A small but growing litany of evidence points towards this possibility. People are reluctant to take pharmaceutical enhancements for traits that are considered fundamental to the self; two moral traits (empathy and kindness) top this list (Riis, Simmons, & Goodwin, 2008). Children judge moral goodness to be a more stable dispositional trait than other personality traits, including intelligence (Heyman & Dweck, 1998; Haslam et al., 2004), and moral attributes are predominant in person perception (Skitka, Bauman, & Sargis, 2005; Goodwin, Piazza, & Rozin, 2014). Willingness to attribute moral

change to the true self is contingent upon one's pre-existing moral beliefs (Knobe, 2005; Newman, Knobe, & Bloom, 2014). Finally, the concept of the soul—by some counts a placeholder for the self at its very pith—carries with it strong moral connotations (Shweder, Much, Mahapatra, & Park, 1997; Bering, 2006; Richert & Harris, 2006).

Our aim in this paper is to take a systematic approach to determining which parts of the mind are most central to personal identity. Of particular interest is the possibility that moral traits are more essential than any other mental feature, including those that provide functionality, distinctiveness, or personal narrative (henceforth 'the essential moral self hypothesis').

We test this hypothesis across five diverse scenarios. In Study 1, we examine how different forms of neuropsychological impairment impact judgments of numerical identity. In Study 2, we look at how change to a broad array of traits brought about by voluntary pharmaceutical intervention affects identity persistence. In Studies 3 and 4, we extend these findings to the soul concept, by probing intuitions about the properties a soul exports when it leaves the body and is placed into a new one. Finally, in Study 5, we measure identity continuity in the face of age-related cognitive change.

## 1. Study 1: The brain transplant

This study examined the impact of cognitive impairments upon willingness to attribute continued identity to a person following brain trauma. We adapted a method used in previous research for studying numerical identity (Blok et al., 2005). In the original paradigm, subjects read a story where a man living in the not-distant future, Jim, gets into a car accident and needs a brain transplant. After the brain transplant, Jim is either psychologically identical or has lost his memories. The authors found that identity (to wit: "Is the transplant recipient still Jim?") is more likely to change after a discontinuity of mental content than a discontinuity of brain matter alone (Blok et al. (2005)).

As we were interested in distinguishing among different kinds of psychological change, we constructed variations where Jim selectively loses different parts of his mind. In the wake of the surgery, Jim either experiences no cognitive change (control condition), inability to recognize objects (visual object agnosia), loss of autobiographical memories (amnesia), loss of desires (apathy), or loss of moral conscience (see Appendix A for text of the stories).

If physical continuity alone is sufficient to disrupt identity, then these cognitive changes should have no additional effect on identity. Likewise, if the mind is uniformly more connected with the self than the body, then all psychological changes should alter identity more than physical changes. The inclusion of visual object agnosia was intended as a test of whether all cognitive deficits—including, in this case, one leading to a profound functional impairment—would dramatically affect identity.

Apathy is a common clinical syndrome in neuropsychiatric disorders, and is distinct from clinical depression (Starkstein, Petracca, Chemerinski, & Kremer, 2001).

Broadly speaking, a person's unique pattern of desires and preferences fulfills the distinctiveness principle of identity (Vignoles et al., 2000). To remove desire is to remove these potentially distinguishing characteristics.

Amnesia provides another opportunity to test the distinctiveness principle. Although past work on psychological continuity speaks of “memory” in generic terms, it seems that episodic memory is what such discussions are intended to draw out. For example, (Locke's 1690/2009) story of the prince and the cobbler describes memory as “the consciousness of the prince's past life.” Recollections of one's own experiences—unlike knowing how to do a cartwheel or knowing the capital of France—are expressly linked with the self. When subjects consider cases that refer simply to “memory loss” as in Blok et al., 2005, they may assume that this refers to autobiographical memory, or they may interpret the question more inclusively. We wanted to ensure there was no ambiguity on this point.

While memories have been a focal point in many philosophical debates over identity, moral capacities may also be central. There are several ways that a person's moral capacities can change; in this study we ask what will happen following a wholesale loss of the moral faculty.

Study 1 has a few novel predictions. The first is that not all cognitive deficits will be viewed as equally injurious to the self, with a low-level cognitive deficit (agnosia) having a negligible effect on identity. Loss of desires and memories will be moderately linked to identity, although a moral deficit will lead to the largest identity disruption.

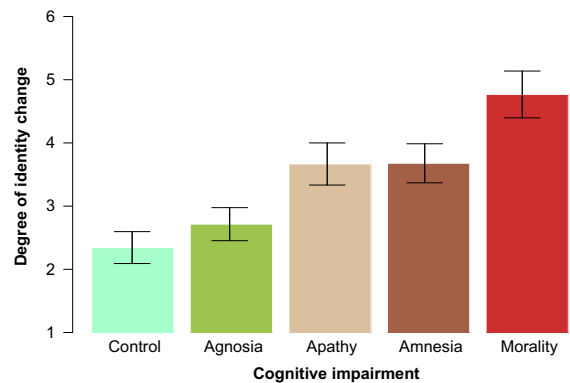
### 1.1. Methods

148 American participants (*Mdn* age = 30.5, 58% female) participated in an online study. In a between-subjects design, subjects were presented with one of five versions of the partial brain transplant story (see Appendix A). After reading the story, subjects were asked to what extent they agreed with the statement “The transplant recipient is still Jim.” They responded on a 1–7 Likert scale, where 1 was “Completely agree”, 7 was “Completely disagree”, and 4 was “Neutral”. Afterwards, participants were asked to explain their answer.

### 1.2. Results and discussion

Results were in line with our predictions. The low-level cognitive deficit (agnosia) did not lead to a change in identity ( $M = 2.7$ ) compared with control ( $M = 2.3$ ;  $t(57) = 1.02$ ,  $p = .31$ ). Autobiographical amnesia led to greater identity loss compared to the control condition ( $M = 3.7$ ;  $t(58) = 3.34$ ,  $p = .001$ ), as did apathy ( $M = 3.7$ ;  $t(50) = 3.16$ ,  $p = .003$ ). Loss of the moral faculty produced the most dramatic change in identity. While there was no significant difference in the average response between amnesia and apathy ( $t(55) = .02$ ,  $p = .98$ ), a planned contrast revealed that a person with a moral deficit is considered more profoundly changed than one who loses memories or desires ( $M = 4.8$ ;  $F(1,146) = 6.94$ ,  $p = .009$ ). All other pairwise comparisons were significant (see Fig. 1).

The written responses provide an opportunity to probe this effect a bit deeper. Some participants explained their



**Fig. 1.** The impact of cognitive impairment on identity loss, with standard error bars. Higher score indicates lesser agreement that the transplant recipient is the same person as before surgery. “4” represents the “Neutral” midpoint of the scale. All pairwise comparisons are significant ( $p < .01$ ), except Control vs. Agnosia and Apathy vs. Amnesia.

answer in terms of the cognitive deficit itself, whereas others mentioned the incidental effects of the surgery. For example, this respondent pinpointed memory *qua* memory as identity-conferring: “Memories and relationships are a huge part of who you are.” Another subject reported that amnesia alters identity only to the extent that it leads to a deeper change: “How can he still be Jim when he does not remember anything taught or learned from the time he grew up until the event? He would have a different character/personality. Therefore he would be a different person.” We conducted a secondary analysis see if explanations for what led to identity change were systematically different across conditions.

Three coders blind to the purposes of the study categorized subject explanations for whether identity change was a function of the deficit itself, or whether it was due to something else. An interrater reliability analysis using the Fleiss Kappa statistic was performed; this test revealed high agreement among raters (Kappa = 0.92,  $Z = 21.6$ ,  $p < .0001$ ). Within the morality condition, 61% of respondents spontaneously volunteered that morality is what gives a person his identity. For the memory and apathy conditions, this number was 34% and 23% respectively. 13% of respondents thought that object recognition itself confers identity, and no respondents in the control condition mentioned that brain tissue confers identity.

