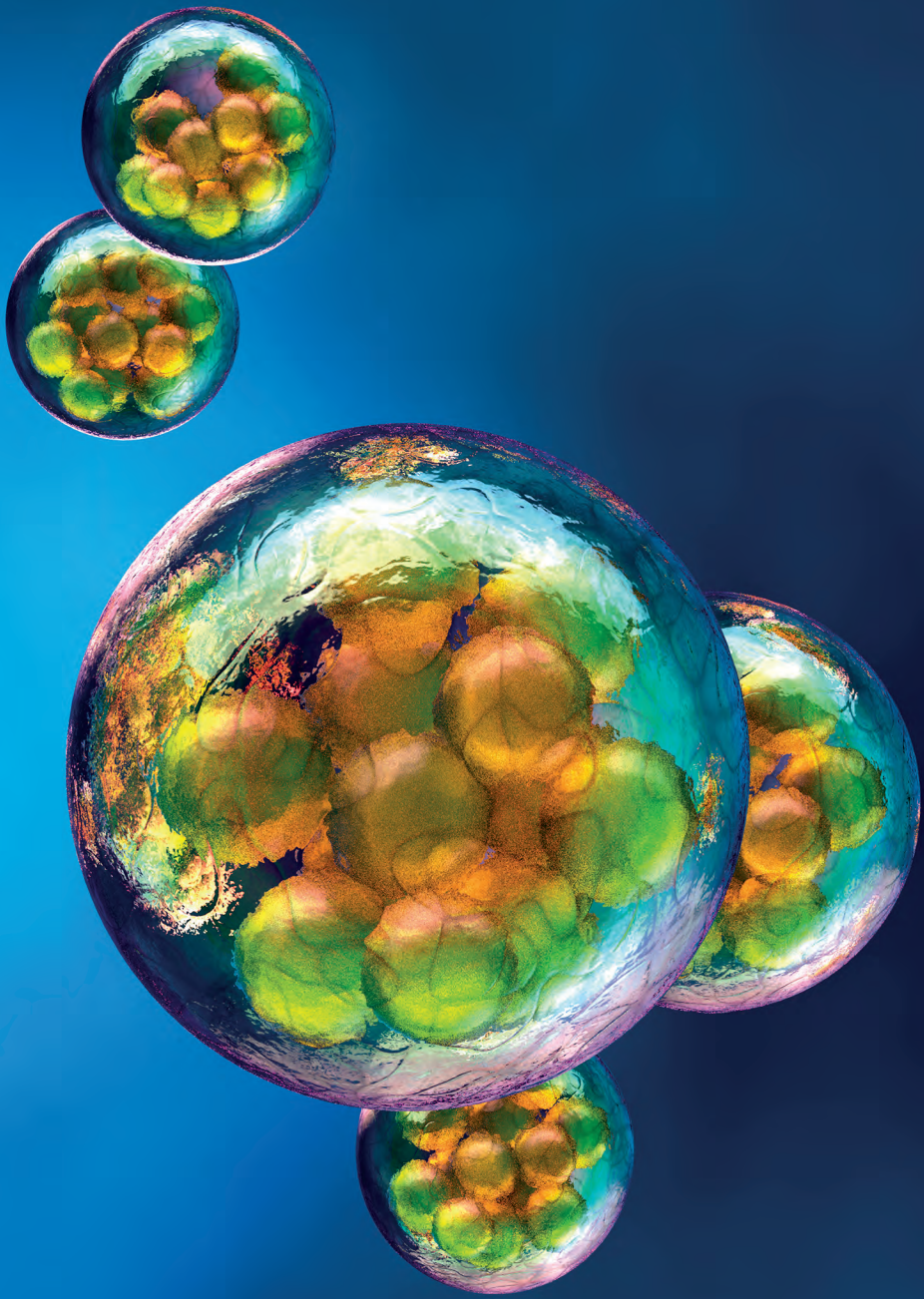
compos mentis

Undergraduate Journal of Cognition and Neuroethics
Volume 6, Issue 2



compos mentis

Student Editor
Cody Hatfield-Myers

Production Editor
Zea Miller

Publication Details
Volume 6, Issue 2 was digitally published in June of 2018 from Flint, Michigan, under ISSN: 2330-0264.

© 2018 Center for Cognition and Neuroethics

The *Compos Mentis: Undergraduate Journal of Cognition and Neuroethics* is produced by the Center for Cognition and Neuroethics. For more on CCN or this journal, please visit cognethic.org.

Center for Cognition and Neuroethics
University of Michigan-Flint
Philosophy Department
544 French Hall
303 East Kearsley Street
Flint, MI 48502-1950

Table of Contents

1	**Culture on Self-Construal: Personhood and Identity** Adina Cianciotto	1–26
2	**Relevance Realization in the Formation of Mental Representation: Analyzing Discrete Thought: Why Cognition Must Use Discrete Representations** Tong (Victoria) Liu	27–37
3	**Who Decides?** Rhea Thukral	39–54
4	**Human Intuition in Folk Mereology** Yufan Yang	55–68

compos mentis

Culture on Self-Construal: Personhood and Identity

Adina Cianciotto
Northwestern University

ABSTRACT
Cross-cultural research suggests that our concept of personal identity is partly constituted by one's culture in a cycle of mutual constitution (Markus and Kitayama 2010). While this concept has been widely referenced in cultural psychology, contemporary work on personal identity continuity has largely ignored this cultural component. I hypothesize that cultural differences between East Asian and Western cultures promote different intuitions about identity continuity in the context of thought experiments on personhood and identity. In this paper, I test a series of scenarios which explore whether culture affects self in personhood, name, and identity ratings. Results of the study (n = 15) showed 13% of the frames showed a significant cultural effect, and contrary to the hypothesis and supportive research, East Asians and Western cultures do not appear to exhibit significantly consistent differences in attribution of personhood and self. This result undermines some widely held philosophical assumptions about the influence of culture on self-construal, suggesting that judgments of personal identity and personhood may be resistant to cultural bias.

KEYWORDS
Identity, Cross-cultural, Personhood, Self, Thought Experiments

INTRODUCTION

For centuries, philosophers have debated personal identity, addressing the question of what makes one's identity, and what is required for it to persist over time. Historically, notions of the essential self (self-as-soul) dominated Western philosophy (Martin and Barresi 2006), however this later switched focus to the psychological view (John Locke 1975), arguing that memory ensures the continuity of a person. Subsequent research by Nichols and Bruno (2010) extrapolated that according to folk judgment, psychological continuity (memories) is necessary for preservation of personal identity. In a response to Nichols and Bruno, Berniunas and Dranseika (2016) presented two challenges to their study. They found that folks do not use a unitary concept of personal identity, but rather division into thinner notions of "person" and "identity of individual" (Blok, Newman, Behr, and Rips 2001), and that psychological continuity is important, but not necessary for personal identity judgments. In investigating what changes have the potential to alter someone's identity and sense of personhood, as illustrated, the focus has been on mainly internal attributes. Oddly enough, these very studies have alluded to another attribute of notable significance, though, that being, cross-cultural significance. Amongst these studies, researchers have consistently noted the emergence/consideration of cross-cultural self-construal. Nichols and Bruno (2010) noted in their paper that "these experiments only focused entirely on Western graduates" and "it is quite possible that people in different culture or socioeconomic groups will respond differently". Berniunas and Dranseika (2016) also noted in a free-listing experiment the emergence of a "social" and "moral" dimension, recognizing research (Markus and Kitayama 1991) that cross-culturally individuals conceptualize the self differently, and that one's moral judgments can actually influence one's intuitions in folk psychology and causal cognition (2016)[1]. This brings up an interesting discussion—research and researchers have suggested the complexity of self, and the important consideration of cultural factors, but yet so far it has been widely ignored in the personal identity debate. This appears to

1. While Berniunas and Dranseika (2016) attempted to explore moral and social dimensions to personal identity through thought experiments (arguably dimensions highly relevant/embedded in culture), their results emerged inconsistent across studies (no significant effects in study 1, sometimes only trends in predicted directions). Furthermore, they only explored this dimension across a single culture, not across two cultures. This leaves opportunity for a cross-cultural design to achieve further insight/leverage into this debate on personal identity.

expose caveats to prior research, and opportunities for future research to explore the implications of culture on self-construal.

Cross-cultural research argues the importance of considering culture in cognition in that it is impossible to understand psychological processes without the consideration for specific cultural background in which these psychological processes are embedded (Shweder 1993). In a cycle of ongoing mutual constitution, cultures and selves are seen to define and build upon each other (Markus and Kitayama 1998), and through this, develop through symbolically mediated, collaborative interaction with others and the social environment. According to Markus and Kitayama (1998), each of these divergent construals have a set of specific consequences for cognition, emotion, and motivation.

Following these concepts of cross-cultural constructions of self, some pioneering research has delved deeper into these exact cultural influences. Research that has been done on personal identity (Kung et al. 2016) suggests that in cultures with high social rigidity, lay beliefs that the world is fixed (vs. malleable) predict identity continuity. This research contributed to the literature on identity continuity by highlighting that people of different cultures may have different intuitions about what features of self are essential for maintaining identity continuity.

Notable cross-cultural distinctions have also been found in independent vs. interdependent concepts of self, in that in interdependent (East Asian) cultures, responding to one's environment requires awareness of the relatively larger role of others in influencing who you are and what you should be doing (Norenzayan, Choi, and Nisbett 2001). While this situational information greatly influenced East Asian individuals, this information had no effect for Americans from Western cultures (Norenzayan, Choi, and Nisbett 2001). Some further ways in which East Asian and Western cultures assess self differently are seen on scales of individualism and collectivism, senses of agency, tightness and looseness of society and norms, and malleability of self vs. world (Heine 2008). As demonstrated through East Asian and Western cultures' highly divergent models of self, they are popular subjects for cross-cultural research.