The moral faculty is part of the mind most likely to be seen as the ultimate explanation for whether a person's identity endures or fades away. That is, subjects in the moral deficit condition did not, by and large, recourse to explaining its importance in terms of other underlying factors. A typical explanation reads: “Our moral conscience, our moral compass, is a huge component of what makes up our identity and our soul.” Another subject explains (emphasis ours), “The surgery resulted in Jim losing his moral conscience and his ability to empathize with the sufferings of others. These are *essential* aspects of personal identity so I concluded that in a profound way Jim is no longer himself after the surgery.”

Overall, these findings support the essential moral self hypothesis. Although this study is illuminating, it only

examines a few cognitive abilities, and collapses across traits which may look quite different when considered separately. For example, every type of desire and memory may not be equivalent in its contribution to the self. The goal of the following studies is to consider these issues in finer detail.

## 2. Study 2: The silver bullet

In the previous study, we looked at identity change brought on by traumatic injury. But sometimes cognitive change is intentional, as when people seek out psychopharmaceutical interventions. In such cases, identity change can be a concern (indeed, it is sometimes the goal; Kramer, 1993).

Despite the differences in these two types of case—one pathological and organic, the other therapeutic and chemical—we may nonetheless observe identity change across the same set of traits. The aim of this study was to see if the same basic pattern observed in Study 1 would obtain for a radically different mode of brain change. Additionally, we wanted to examine a range of specific traits to learn more about the categories under investigation.

### 2.1. Methods

209 Americans (*Mdn* age = 27, 59% female) were paid to participate in an online study.

Participants were asked to imagine that pills had been developed that, once swallowed, would permanently alter only one part of a person's mind, without affecting anything else. Subjects rated how much a person would change after taking a pill that selectively removed one of 62 cognitive/behavioral traits, on a sliding scale from 0% ("They're the same person as before") to 100% ("They're completely different now"). The slider position could be adjusted with a mouse. The numerical score appeared to the right of the item and the slider, so that traits could be compared. All traits appeared on the same page, and in randomized order between subjects. All subjects rated all items. Afterwards, participants answered a short battery of individual differences and demographic measures.

Items were selected to span a relatively representative sample of cognitive capacities, including those that are the central focus of this paper: perception, desires and preferences, memories, and morality. We included a few types of memory, as the results from the previous study leave open the possibility that some memory types are more connected to identity than others. It may be the case that memory *per se* is not what people think is essential to the self, but specific types of memory, or memory that contains certain sorts of information.

To determine categories for analysis, five coders blind to the aims of the study rated which of the four categories each of the 62 traits fit into. An item could be placed in more than one category, and if an item did not fit any of the categories, it was excluded from the analysis. In order to earn placement in a category, at least 4 out of the 5 coders needed to be in agreement over category membership. The "desires and preferences" category contained several

traits that scored high in both the morality and the desires categories (e.g. pedophilia). To ensure that only non-moral desires and preferences were included in this category, the desires score was determined by subtracting the moral category score from desires category score. A score of 4 or higher was then used to determine inclusion in the desires category. For the selected items and their categories, see Table 1.

### 2.2. Results and discussion

Paired *t*-tests were performed for all pairwise comparisons among category of mental change; *p*-values were adjusted using the Holm correction to account for multiple comparisons (Holm, 1979). Subjects considered a person most dramatically changed if the pill altered moral traits, with all other traits leading to less dramatic change (see Fig. 2). The morality condition ( $M = 60.7$ ,  $SD = 30.5$ ) resulted in greater identity change than perception ( $M = 37.7$ ,  $SD = 31.4$ ,  $t(208) = 13.9$ ,  $p < .0001$ ), desires ( $M = 40.4$ ,  $SD = 31.4$ ,  $t(208) = 15.4$ ,  $p < .0001$ ), memory ( $M = 39.2$ ,  $SD = 33.0$ ,  $t(208) = 16.6$ ,  $p < .0001$ ), and personality ( $M = 51.8$ ,  $SD = 28.7$ ,  $t(208) = 8.7$ ,  $p < .0001$ ).

Personality traits were significantly more related to identity than memory ( $t(208) = 9.7$ ,  $p < .0001$ ), desires ( $t(208) = 9.6$ ,  $p < .0001$ ), or perception ( $t(208) = 8.7$ ,  $p < .0001$ ). Consistent with Study 1, no significant difference was found between the memory and desires categories ( $p = .40$ ). The lower-level cognitive category in this study—perception—did not score significantly lower than memories ( $p = .40$ ) or desires ( $p = .08$ ). For individual item scores, refer to Table 1.

In Study 1, we focused on pathology: what would happen to identity when a cognitive capacity was removed wholesale. In the present study, the pill could cure ailments (e.g. pedophilia) or create them (e.g. colorblindness), as well as confer more neutral results (e.g. changing music preferences). Regardless of whether the outcome was pathological, we observe the same general pattern of results.

Traits that scored unusually high within the memory, desires, and perception categories provide additional insight. In the desires category, none of our coders rated homosexuality as related to morality, an attitude that might not have been shared by the participants of the study: sexual practices are often moralized, especially by more conservative populations (Haidt & Hersh, 2001). In perception, ability to feel pain had an unusually high score; perhaps this is because of pain's relationship to the moral status of sentient beings (Wallace, 2004; Gray, Young, & Waytz, 2012).

In Study 1, episodic memory amnesia led to identity change, leaving open the question of how this effect would generalize to other forms of memory. In the present study, the extent to which we are defined by our memories depends on what kind of memories they are. The lowest-ranked of the 62 items were related to memory (knowing how to ride a bike and memories of time spent commuting to work); these scored lower even than the perceptual traits. Put another way, people believe that putting on a pair of glasses constitutes a greater threat to identity than

**Table 1**

Individual trait scores for the four psychological categories in Study 2. Score represents the percent change a person would undergo if they took a drug that would remove that trait.

CATEGORY	ITEM	MEAN (SD)
MORALITY	Psychopathy	73.6 (27.9)
	Pedophilia	70.4 (33.2)
	Criminality	64.8 (30.2)
	Empathy for the suffering of others	64.3 (30.3)
	Being a jerk	64.0 (29.5)
	Conscientiousness	61.1 (29.7)
	Virtuousness	57.1 (29.8)
	Politeness	56.4 (28.3)
	Womanizing	55.5 (29.2)
	Kleptomania	55.3 (28.8)
	Cowardice	53.5 (27.9)
Religious piety	52.1 (33.1)	
PERSONALITY	Shy	57.3 (29.3)
	Anxious	53.8 (28.5)
	Industrious	50.5 (27.2)
	Absentminded	45.7 (28.7)
MEMORIES	Traumatic memories	65.5 (30.7)
	Cherished memories of time spent with parents	53.4 (32.0)
	Knowledge of math	44.5 (30.3)
	Knowing how to play the piano	29.5 (27.7)
	Memories of time spent commuting to work	21.7 (25.0)
Knowing how to ride a bike	20.4 (24.4)	
DESIRES & PREFERENCES	Homosexuality	58.1 (35.6)
	Enjoyment of sex	52.4 (29.6)
	Desire to eat healthy	43.8 (28.1)
	Craving cigarettes	38.9 (31.5)
	Desire to quit smoking	36.4 (31.0)
	Wanting to be a doctor	35.5 (29.0)
	Enjoyment of rock music	29.3 (25.5)
Enjoyment of a favorite food	28.4 (27.8)	
PERCEPTUAL	Ability to feel pain	57.6 (31.8)
	Ability to see color	40.6 (31.2)
	Ability to smell	39.5 (33.7)
	Ability to appreciate art	38.7 (29.3)
	Nearsightedness	32.0 (30.6)



**Fig. 2.** The impact of psychoactive ‘magic bullet’ drugs on identity change, with standard error bars. Higher score indicates greater percent change to the self for the set of drugs impacting traits within that category. All pairwise comparisons between morality and the other categories are significant ( $p < .0001$ ).

wiping out an entire tract of experiential memories! On the other hand, episodic memories with social and personal reverberation—whether positive or negative—are

considered quite important. These results may be understood as consistent with the essential moral self hypothesis: memory is not important unto itself, but rather for the connections it affords us to our socio-moral core.

Two additional items, which were related to object recognition, hit this point home further. We asked how changed someone would be if they lost the ability to recognize celebrities (a sort of Hollywood prosopagnosia) or the ability to recognize their loved ones. While these are both the same sort of cognitive deficit, agnosia for loved ones leads to a much greater identity shift than agnosia for celebrities ( $M_{family} = 69.0$ ,  $M_{celebrities} = 18.9$ ,  $t(208) = 18.3$ ,  $p < .0001$ ).

The pattern of findings in this study again suggests that the moral self is central to identity. To the extent that other mental traits are considered important to identity, this appears to be driven, at least in part, by that trait’s relationship to our socio-moral lives. This data puts a finer point on past discussions of the role of memory in personal identity (e.g. Locke, 1690/2009): not all memories contribute equally to identity, nor is autobiographical memory the most important part of the self. Indeed, some types of memory are less important even than perceptual faculties.

### 3. Study 3: The soul switch

Folk notions of personal identity take a variety of forms. One of the most noteworthy of these is the soul, an immaterial entity that represents an individual's essence (Richert & Harris, 2006; Ward, n.d.). Many religions propose that the soul persists across physical transformations and travels independently of the body, for example during reincarnation or spirit possession (Boyer, 2001). Although the soul has spiritual functions, it also has mental content: Tibetans presume psychological continuity when judging the identity of reincarnated religious leaders (White, Sousa, & Berniunas, in press) and ghosts are thought to retain the mental states of their living predecessors (Bering, 2002).

Instances where the soul takes residence in a new body are not limited to religious contexts, but are also a mainstay of art and popular culture. Movies (*Freaky Friday*, *All of Me*, *Heaven Can Wait*, *Child's Play*) and literature (*Vice Versa*, *The Lovely Bones*, *Prelude to a Kiss*) frequently trade on the implicit beliefs we hold about the soul and the properties it retains when it departs from its too, too solid flesh. One need not believe in souls, or that it's possible for one soul to inhabit the body of another, in order to engage with these stories.<sup>2</sup>

The present study allows us to address whether the essential moral self hypothesis extends to intuitions about the soul, thus broadening the scope of our investigation. We predicted that, when the soul leaves the body, people would judge it to conduct moral traits more reliably than other mental traits, and that this pattern of results would obtain regardless of an explicit belief in God, an immaterial soul, or spiritual possession.

#### 3.1. Methods

318 Americans (*Mdn* age = 33, 61% female) were paid to participate in an online study.

Subjects were asked to imagine that it's possible for a soul to leave one body and inhabit another one, and that this had just happened to a man named Jim, whose soul now inhabited a new body with none of Jim's original traits. Participants reported the extent to which they thought a variety of traits (66 in all) would be carried over into the new body, on a Likert scale from 1 ("Strongly Disagree") to 7 ("Strongly Agree"). The statements took the form of a hypothetical, e.g. "If John was a pedophile in his original body, he will be a pedophile in the new body." Building upon the items presented in Study 2, the traits spanned a variety of psychological categories, as well as somatic traits, so that the two could be directly compared. The somatic traits were designed to be of similar

'distinctiveness' to the mental traits (such as fingerprints or medical conditions), to maximize the chance of their being associated with personal identity. At the end of the study we collected demographics information and individual differences, including belief in souls, spiritual possession, and dualism, to see if endorsement of the premise of the story made any difference in responding.

Item placement in categories for analysis was determined by blind coders, using the same methods outlined in Study 2. The categories were similar to those appearing in Study 2, with the addition of a "somatic" category. One item was rated as both moral and somatic (virginity), so it was excluded from the final analysis. For the selected items and their categories, see Table 2.

#### 3.2. Results and discussion

Paired *t*-tests were performed for all pairwise comparisons between the five categories, with Holm corrections applied to the *p*-values to account for multiple comparisons. Consistent with predictions, participants reported that moral traits are the most likely to move when a soul traveled to a new body (see Fig. 3). Moral traits ( $M = 4.73$ ,  $SD = 1.71$ ) were more associated with the soul than somatic traits ( $M = 2.19$ ,  $SD = 1.34$ ,  $t(317) = 28.2$ ,  $p < .0001$ ), perception ( $M = 2.36$ ,  $SD = 1.32$ ,  $t(317) = 25.4$ ,  $p < .0001$ ), preferences and desires ( $M = 4.14$ ,  $SD = 1.73$ ,  $t(317) = 11.7$ ,  $p < .0001$ ), or memory ( $M = 4.21$ ,  $SD = 1.87$ ,  $t(317) = 8.1$ ,  $p < .0001$ ). Consistent with the previous two studies, the items in the memory and desire categories were not significantly different from one another ( $p = .26$ ). Desires ( $t(317) = 22.0$ ,  $p < .0001$ ) and memories ( $t(317) = 18.9$ ,  $p < .0001$ ) were both significantly different from perceptual traits.

Intriguingly, somatic traits were thought to have a non-zero chance of being imported to the new body. This echoes a finding from Study 1, that physical traits are associated with identity to some extent. Nonetheless, even perceptual traits were more associated with the soul than purely physical features ( $t(317) = 5.4$ ,  $p < .0001$ ). The highest-rated somatic traits (insomnia, hunger) are related to phenomenal experiences, which may explain their higher ranking on the list.<sup>3</sup> Distinctive physical traits are not especially associated with the soul—indeed, of all the traits, fingerprints were deemed *least* likely to transfer to the new body—suggesting that identifying features are not necessarily identity-giving.

Belief in dualism does not lead to an overall tendency to place traits in the new body more readily (Pearson's  $r = .01$ ,  $p = .81$ ), nor does belief in a soul ( $r = -.08$ ,  $p = .14$ ) or belief that souls can possess other bodies ( $r = .06$ ,  $p = .30$ ), or religiosity ( $r = -.09$ ,  $p = .15$ ). Furthermore, the relative ranking in importance of the five categories described in Fig. 3 is no different for those who believe in dualism, souls, and soul possession, and those who do not. This is consistent with previous research showing that declared religious beliefs

<sup>2</sup> Philosophers have frequently tapped these intuitions in thought experiments about identity. II.xxviii.xv Locke (1690/2009) illustrates his argument in favor of psychological continuity by using a tale of a prince's soul inhabiting the body of a cobbler: "every one sees he would be the same person with the prince, accountable only for the prince's actions." Contemporary work has also propelled its arguments on identity with examples based on body-swapping (Shoemaker, 1963; Parfit, 1984; Nichols & Bruno, 2010).

<sup>3</sup> Bear in mind that category membership was determined by an independent set of raters given a pre-determined set of categories to choose from. These two traits would perhaps more accurately be classed as 'psychobiological processes' (viz. Bering, 2002).

**Table 2**

Individual trait scores for the four psychological categories in Study 3. Items were presented in the form, “If John ITEM in his original body, he will ITEM in the new body.” Score represents the extent to which people agreed with this statement; higher score indicates higher agreement.