While there have been studies on comparisons on East Asians and Western cultures on identity continuity (Kung and Eibach 2016), none have yet been done through thought experiments. The thought experiment is one important tool that philosophers have frequently used to investigate personal identity, which is

believed to help to reveal "the criteria of personal identity that we actually use" (Parfit 1971). Influential thinkers have considered thought experiments like this to debate which features of the self are critical for determining self are critical for determining personal identity continuity (Parfit 1971). While prior cross-cultural research provides promising evidence of culture's influence on constructions of self, how culture would affect self in thought experiments remains unknown. This study aims to explore whether these cultural worldviews manifest in thought experiments, and whether the processes involved in evaluating personal identity are alterable by culture. The present study addresses these questions by taking a novel thought experiment approach to cross-cultural selves, and also by utilizing novel methods into identity gleaned by Berniunas and Dranseika (dissociation between personhood and identity). Broadly speaking, such an investigation could ultimately support at least two different conclusions. On the one hand, it could turn out that culture affects personal identity in thought experiments under no circumstances. On the other hand, it can demonstrate culture can play a role in all, some, or under certain conditions in thought experiments. This research is significant in that cultural reflection on judgments in philosophical thought experiments could progress or reinvent research on identity in cognitive science. As detailed in the research overview, this study predicts that cultural differences between East Asian and Western cultures promote different intuitions about identity continuity in regards to thought experiments on ratings of name, personhood, identity.

RESEARCH OVERVIEW

In this study, I aim to replicate and expand/adapt on thought studies done by Berniunas and Dranseika (Nichols and Bruno 2005; Block, Newman, and Rips 2005; Grey, Knickman, Wegner 2011; Berniunas and Dranseika 2016), and Weaver and Turri (Parfit 1984; Weaver and Turri, In Press). To test whether cultures differ in their personal identity intuitions, and see how these results inform about the debate on self-construal, this study compares individuals with representative East Asian or Western background.

In my design, I took 4 different approaches to personal identity. The cases are as follows: Frame #1 Brain Transplant, Frame #2 Half Brain Procedure, Frame #3 Persistent Vegetative State, #4 Persisting as Many. In each case, I adapted the model vignette with the addition of a cultural condition (two conditions: neutral

and cultural). While in Weaver and Turri's initial experiment, participants were asked to choose options to best describe that happens in the story (Derek is in the West Recovery room, etc.), I maintained the same evaluation criteria from Berniunas and Dranseika (2016) for constancy. Across all these studies, I looked at the cultural effect on personhood, identity, and name. In this asking, do East Asian and Western cultures approach intuitions about personhood, identity, and name differently?

Across all vignettes, a neutral condition was included as to serve as a control, and also as means to verify the results of the original study by extending it to a cross-cultural frame. Conditions were broken into cultural conditions in order to get at "latent" intuitions on personhood and self between cultures. As with more discreet concepts, like culture, many notable distinctions are not illuminated without the appropriate scenario, environment, or frame of mind. For instance, one may not realize there is no concept equivalent of "chi" in Western cultures unless in a tai-chi class. This has been supported by Williams (1970) who argues that our intuitions about personal identity vary depending on how a given thought experiment is framed. In this, I propose that in order to get the "true" scope of culture on self-identity intuitions, these latent cultural intuitions must be triggered by particular framing scenarios. Due to the "standardized" nature of the neutral condition, culturally-relevant differences in concepts, values, and perspectives may go ignored. I created the cultural condition (merge of social/moral aspects) to bring to the forefront some culturally-divergent topics/perspectives. These focused on relation of self to others; independent vs. interdependent relationships, malleability of self vs. world, cultural core values, analytic vs. holistic reasoning style (in contradictions), and overall treatment of personal continuity.

Before each condition, I present background research on the topic, my specific hypotheses, the neutral and cultural frame, and the research questions the scenario aims to address. While the background research is varied to each frame, as culture consists of many overlapping aspects, points and research from all frames were considered in testing and design. Results and discussion for all questions will be discussed altogether at the end.

For all conditions and frames, I start with a few general hypotheses: I predict that neutral conditions will not elicit any statistically significant differences between East Asian and Western cultures (across all 3 statements). In contrast, a portion of cultural conditions will. In the circumstance that a neutral condition does produce

a significant result, I anticipate that for whatever effect is statistically significant in the neutral condition, that effect will be greater in the cultural condition. I also predict that overall the standard deviation will be larger for Western conditions, in that answers will tend to fall on either extreme end of the scale (0, or 10) due to preference for rule-based reasoning (Heine 2008). In contrast, I predict that East Asians will have less disparity in their answers of personhood and identity, due to a "thicker" notion of self (Heine 2008). This would suggest a more unitary concept of self.

While I am interested in how culture affects ratings of personal identity between name, identity continuity, and personhood, in analysis, I primarily investigate whether, cross-culturally, there is a similar pattern for ratings across all 3 statements, and am primarily concerned with whether any of these statements (compared on their own) result from significant effects between cultures.

METHODS

Participants

Participants were undergraduate students at Northwestern University. 26 participants were run, but 11 were excluded from analysis on the basis of representativeness (leaving N = 15, 60% female, mean age = 21). The study was administered as an online survey via Qualtrics.