CATEGORY	ITEM	MEAN (SD)
MORALITY	Was honest	5.40 (1.38)
	Was evil	5.28 (1.52)
	Was conscientious	5.14 (1.51)
	Was a coward	4.86 (1.51)
	Was a jerk	4.83 (1.67)
	Thought having sex was wrong	4.79 (1.69)
	Was a womanizer	4.70 (1.64)
	Was a pedophile	4.58 (1.70)
	Was a thief	4.58 (1.66)
	Was a kleptomaniac	4.53 (1.71)
MEMORIES	Had traumatic memories	4.50 (1.83)
	His memories <sup>a</sup>	4.26 (1.92)
	Knew how to ride a bike	3.86 (1.82)
DESIRES & PREFERENCES	Wanted to eat healthy	4.71 (1.64)
	Liked rap music	4.67 (1.75)
	Wanted to be a doctor	4.51 (1.77)
	Was gay	4.46 (1.67)
	Liked the taste of coffee	3.81 (1.62)
	Hated broccoli	3.77 (1.70)
	Didn't enjoy sex	3.72 (1.55)
	Craved cigarettes	3.49 (1.64)
PERCEPTUAL	Had a good sense of smell	2.64 (1.42)
	Was color blind	2.24 (1.28)
	Was farsighted	2.21 (1.23)
SOMATIC	Had insomnia	3.60 (1.59)
	Was a virgin <sup>b</sup>	3.29 (1.85)
	Was hungry	3.18 (1.72)
	Had high blood pressure	2.19 (1.18)
	Had trouble digesting onions	2.16 (1.17)
	Had high cholesterol	2.12 (1.21)
	Was prone to sunburns	2.05 (1.12)
	Had a large brain	2.04 (1.17)
	Had diabetes	2.03 (1.15)
	Had liver disease	1.97 (1.16)
	Had a birthmark	1.94 (1.25)
	Had a brain tumor	1.93 (1.10)
	Had long legs	1.92 (1.14)
Had a long tongue	1.90 (1.12)	
His fingerprints	1.61 (1.06)	

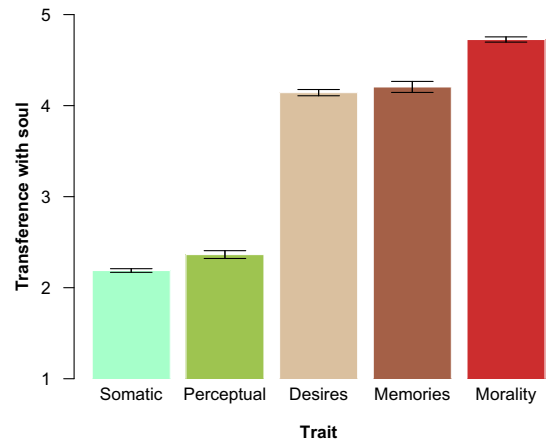
<sup>a</sup> The wording for this question was “John's memories in his original body will be the same in the new body.”

<sup>b</sup> This item was rated as both moral and somatic, so it was excluded from the final analysis.

do not always manifest at an implicit level. For example, people can maintain an explicit belief that God is omnipresent, yet tacitly treat Him as an agent limited by space and time (Barrett, 1999; Boyer, 2001). Similarly, Bering (2002) reports that people attribute to ghosts certain traits (emotions, desires, and beliefs) more than others (perception, psychobiological processes), a pattern that obtains regardless of whether one believes that personal consciousness ceases at death.

#### 4. Study 4: Reincarnation

Bloom and Gelman (2008) describe a case where the 14th incarnation of the Dalai Lama was selected on the basis of his ability to pick out the true artifact owned by the Dalai Lama during his previous life. This tradition may



**Fig. 3.** Association of trait types with the soul, with standard error bars. Higher score indicates greater agreement that the trait will move with the soul to a new body. All pairwise comparisons are significant ( $p < .0001$ ), except Desires vs. Memories.

stem from beliefs about psychological continuity: only the true Dalai Lama would remember interacting with the objects in question (White et al., in press).

As with soul switches, reincarnation is an area where intuitions about numerical identity are drawn upon. In the present study, we use the premise of reincarnation as a springboard for looking at the relative connectedness to the self of two psychological traits: moral character and personality.

Theories of identity and the self that rely on distinctiveness (e.g. Vignoles et al., 2000) would predict that any distinctive dispositional trait should strongly contribute to identity, regardless of its relationship to morality. In Study 2, we found that four personality traits (shy, anxious, industrious, and absent-minded) were not considered as determinative of identity as moral traits. However, these items were not tightly controlled with the moral items, and not as representative in their breadth. The present study focuses on moral and personality traits directly and systematically, so that stronger inferences about their relative contribution to the self can be made.

Some dispositional traits are highly related to moral behavior, such as conscientiousness, and others, such as extroversion, are not (Cawley, Martin, & Johnson, 2000; Strohminger & Gonzalez, 2012). In this study, we pitted highly moralized character traits (such as honesty and loyalty) against those with low moralization (such as creativity and sense of humor). This design also allows us to compare moral traits to general intelligence, which some have asserted is essential to personal identity (Gelman et al., 2007; though see Heyman & Dweck, 1998).

The present study therefore has two chief aims: to see how implicit beliefs about personal identity are accessed for intuitions about reincarnation, and to systematically test whether morality traits are considered more essential to the self than personality traits.

#### 4.1. Methods

118 Americans (*Mdn* age = 26, 33% female) were paid to participate in an online study.

Participants were told that many religious traditions hold that humans can be reincarnated after death into a new human body; they were asked to imagine, for the purpose of the study, that someone has been reincarnated. During this reincarnation process, their “true self”—who they really are, deep down—had been preserved. Some of their more superficial traits may have changed, but their real identity has not. Participants were given pairs of items and asked to choose which of the two items is more likely to have been preserved during the reincarnation.

The pairs consisted of 10 character traits that were strongly related to morality, and 10 that were not strongly related to morality (Strohminger & Gonzalez, 2012), resulting in ten total trials. All subjects saw all items; pairing was counterbalanced between subjects. The items are listed in Table 3.

At the end of the study we collected demographics information, including religiosity (on a sliding scale from 0 to 100), belief in reincarnation (Yes or No), and belief in souls (on a scale of 1–7).

#### 4.2. Results and discussion

A binary logistic regression binning character trait type (moral or personality) into two categorical factors, which included subject and item as random factors (Stiratelli, Laird, & Ware, 1984), revealed that subjects were significantly more likely to choose moral items as surviving after reincarnation ( $t = 55.68$ ,  $p = .0001$ ). Indeed, every single item in the moral category was more likely to be selected than any of the personality items, including traits that have broad impact on overall social and intellectual competence, such as intelligence and extroversion (see Table 3).

**Table 3**

Frequency with which moral and personality traits were selected as being more likely to persist after reincarnation (Study 4).

TRAIT TYPE	TRAIT	% TRIALS SELECTED
MORALITY	Honest	89.0
	Trustworthy	83.1
	Loyal	79.7
	Compassionate	78.0
	Generous	77.1
	Brave	68.6
	Forgiving	66.9
	Fair	66.1
	Grateful	59.3
	Wholesome	58.5
PERSONALITY	Creative	46.6
	Smart	38.1
	Ambitious	33.1
	Relaxed	28.8
	Outgoing	28.0
	Cautious	28.0
	Funny	22.9
	Energetic	22.9
	Assertive	19.5
	Messy	5.9

For the results of all one hundred pairwise comparisons, see Appendix B.

Belief in reincarnation did not moderate either the strength or the direction of these effects when added to the logistic regression model ( $t = .93$ ,  $p = .31$ ), nor did belief in the soul ( $t = .17$ ,  $p = .86$ ) or overall level of religiosity ( $t = .15$ ,  $p = .89$ ).

Social desirability cannot fully account for these effects. For instance, intelligence and creativity are considered more desirable than compassion and wholesomeness (Strohminger & Gonzalez, 2012) yet the latter pair were selected more often. This study demonstrates that moral traits are seen as more constitutive of the true self than personality traits. This is a robust effect: across twenty dispositional traits, moral traits were, without exception, judged to be more likely to survive the radical transformation of existence after death.

#### 5. Study 5: Golden years

The previous four studies presented subjects with hypothetical scenarios that were speculative or unusual in some way. It would also be important to demonstrate that the moral self hypotheses holds true in more everyday settings. Additionally, two of the studies presented here (3 and 4) tap supernatural intuitions, which may be more associated with religious, and by extension moral, traits. For the final experiment, we selected a circumstance—the cognitive and behavioral changes that accompany aging—which is safely within the bounds of normal experience.

Many theories of identity have suggested that uniquely identifying traits are what drives perceptions of the self (Vignoles et al., 2000; Blanton & Christie, 2003). This is unlikely to be the entire story, as Studies 1–4 demonstrate that highly distinctive traits, like traumatic memories or hating broccoli, do not affect identity as much as near-universal human traits, like the capacity for empathy. However, distinctiveness may also be judged in the change itself: memory loss may be seen as more commonplace than change to moral character (for example after brain damage or pill popping; Studies 1 and 2). The present study therefore set out to measure the perceived frequency of mental change in addition to identity change. This will let us see whether the distinctiveness of a change is related to its ability to determine identity. It will also allow us to rule out the possibility that the prominence of morality in the previous studies is due to moral change being perceived as more special.

One reason for morality's dominance in determining identity judgments may be that its effects on behavior are broad-ranging. Some aspects of previous studies provide evidence against this possibility; for example, general intelligence is considered less likely to be reincarnated in the next life than any other moral trait (Study 4). To push on this idea a bit harder, the present study introduced a new category of mental items: basic cognition. If mental content is considered important to the self simply because of its pervasive effects across many domains, then general and higher-level cognitive abilities—such as attention,



planning, and learning abilities—may well be perched at the summit.

In sum, the present experiment aims to tie up four loose ends not fully addressed in the previous studies. One, to demonstrate that personality traits continue to be seen as less crucial to identity when placed in a non-supernatural context. Two, to test whether the exceptional nature of a mental change itself can fully explain our pattern of results. Three, to show that fundamental changes to the mind—such as those engendered by basic cognition—are not the driving force behind the prevalence of moral traits in numerical identity perception. And four, to show that our results hold true in an ordinary context: change to the mental landscape with age.

### 5.1. Methods

79 Americans (*Mdn* age = 32, 56% female) participated in an online paid survey.

Participants were asked to imagine they were meeting an old friend they had known when they were 25, and whom they had not seen in 40 years. Subjects were presented with a list of changes the friend had undergone, and instructed to indicate the degree to which each change would impact the friend's identity. The response to each item was given on a scale from 0% ('This change has no impact on his true self') to 100% ('This change completely alters his true self'). Subjects were asked to assume that for each trait listed, the friend was not that way when they had known him last.

There were 56 items in all, which were presented in randomized order between subjects. Items fell within six categories: morality, personality, basic cognition, memory, desires and preferences, and perception. Positive and negative traits were included within category, where relevant. For example, there were positive moral traits (honest, generous) as well as negative moral traits (racist, cruel). There were cognitive improvements (better at multitasking, improved vocabulary) as well as decrements (worse attention span, forgetfulness). See Table 4 for the full list.

After rating each item for its impact on identity, subjects rated each of these age-related cognitive changes for its perceived frequency. Answers were given on a sliding scale from 0% ('No one experiences this change') to 100% ('Everyone experiences this change').

### 5.2. Results

Pairwise comparisons using paired *t*-tests were performed, with the Holm adjustment to account for multiple comparisons. Moral trait change ( $M = 62.30$ ,  $SD = 20.44$ ) led to significantly more identity change than all other categories of mental change: Perception ( $M = 29.73$ ,  $SD = 20.63$ ,  $t = 10.26$ ,  $p < .0001$ ), Preferences ( $M = 31.12$ ,  $SD = 20.95$ ,  $t = 10.46$ ,  $p < .0001$ ), Memories ( $M = 36.08$ ,  $SD = 19.70$ ,  $t = 8.82$ ,  $p < .0001$ ), Basic cognition ( $M = 38.08$ ,  $SD = 21.20$ ,  $t = 8.79$ ,  $p < .0001$ ), and Personality ( $M = 45.26$ ,  $SD = 21.26$ ,  $t = 9.18$ ,  $p < .0001$ ).

**Table 4**

Individual trait scores for the six psychological categories in Study 5. Score represents the extent to which people said the item contributed to the true self.

CATEGORY	ITEM	MEAN ( <i>SD</i> )
MORALITY	Racist	74.1 (28.7)
	Cruel	70.6 (29.9)
	Rude	65.7 (26.6)
	Spiritual	64.0 (30.2)
	Selfish	63.1 (26.5)
	Hypocritical	62.7 (27.8)
	Forgiving	61.9 (30.5)
	Honest	61.0 (33.0)
	Generous	61.0 (30.6)
	More likely to steal	59.9 (32.8)
	Adulterous	58.2 (31.6)
	Tenderhearted	58.0 (31.0)
	Empathetic	57.7 (31.9)
	Humble	54.3 (31.5)
PERSONALITY	Adventurous	50.2 (30.4)
	Laid-back	50.1 (31.9)
	Ambitious	48.8 (30.9)
	Shy	48.0 (30.9)
	Playful	47.4 (28.8)
	Incurious	43.9 (31.3)
	Likes routine	43.6 (32.8)
	Artistic	42.5 (28.9)
	Detail oriented	40.9 (28.4)
A slow learner	37.0 (27.9)	
BASIC COGNITION	Worse attention span	44.6 (28.1)
	Better at planning ahead	39.2 (26.6)
	Worse at picking up new skills	37.6 (31.5)
	Worse at identifying objects quickly	36.9 (30.2)
	Better at multitasking	32.2 (26.6)
MEMORY	His childhood memories have faded	48.5 (29.3)
	Forgetful	47.1 (30.0)
	Bad at remembering new information	41.3 (30.1)
	Remembers experiences more vividly	38.1 (29.1)
	Knows the meaning of more words	34.3 (28.6)
	Knows how to speak French	32.1 (27.1)
	Forgets how to do calculus	25.1 (29.9)
	Forgets how to ride a bike	22.2 (26.8)
DESIRES & PREFERENCES	Reduced libido	38.4 (32.6)
	Hates throwing things out	37.1 (28.0)
	Likes classical music	35.6 (30.4)
	Likes cooking	33.5 (28.4)
	Hates exercise	32.7 (27.9)
	Hates fixing things	31.5 (28.4)
	Likes shopping sales	30.9 (29.7)
	Likes contemporary art	30.9 (26.8)
	Likes riding motorcycles	29.7 (28.9)
	Hates sports	29.7 (30.5)
	Likes talk radio	29.3 (29.3)
	Likes vegetables	26.9 (27.6)
	Hates gardening	18.5 (24.1)
PERCEPTION	No sensation in arms or legs	43.5 (32.1)
	Moderate hearing loss	36.5 (30.9)
	Needs prescription eyeglasses	30.5 (34.6)
	Can taste food better	24.8 (25.6)
	Colorblind	24.2 (28.1)
	Sharpened night vision	19.0 (20.3)

While personality traits were not as important as any moral traits under study, personality was the second most important contributor to identity, beating out basic cognition ( $t = 3.68, p < .002$ ), memories ( $t = 3.96, p = .0008$ ), preferences ( $t = 6.86, p < .0001$ ), and perception ( $t = 5.94, p < .0001$ ).

Basic cognition, such as attention span and learning aptitude, was found to be more important than preferences ( $t = 4.99, p < .0001$ ) and perception ( $t = 5.14, p < .0001$ ). There was no difference between basic cognition and memories ( $p = .29$ ).