Materials and Procedure

Each vignette was split into neutral and cultural conditions. Using a within-subjects design, participants were exposed to every neutral and cultural condition (4 vignettes x 2 conditions = 8 scenarios total). Following Berniunas and Dranseika's study (2016), participants rated their agreement with whether the subject in each vignette "is still -insert name-" " is still the same person", "and is still a person" on a scale of 0 (strongly disagree) to 10 (strongly agree). Comprehension of scenarios was checked for each and varied to reflect changes in cultural/neutral conditions (2 failed and were excluded from analysis). Scenario subject names were also balanced for cultural familiarity, subject gender was varied, and wording of the likert scales were matched to bodily state and time reference. Questions were asked in the same order and the order of response options were all counterbalanced. I compared participants using both English-language materials for both samples,

eliminating language confounds (Grossman & Na 2014). The story remained at the top of the screen throughout, and participants could not return to a previous screen to change their answers.

Following the survey, participants filled out a brief demographic survey providing the context for group assignment. Participants were put into either East Asian or Western cultural conditions on the basis of race, ethnicity, language, generation, personal identity ratings, time abroad, and familiarity with Eastern/ Western ideals (East Asian condition N = 6, Western condition N= 9) (*see appendix for exact questions/grouping criteria). These same basic procedures were used for all subsequent questions reported here.

FRAME TYPE #1: BRAIN TRANSPLANT CASE

Research by Markus and Kitayama (2010) suggests divergent models in which East Asian and Western cultures reference sense of self. East Asian schemas, otherwise known as interdependent schemas, organize behavior in reference to the thoughts, feelings, and actions of the related person. This interaction with others creates a sense of self connected, or interdependent with others, in other words, viewing social relationships as core to self (Markus and Kitayama 2010). In contrast, Western schema of self organizes behavior in respect to the individual's own thoughts, feelings, and actions. This individualistic approach produces a sense of self existing separately from social environment and all other selves, hence, assuming a weaker role of relationships in self-identity (Markus and Kitayama 2010).

A particularly interesting finding of this research is the idea that in interdependent schemas of self, people attribute a sense of themselves as part of their social relationships (Markus and Kitayama 2010). This is consistent with a sense of one's self being related to others, and has been supported by research that for East Asians, the same region of the brain is activated by both significant other's (ex. mother) and for the self (Zhu et al. 2007). This was also seen in that East Asian individuals weigh the words of their mother as much as if they had decided it themselves (Zhu et al. 2007). This is illustrated in Figure 1, where in interdependent cultures, people view "in-group" individuals (mother, father, friend) as an extension of oneself. In contrast, independent cultures view self as isolated from external relations.

In my first frame, I wanted to explore these diverging models of self, particularly in regards to interpersonal relationships, and the tendency for interdependent cultures to see others as a part of what makes "them". To explore this, I first adapted a vignette from Blok, Newman, and Rips (2005) in which an individual must undergo a brain transplant to a stock body to survive. I used the original frame for my neutral condition (making slight alterations for conditional constancy). I then further adapted the vignette for the cultural condition, altering the stock body used for the transplant to be the individual's father instead. This addressed the question of whether the vessel of the brain transplant (in this case, the father) has different implications cross-culturally on self. Utilizing the interdependent schema (Figure 1), I predicted that due to the overlapping of self and others, East Asian individuals would approach their father's body in a different manner (responding with higher or lower identity ratings than in the neutral condition), whereas Western individuals would respond with constancy between conditions in their attributions of identity. Once again, as research suggests Western cultures view self-identity separately from relationships and one's social environment, the transplant into the father's body should not interfere with one's own concept of self. Additionally, as these referential schemas focus on one's "sense" of self, I do not predict that this manipulation will alter one's ratings for personhood in either culture.

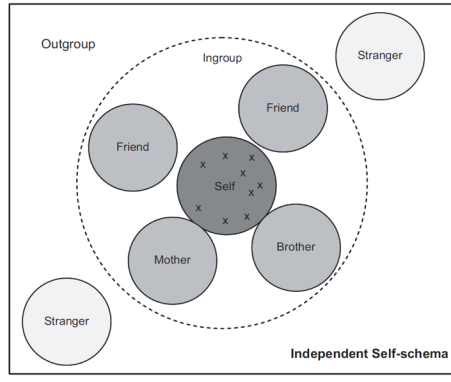


Figure 1a. Independent and interdependent self-schemas. Figure adapted from Markus and Kitayama (1991) and Heine (2008).

In this frame, the neutral condition is as follows:

> David is severely injured in a tragic car accident. His only chance for survival is participation in a "Type 2 transplant" procedure. In a "Type 2 transplant" procedure, David's brain is removed and carefully placed into a stock body. David agrees to the operation. David's original body is destroyed in the operation. After the operation, all the right neural connections between the brain and body have been made. The doctors test all physiological responses and determine that the transplant recipient is alive and functioning. David's brain is successfully transplanted and all his memories from before the operation are intact.