Memories were linked more with identity than preferences ( $t = 3.36, p = .004$ ) or perception ( $t = 4.32, p = .0002$ ). Perception and preferences were not reliably different in their contribution to identity ( $p = .41$ ). For the overall pattern of relationships between group (mental category) means, see Fig. 4.

As with Study 4, every moral item scored higher than every other item, including the personality traits (see Table 4 for individual item scores). Note, too, the importance of pain perception compared with other sensory modalities, and the importance of sexualized desires compared with other desires, patterns appearing in Study 2.

Participant age ranged from 18 to 79 years, with one-third of participants over the age of 53. Since older participants are likelier to have witnessed the kinds of changes described in this study, it's possible that age could affect the overall pattern of responding. A linear mixed-effects model was created, with mental change type and age as independent variables, and item and subject as random factors. Participant age contributes no predictive value to this model: there was neither a main effect of age on identity change judgments, nor an interaction between age and mental category ( $t = .48, p = .63$ ; all interactions NS).

A similar model was constructed to examine how frequency of mental change affects identity judgments, replacing age with frequency ratings in the mixed-effects model. Contrary to the distinctiveness hypothesis, the rarer

the change was, the *less* it was seen as relating to the self ( $t = 2.51, p = .01$ ). This result is reminiscent of the finding that essentialized personality traits are considered more universal (Gelman et al., 2007).

## 6. Discussion

The studies described herein illustrate several points about lay theories of personal identity. The first, most basic, point is that not all parts of the mind are equally constitutive of the self, challenging a straightforward view of psychological continuity. Identity does not simply depend on the magnitude of retained mental content; indeed, certain cognitive processes contribute less to identity than purely physical traits.

Across five experiments, we find strong and unequivocal support for the essential moral self hypothesis. Moral traits are considered more important to personal identity than any other part of the mind. This holds true whether people are considering the fate of someone who suffers brain trauma (Study 1), takes a psychoactive drug (Study 2), moves from one body to another (Study 3), is reincarnated after death (Study 4), or undergoes age-related cognitive changes (Study 5). While morality is consistently revealed as the most important part of the self, modest importance is bestowed upon personality, memories, and desires. To a lesser extent, people also consider low-level cognition and perceptual faculties to be part of the self. This relationship was obtained for both positive and negative traits.

While memory has featured prominently in discussions of personal identity (Locke, 1690/2009; Shoemaker, 1959; Hood et al., 2012) this work has rarely made a distinction among either types or contents of memory. The research presented here suggests that memories are important to the extent that they have resonance with personal relationships. While episodic memory appears to be the most important memory process, this is largely driven by

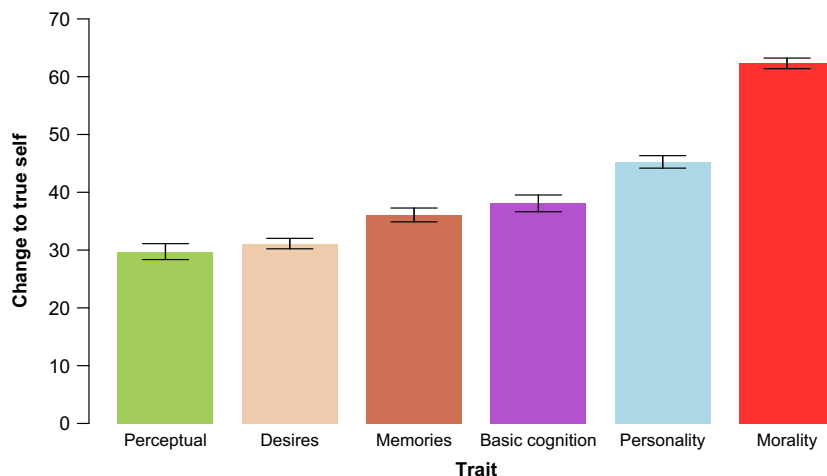


Fig. 4. Average within-category response for how much a trait change in an aging friend would alter his identity, with standard error bars. Higher score indicates greater judged change to his identity. All pairwise comparisons are significant except between basic cognition and memory, and perception and desires.

memories relevant to a person's place in the social world, as evidenced by the complete devaluation of autobiographical memories that do not contain these properties (Study 2). The moderate importance of semantic memory may be due to its perceived relationship to intelligence, which previous research has found to relate to the self concept (Haslam et al., 2004; Gelman et al., 2007). These stratifications within memory lend insight into the contours of how personal identity is constructed. It is not merely the unique, haphazard collection of memories a person acquires across a lifetime, but a particular strain of remembrance of times past. This strengthens the overall view that people are keying into a self that is, at its core, about social relationships, both past and possible.

Previous work on group identity has emphasized the importance of distinctiveness (Nelson & Miller, 1995; Vignoles et al., 2000; Blanton & Christie, 2003). Our results show that the extent to which a trait allows us to be picked out from a crowd is not a critical criterion for determining numerical identity. Distinctiveness alone cannot explain our findings, otherwise memories and preferences would have topped the list, and standard cognitive equipment, like empathy and possession of a moral compass, would have fallen near the bottom. Furthermore, the uniqueness of a cognitive change makes it less likely to be associated with the self (Study 5).

People regard morality as central to identity. Why might morality occupy such a place of privilege? One possibility is that our moral selves are central to what it means to be human—even *C. elegans* has memory. Another, perhaps more substantial, factor is that moral traits are a reliable predictor for how individuals will fare as potential partners for cooperation and affiliation. One of the chief reasons we make distinctions among persons in the first place is to monitor suitable social partners (Baumeister, 1998), and indeed, a person's moral character—as compared with, say, their personality or shared interests—is the ultimate dimension by which we judge friends, business associates, and mates (Buss, 1989; Nowak, Page, & Sigmund, 2000; Goodwin et al., 2014).<sup>4</sup> It would make sense, then, for moral traits to have the most substantial impact on identity persistence.

Perhaps the vaunted status that memories have enjoyed in previous philosophical discussions can be explained by a conflation between the first-person experience of the self and assessment of another person's identity, which may operate by different rules. If you lose all your episodic memories tomorrow, you may feel unmoored in a way that is more salient to you than to others. If your friend loses all his episodic memories, his identity may have changed less in relation to the factors you consider important to your friendship and love for him. A similar rule may be applied to desires and preferences: it's possible these distinguishing features define us more to ourselves than to others.

<sup>4</sup> The ability to keep track of individuals over repeated interactions is, perhaps not coincidentally, one of the necessary preconditions for the emergence of a moral faculty (Trivers, 1971; Nowak, 2006). Relatedly, philosophers have long pointed out that the building blocks of morality—rights, duty, responsibility—would be impossible without the ability to ascribe stable identity to persons (Reid, 1785/1850).

Future work ought to examine how identity from the first and third person perspective differ on this point, if at all.

The soul represents a manifestation of a person's essence (Gelman, 2003). But why does the soul so often take on a moral significance? Children queried about the soul appeal not only to identity and vitalism, but moral-type capacities such as communion with God and knowing the difference between right and wrong (Richert & Harris, 2006). These findings are not limited to Westerners; eastern traditions such as Hinduism conceive of an inner essence that is, at its core, moral (Mehta, 1997). (The fact that religion adjudicates the contours of the soul is itself revealing.) Our results suggest that this relationship may not be coincidental, but part of a natural tendency to associate moral traits with the most essential parts of the self.