Following the passage, participants rated their agreement with the following statements on a scale of 0 (strongly disagree) to 10 (strongly agree):

1. After the event, the Type 2 transplant recipient is still David
2. After the event, the Type 2 transplant recipient is still the same person as before the event
3. After the event, the Type 2 transplant recipient is still a person
4. After the event, the Type 2 transplant recipient has the same memories as before the procedure [comprehension check]

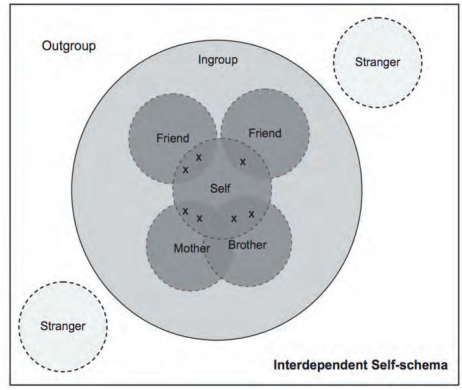


Figure 1b. Independent and interdependent self-schemas. Figure adapted from Markus and Kitayama (1991) and Heine (2008).

For the cultural condition, in which stock body switched to be the individual's father's body, the scenario is as follows:

> David is severely injured in a tragic car accident. His only chance for survival is participation in a "Type 2 transplant" procedure. In a "Type 2 transplant" procedure, David's brain is removed and carefully placed into another body. However, there is no body to transfer David's body into. David's father's body was donated to science when he passed away. David's father's body is used as a stock body. David's brain is successfully transplanted into the body and all his memories from before the operation are intact. David's original body is destroyed in the operation. After the operation, all the right neural connections between the brain and body have been made. The doctors test all physiological responses and determine that the transplant recipient is alive and functioning.

Once again, on a scale of 0 (strongly disagree) to 10 (strongly agree) participants rated their agreement with the following statements:

1. After the event, the Type 2 transplant recipient is still David
2. After the event, the Type 2 transplant recipient is still the same person as before the event
3. After the event, the Type 2 transplant recipient is still a person
4. After the event, the Type 2 transplant recipient has the same physical appearance he did before [comprehension check]

FRAME TYPE #2: HALF BRAIN CASE;
FRAME TYPE #3: PERSISTENT VEGETATIVE STATE

Cross-cultural studies suggest that East Asian and Western cultures have conflicting stances on the malleability of self vs. world—Westerners approaching malleability of world relative to self, whereas East Asians view malleability of self relative to world. This is demonstrated in research by Chiu (1997), which showed how Chinese and Americans possess divergent constructs of society. If presented with the problem of building a stone wall, there are two considerations, the shape of the stones and the design for the structure. Chiu describes one method, which

refers to the shape of each stone altered to accommodate for the design of the wall, and the second method, where the shape of the stone is preserved, and the blueprint is altered to accommodate for the stones. Chiu's research suggests that the context-specific model conforms more closely to Chinese and East Asian societies, in that East Asians hold a stronger belief in the fixedness of the social world (roles, positions, relationships) and the fluidity of personal qualities (Chiu 1997). This references the importance in Confucian values to "fit in" to one's social environment, as it is fixed. Westerners, on the other hand, viewed our core selves as unchanging and constant (Markus and Kitayama 2001), and perceived that individuals shape the world and their situations. Other research supports these findings as well, for instance, Kung (2016) found that cultures with fixed-world beliefs perceive more identity discontinuity than Americans when one's relationships are altered (Kung et al. 2016). In the context of East Asian cultures, as with many other points, this can also be understood as an adherence to Confucian norms, which prioritizes group/intrapersonal harmony of "Li" (Markus and Kitayama 2010). Research by (Lee, Hallahan, and Herzog 1996) also revealed that Western cultures assume people have fixed internal attributes, and one's identity affirms these inner attributes. For instance, Americans tend to judge changes in moral and personality traits (psychopathy, shyness) as most indicative of psychological discontinuity (Strohminger and Nichols 2014). This exaggeration of dispositional information leads to a drive for consistency, where consistent behavior is considered important to psychological well-being, and too many changes in behavior lends to a "loss of self" (Lee, Hallahan, and Herzog 1996).

In my experimental design, I took two separate approaches to these concepts—one catering towards East Asian fixedness and importance of the external world (relationships), and Westerners de-emphasis on the social world/situation—the other catering towards Westerners emphasis on the fixedness of self, and East Asians de-emphasis/fluidity on a consistent self. In design, this emerged as one scenario where there was change in one's external relations, and in the other, one where there is a change in one's internal traits.

To test this, I adopted a vignette from Weaver & Turri (in press) on Half Brain procedures, and a vignette from Grey, Knickman, Wegner (2011) on Persistent Vegetative states for my neutral condition. For my cultural condition, I further adapted the scenarios, in the Half Brain frame, creating it so that following the procedure, the individual (Sara), experiences consequences that result in her

removal from society and close social ties. In the PVS condition, I adapted it so that upon re-consciousness, the individual (Anh) undergoes severe moral and internal character changes.