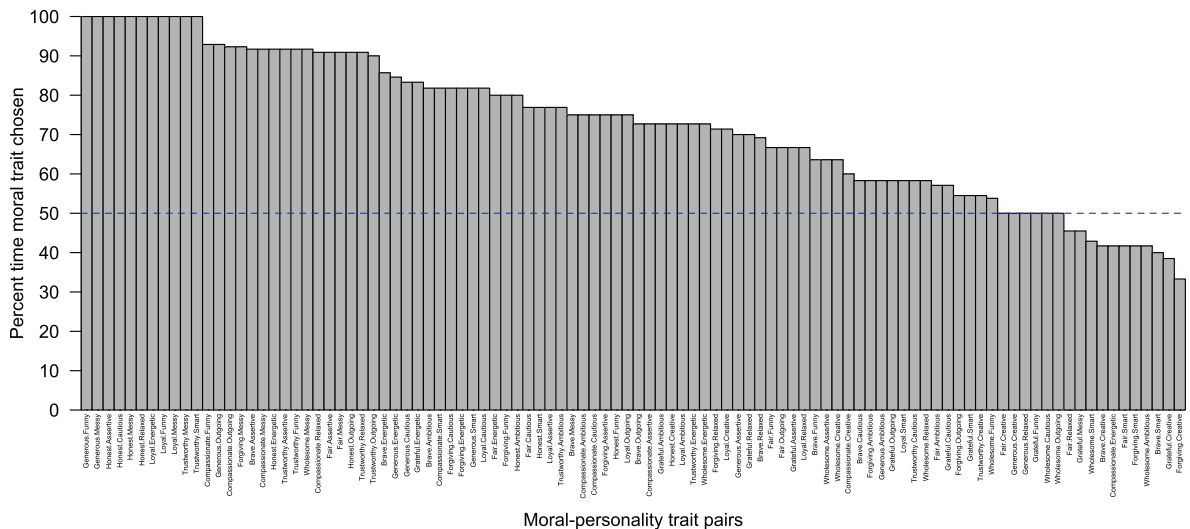
A subsidiary interest of this project was to determine whether one need believe in souls, spiritual possession, reincarnation, or dualism in order to engage with hypothetical scenarios featuring these conditions. Our results show that subjects who do not endorse these ideas nonetheless attribute the same sorts of properties to them. Participants appear to be drawing on a common wellspring of intuitions unrelated to their explicit belief system in order to engage with the task. To make sense of fictional worlds, we do not suspend disbelief indiscriminately; rather, we bypass explicit disbelief in order to access implicit belief structures (viz. Kelly & Keil, 1985, Ward, 1994, Weisberg & Goodstein, 2009, Liao, Strohminger, & Sripada, in press).

These findings also bear on less exotic cases. Every year, millions are diagnosed with Pick's disease, Alzheimer's, and other neurological disorders. Phineas Gages are rare, and Freaky Fridays are fictional, but age-related dementia is ubiquitous. Increasingly common are medications that affect cognitive functioning and behavior: antidepressants, antipsychotics, 'smart drugs', and the like. Disruptions of identity due to medical syndromes and their treatments are likely pervasive; our work suggests that they will be particularly dramatic for outcomes that affect moral faculties.

Whether we see a person as persisting across mental change—even radical, permanent change—depends on the capacity that has been affected. Had Phineas Gage's injury merely eliminated his memory for how to lay down railroad ties, it seems likely his friends would have seen the same old Gage shining through beneath his impairment. Sacks (1985) eventually concludes that his amnesic patient was not a "spiritual casualty" after all, upon witnessing his transcendence while taking the sacrament at Mass. The self is not so much the sum of cognitive faculties as it is an expression of moral sensibility; remove its foothold on that world, and watch the person disappear with it.

## Acknowledgements

Thanks to Susan Gelman, Jesse Graham, Josh Knobe, Meredith Meyer, Nicholas Buttrick, Scott Clifford, Victor Kumar, Wesley Buckwalter, and Stephen Stich for valuable discussions. For help with coding and other research assistance, we thank Garrett Marks-Wilt, Mohini Bhargava, Jeffrey Chang, Elizabeth Cushing, Deepti Joshi, Kara Post, Jingxian Zhang, Irene Lee, and Alexandria Seekely.



**Fig. B1.** The result of all pairwise competitions between moral and personality traits in Study 4. Dotted line indicates chance.

## Appendix A. Materials used in Study 1

Jim is an accountant living in Chicago. One day, he sustains a severe head injury from a car accident. His only chance for survival is participation in an advanced medical experiment called a Type 2 transplant procedure. It is the year 2049 and scientists are able to grow different parts of the brain if they become damaged. A stock of brain tissue is kept cryogenically frozen to be used as spare parts in the event of an emergency. In a Type 2 transplant procedure, a team of doctors removes the damaged parts of the brain and carefully replaces it with the stock brain tissue. The damaged brain tissue is destroyed after it has been removed. After the operation, all the right neural connections between the old brain and the replacement brain tissue have been made. The doctors test all physiological responses and determine that the patient is alive and functioning. The doctors scan the brain of the transplant recipient and run some standard psychological tests. [VARIABLE] Indicate the extent to which you agree with the following statement: *After the surgery, the transplant recipient is still Jim.*

Text for VARIABLE:

CONTROL: “They discover that the transplant recipient thinks and acts the same way as before the accident.”

AGNOSIA: “They discover that the transplant recipient has lost his ability to recognize objects—he can see perfectly fine, but his ability to identify objects has disappeared. Aside from this, he thinks and acts the same way as before the accident.”

APATHY: “They discover that the transplant recipient has lost all his desires—he no longer wants or desires anything. Aside from this, he thinks and acts the same way as before the accident.”

AMNESIA: “They discover that the transplant recipient has lost his memories—he can no longer remember anything that happened before the accident. Aside from

this, he thinks and acts the same way as before the accident.”

MORALITY: “They discover that the transplant recipient has lost his moral conscience—he is no longer capable of judging right from wrong, or being moved by the suffering of others. Aside from this, he thinks and acts the same way as before the accident.”

## Appendix B

See Fig. B1.

## References

- Allport, G.W. (1937). *Personality: A psychological interpretation*. Oxford, England: Holt.
- Ayer, A. J. (1936). *Language, truth, and logic*. London: Gollancz.
- Barrett, J. (1999). Theological correctness: Cognitive constraint and the study of religion. *Method and Theory in the Study of Religion*, 11(4), 325–339.
- Bartels, D. M., Kvaran, T., & Nichols, S. (2013). Selfless giving. *Cognition*, 129(2), 392–403.
- Bartels, D.M., & Urminsky, O. (2011). On intertemporal selfishness: How the perceived instability of identity underlies impatient consumption. *Journal of Consumer Research*, 38(1), 182–198.
- Baumeister, R. (1998). The self. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680–740). New York, NY: McGraw-Hill.
- Bering, J. (2002). Intuitive conceptions of dead agents' minds: The natural foundations of afterlife beliefs as phenomenological boundary. *Journal of Cognition and Culture*, 2 (4), 263–308.
- Bering, J. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, 29(5), 453–462.
- Blanton, H., & Christie, C. (2003). Deviance regulation: A theory of action and identity. *Review of General Psychology*, 7(2), 115.
- Blok, S., Newman, G., & Rips, L. (2005). Individuals and their concepts. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin* (pp. 127–149). Washington, DC: American Psychological Association.
- Bloom, P., & Gelman, S. (2008). Psychological essentialism in selecting the 14th Dalai Lama. *Trends in Cognitive Sciences*, 12(7), 243–243.
- Boyer, P. (2001). *Religion explained*. New York, NY: Basic Books.

- Buss, D. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, 12(1), 1–14.
- Butler, J. (1736/2008). Of personal identity. In *The analogy of religion*. New York, NY: Robert Carter and Brothers (Reprinted in Perry, 2008, pp. 99–106).
- Cawley, M.J., Martin, J.E., & Johnson, J.A. (2000). A virtues approach to personality. *Personality and Individual Differences*, 28(5), 997–1013.
- Corriveau, K., Pasquini, E., & Harris, P. (2005). "If it's in your mind, it's in your knowledge": Children's developing anatomy of identity. *Cognitive Development*, 20(2), 321–340.
- Dennett, D. C. (1978). *Brainstorms*. In (pp. 310–323). Cambridge, MA: Bradford Books.
- Gelman, S. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.
- Gelman, S., Heyman, G., & Legare, C. (2007). Developmental changes in the coherence of essentialist beliefs about psychological characteristics. *Child Development*, 78(3), 757–774.
- Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, 106(1), 148–168.
- Gottfried, G., Gelman, S., & Schultz, J. (1999). Children's understanding of the brain: From early essentialism to biological theory. *Cognitive Development*, 14(1), 147–174.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23(2), 101–124.
- Haidt, J., & Hersh, M. (2001). Sexual morality: The cultures and emotions of conservatives and liberals. *Journal of Applied Social Psychology*, 31(1), 191–221.
- Haslam, N., Bastian, B., & Bissett, M. (2004). Essentialist beliefs about personality and their implications. *Personality and Social Psychology Bulletin*, 30(12), 1661.
- Heyman, G.D., & Dweck, C.S. (1998). Children's thinking about traits: Implications for judgments of the self and others. *Child Development*, 69(2), 391–403.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2), 65–70.
- Hood, B., Gjerroe, N., & Bloom, P. (2012). Do children think that duplicating the body also duplicates the mind? *Cognition*, 125(5), 3466–3474.
- Hume, D. (2000). *An enquiry concerning human understanding: A critical edition* (Vol. 3). USA: Oxford University Press.
- Johnson, C. (1990). If you had my brain, where would I be? Children's understanding of the brain and identity. *Child Development*, 61(4), 962–972.
- Kelly, M. H., & Keil, F. C. (1985). The more things change...: Metamorphoses and conceptual structure. *Cognitive Science*, 9(4), 403–416.
- Knobe, J. (2005). Ordinary ethical reasoning and the ideal of 'being yourself'. *Philosophical Psychology*, 18(3), 327–340.
- Kramer, P. (1993). *Listening to prozac*. Viking: New York, NY.
- Liao, S., Strohminger, N., & Sripada, C. (in press). Empirically investigating imaginative resistance. *British Journal of Aesthetics*.
- Libby, L., & Eibach, R. (2007). How the self affects and reflects the content and subjective experience of autobiographical memory. In C. Sedikides & S. Spencer (Eds.), *The self*. New York, NY: Psychology Press.
- Locke, J. (1690/2009). *An essay concerning human understanding*. New York, NY: WLC Books.
- Macmillan, M. (2000). Restoring phineas gage: A 150th retrospective. *Journal of the History of the Neurosciences*, 9146–9166.
- McGuire, W.J., & Padawer-Singer, A. (1976). Trait salience in the spontaneous self-concept. *Journal of Personality and Social Psychology*, 33(6), 743.
- Mehta, K. K. (1997). The impact of religious beliefs and practices on aging: A cross-cultural comparison. *Journal of Aging Studies*, 11(2), 101–114.
- Nelson, L.J., & Miller, D.T. (1995). The distinctiveness effect in social categorization: You are what makes you unusual. *Psychological Science*, 6(4), 246–249.
- Newman, G., Knobe, J., & Bloom, P. (2014). Value judgments and the true self. *Personality and Social Psychology Bulletin*, 40(2), 203–216.
- Nichols, S., & Bruno, M. (2010). Intuitions about personal identity: An empirical study. *Philosophical Psychology*, 23(3), 293–312.
- Nowak, M. (2006). Five rules for the evolution of cooperation. *Science*, 314(5805), 1560–1563.
- Nowak, M., Page, K., & Sigmund, K. (2000). Fairness versus reason in the ultimatum game. *Science*, 289(5485), 1773–1775.
- Olson, E. (2003). Personal identity. In S. Stich & T. Warfield (Eds.), *The Blackwell guide to the philosophy of mind*. Oxford, UK: Blackwell.
- Parfit, D. (1971). Personal identity. *The Philosophical Review*, 803–827.
- Parfit, D. (1984). *Reasons and persons*. Oxford, England: Oxford University Press.
- Perry, J. (2002). *Identity, personal identity, and the self*. Indianapolis, IN: Hackett.
- Prinz, J. (in press). An empirical case for motivational internalism. In Björnsson, G. (Ed.), *Motivational internalism*. Oxford, UK: Oxford University Press.
- Prinz, J., & Nichols, S. (in press). Diachronic identity and the moral self. In Kiverstein, J. (Ed.), *Handbook of the social mind*. London: Routledge.
- Reid, T. (1785/1850). Essays on the intellectual powers of man. In Walker, J. (Ed.). (Abridged ed.) Cambridge, MA: John Bartlett.
- Richert, R., & Harris, P. (2006). The ghost in my body: Children's developing concept of the soul. *Journal of Cognition and Culture*, 3(4), 409–427.
- Riis, J., Simmons, J.P., & Goodwin, G.P. (2008). Preferences for enhancement pharmaceuticals: The reluctance to enhance fundamental traits. *Journal of Consumer Research*, 35(3), 495–508.
- Sacks, O. (1985). The lost mariner. In *The man who mistook his wife for a hat* (pp. 23–42). New York, NY: Touchstone.
- Shoemaker, S. (1959). Personal identity and memory. *The Journal of Philosophy*, 56(22), 868–882.
- Shoemaker, S. (1963). *Self-knowledge and self-identity*. Ithaca, New York: Cornell University Press.
- Shoemaker, S., & Swinburne, R. (1984). *Personal identity*. Oxford, England: Blackwell.
- Shweder, R., Much, N., Mahapatra, M., & Park, L. (1997). The "big three" of morality (autonomy, community, divinity) and the "big three" explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health: Interdisciplinary perspectives* (pp. 119–169). New York, NY: Routledge.
- Skitka, L.J., Bauman, C.W., & Sargis, E.G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology*, 88(6), 895–917.
- Starkstein, S., Petracca, G., Chemerinski, E., & Kremer, J. (2001). Syndromic validity of apathy in Alzheimers disease. *American Journal of Psychiatry*, 158(6), 872–877.
- Stiratelli, R., Laird, N., & Ware, J.H. (1984). Random-effects models for serial observations with binary response. *Biometrics*, 40(4), 961–971.
- Strohminger, N., & Gonzalez, R. (January 2012). Virtues, vices, and personality: Do people reason the same way about all individual differences? In *Presented at the society for personality and social psychology 13th annual meeting, San Diego, CA*.
- Thomson, J. J. (1997). People and their bodies. In J. Dancy (Ed.), *Reading parfit*. Oxford, UK: Blackwell.
- Trivers, R.L. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, 46(1), 35–57.
- Unger, P. (1990). *Identity, consciousness, and value*. Oxford, UK: Oxford University Press.
- Vignoles, V.L., Chryssochoou, X., & Breakwell, G.M. (2000). The distinctiveness principle: Identity, meaning, and the bounds of cultural relativity. *Personality and Social Psychology Review*, 4(4), 337–354.
- Wallace, D.F. (2004). Consider the lobster. *Gourmet Magazine*, 50–64.
- Ward, T. B. (1994). Structured imagination: The role of category structure in exemplar generation. *Cognitive Psychology*, 2(7), 11–40.
- Ward, A. (n.d.). *What is self like?* Unpublished manuscript.
- Weisberg, D. S., & Goodstein, J. (2009). What belongs in a fictional world? *Journal of Cognition and Culture*, 9(1-2), 1–2.
- White, C., Sousa, P., & Berniunas, R. (in press). Psychological essentialism in selecting the 14th Dalai Lama: An alternative account. *Journal of Cognition and Culture*.
- Williams, B. (1970). The self and the future. *The Philosophical Review*, 79(2), 161–180.
- Williams, B. (1973). Problems of the self. In *Personal identity and individuation*. Cambridge, UK: Cambridge University Press (Originally appeared in Proceedings of the Aristotelian Society 57, chapter).