As an overview, I predict that East Asian cultures will attribute changes in one's external relationships to be more identity altering, whereas Western cultures will attribute internal psychological changes to be more identity altering. I hypothesize that in frame type #2 (half brain-cultural), East Asians will be especially susceptible to identity discontinuity following changes in one's external relationships/situation due to a social world fixedness and an emphasis on Confucian values of relationships, harmony, "fitting in", and roles. This will result in lower ratings of both personhood and personal identity relative to both the neutral and Western conditions. On the other hand, I predict Western cultures will be unaffected in personhood and identity ratings following the shift in one's social sphere. Compared to other countries, U.S. culture emphasizes individuals' uniqueness (Markus and Kitayama 1991, 2010; Savani, Morris, and Naidu 2012) and the autonomy to opt into or out of social roles and relationships (Schug, Yuki, and Maddux 2010). Due to the presumed malleability of the social world, and inconsequential nature of environment on self, Westerners thus should not view such external/social changes as undermining a person's identity continuity.

While in the last frame I predicted Westerners to be unaffected, and East Asians largely affected, for frame type #3 (PVS-cultural), I predict the opposite. Due to the drastic internal psychological/moral changes, I believe Westerners will interpret the scenario as more identity altering (lower ratings for personhood and identity). This is because dramatic changes in one's internal traits can threaten identity as it alters a Western belief of essential substance of self (Heine 2008). East Asian individuals, in contrast, have been observed to have a less clear concept of self, and not see consistency as necessary for psychological well being (Heine 2008). Hence, I predict East Asian individuals will see the situation with fluidity (or merely as an unfortunate consequence) and display no change in ratings.

The neutral condition for frame #2 (half brain) is as below:

> The year is 2450 and human civilization has advanced greatly. Sara, a young woman, was recently diagnosed with an incurable wasting condition in her body. But her brain is perfectly fine, so doctors recommend growing a new body to host Sara's

> consciousness. The host body is grown from Sara's own stem cells, so it perfectly matches Sara's DNA and physical stature. Doctors reorganize the information in Sara's brain so that all memories, emotions, and traits are redundantly duplicated in each half of the brain. After administering a sedative, one team of doctors carefully removes one half of Sara's brain and carefully implants it into the new body, which is the wheeled to the West Recovery Room. Simultaneously, another team of doctors carefully implants the other half of Sara's brain into a preservation tank and saves it as a backup. Sara's sister is waiting anxiously in the West Recovery Room. A familiar voice says, "Sara, it's great to see you."

Participants rated their agreement with the following statements on a scale of 0 (strongly disagree) to 10 (strongly agree):

1. After the event, the individual is still Sara
2. After the event, the individual is still the same person as before the event
3. After the event, the individual is still a person
4. After the event, the individual only has one half of her physical brain [comprehension check]

In the cultural condition, the following paragraph was added to the end of the scenario:

> After Sara's half-brain operation, though, an unexpected consequence occurs, and Sara gains the "ability" to see and speak to ghosts. Slowly, the people around Sara start to develop paranoia. Despite their love for Sara, her peers and family all start to avoid her, and she is seen a threat. The government hears about Sara, and sentences her to death. Only through losing her name, undergoing an unrecognizable appearance change, and "ceasing to exist" to her family/friends, will Sara be allowed to live. Sara accepts these conditions, and goes on to live the rest of her life in solitude.

In the cultural condition, the following comprehension check was provided: "After the event, the individual still sees her friends."

The neutral condition for frame #3 (PVS) is as below:

> Anh is a freelance writer. On Anh's way back home, his car was struck on by a truck. The ambulance arrived quickly, but there was not much they could do for Anh. Although Anh did not die, he entered a Persistent Vegetative State. While his body is still technically alive, he will never regain consciousness.

Participants rated their agreement with the following statements on a scale of 0 (strongly disagree) to 10 (strongly agree):

1. After the event, the patient is still Anh
2. After the event, the patient is still the same person as before the event
3. After the event, the patient is still a person
4. After the event, the patient is still breathing [comprehension check]

In the cultural condition, the following paragraph was added to the end of the scenario:

> In a completely miraculous event, Anh regains consciousness. Anh remembers everything from before the accident. However, the mild-mannered, respectable, and expressive Anh all of a sudden awakes foul-tempered, erratic, and disengaged. Anh has also lost all interest in writing, and nurses have reported Anh stealing their belongings.

In the cultural condition: the following comprehension check was provided: "After the event, the patient has the same memories as before the accident".

FRAME TYPE #4: PERSISTING AS MANY

Prior cross-cultural research suggests a distinction between how Western and East Asian cultures reconcile, transcend, and even accept apparent contradictions (Nisbett 1990). According to this distinction, human thinking is guided by two separate classes of cognitive strategies that implement different computational principles. One can be described as intuitive, experience-based, or holistic, whereas the other can be described as formal, rule-based, or analytic (Evans

1996; James 1890; Neisser 1963; Smith, Langston, and Nisbett 1992; Tversky and Kahneman 1983). In the context of culture, research has found that Chinese beliefs of Tao and Yin-Yang have led Chinese to rely less on use of categories and formal logic, and believe that two conditions can be incompatible, but both have merit (Chang 1939; Mao 1937; Norenzayan et al. 2000), Western cultures approach conflicting scenarios with a more analytic, or rule-based logical approaches, tending to detach objects from their context to avoid contradictions (Norenzayan et al. 2000; Nisbett 1990). This suggests that in response to contradictions, East Asians have tendency towards a "middle way", or compromise (dialectic reasoning), while Westerner's drive for consistent logic results in "polarizing" responses to contradictions (one option true, the other false) (Peng and Nisbett, 1999; Davis, Nisbett, and Schwarz 2000). These conflicting approaches presents an interesting dichotomy in culture-cognition, and one I wanted to further explore in personal identity.

In my design, I adapted the famous "quantum teletransportation" thought experiment cited by Weaver and Turri (in press) (Parfit 1984) in which an individual creates a replica of oneself and travels to another dimension or location. The experiment aimed to test the one-person-one place rule, addressing whether or not a person can exist in two places at the same time (Mars and Venus).

The neutral condition remained close to the original framing (minor wording variations), but in the cultural condition, I altered the events to address these culturally-diverging reasoning styles through aiming at cases of apparent contradictions. I further abstracted the experiment through constructing an alteration to the popular "grandfather paradox". In this paradox, an individual time travels and "rewrites" their past by interfering with the origin of their familial line. As a result of this paradox, inconsistencies/conflicts emerge in regards to the time traveler's own existence. I adapted this paradox so that in an (admittedly) absurd scenario, the individual's replica exists in a separate time frame in which it fathers the current individual (Kim). By intentionally obscuring the grounds of this frame, I sought to examine how Eastern and Western cultures implicitly reconcile self in cases of contradiction, inconsistencies, and personal doubt/ambiguity (brought upon by Kim's self-proclaimed existential crisis).

I hypothesize that due to Western cultures "polarizing" reactions to contradictions, the apparent inconsistencies/paradoxes in this study will lead Westerners to respond towards one extreme of the scale. In the scenario, the

individual (Kim) clearly still exists in physical form; the contradiction rather exists on a meta scale. Therefore, I believe Westernized individuals will have trouble arguing Kim doesn't exist, and in a search for constancy, attribute consistently higher ratings of personal identity (mainly 10s). This would be consistent with findings of Kung et. al (2016) which found that Americans have a stronger presumption that personal identity is continuous over time. Comparatively, East Asian dialectic reasoning lends to a less clear concept of self (Heine 2008), and a tendency towards finding the "middle way". In this, I believe East Asians will "compromise" between Kim's physical and mental paradox/conflict, attributing averaged ratings (M=~5) of personal identity (name and individual statements).

The neutral condition is as follows:

> The year is 2450 and human civilization has greatly advanced. Kim is currently on Earth. His mother is on Mars. Feeling somewhat lonely, Kim enters the Quantum Teletransporter in his house on Earth creates a temporary extension of himself, which he scans and sends to visit his mother on Mars. In an instant, the quantum device scans his body and records the exact state of all his cells and brain states. Instantly, the information travels through an information wormhole to Mars where it is perfectly reconstituted in physical form. The traveler steps out of the Teletransporter into Kim's mother's apartment on Mars. With a smile, they hug and she says, "My dear! I'm so happy to see you!". Kim remains in his house on Earth, watching and living vicariously through this interaction.

Participants rated their agreement with the following statements on a scale of 0 (strongly disagree) to 10 (strongly agree)

1. During the event, the traveler is still Kim
2. During the event, the traveler is still the same person as before the event
3. During the event, the traveler is still a person
4. During the event, the traveler recognizes his mom [comprehension check]

In the cultural condition, the following paradox was added:

> Suddenly, however, something goes wrong with the Quantum Teletransporter, spinning and tumbling until it eventually opens another 30 years into the past. The traveler steps out of the Teletransporter into Kim's mother's apartment on Mars. His mother does not recognize the traveler's appearance—stating it must be the first time they met. "It seems you are a lost traveler". She says. Due to the glitch in the Teletransporter, Kim believes some of the extension's mental states and consciousness were not properly coded, but regardless a pleasant conversation with Kim's mother follows. When the time comes, Kim instructs his extension to return to Earth. However, the teletransporter malfunctions again, and his extension cannot return. Watching from Earth, Kim abandons his extension in the past, and shuts off his Quantum Teletransporter. After this event, Kim contacts his mother. "That's coincidental, because 30 years ago on this day, in a rather faithful encounter, I met a strange traveler." She says. "Your father was quite a handsome man." Kim feels a combination of nausea, emptiness, and confusion. Kim confronts the paradox that he is his own father. Kim is at a loss for words.

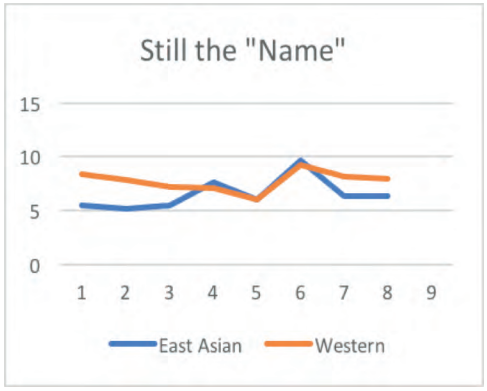
Participants rated their agreement with the following statements on a scale of 0 (strongly disagree) to 10 (strongly agree):

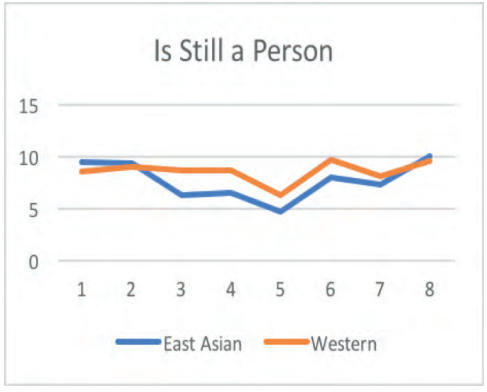
1. After the event, the person on Earth is still Kim
2. After the event, the person on Earth is still the same person as before the event
3. After the event, the person on Earth is still a person
4. After the event, the person on Earth still exists [comprehension check]

This frame addresses the following questions: cross-culturally, do people affirm findings of Weaver and Turri—in that judgments of personal identity are not committed to a one-person-one-place rule? Do different cultures have different interpretations of the idea that the mind can be divided into two streams of consciousness and have two simultaneous streams of experience separate from one another? Finally, do cultures reconcile logical fallacies/ambiguity/existential matters differently in thought experiments on personhood and identity?

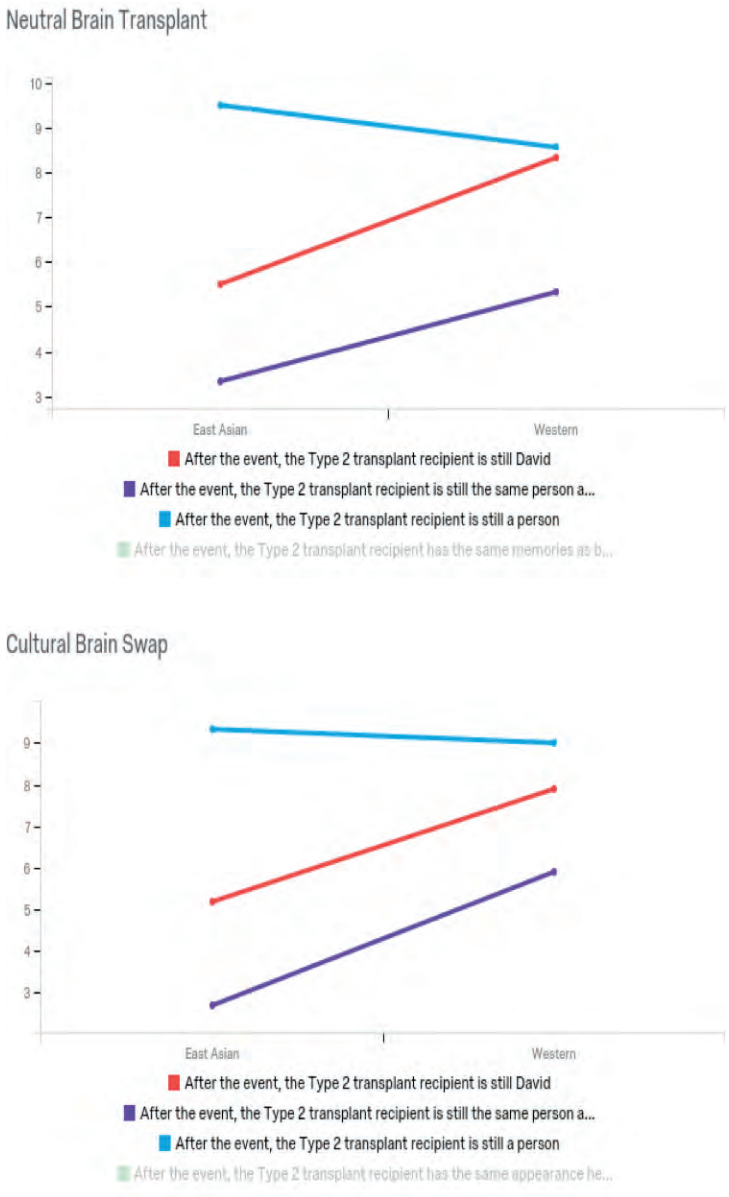
RESULTS

Independent t-tests (2-tailed) were performed on every question frame (8) x every condition (3) to determine whether East Asian and Western cultures displayed significant differences in approaching self along name, personhood, and identity statements. Out of the 16 computations, 3 emerged statistically significant. In the neutral brain transplant frame, East Asians reported weaker agreeance with the statement that the Type 2 transplant recipient "is still David" (M = 5.5, SD=2.43, SE = 0.99) than Westerners (M=8.33, SD=1.83, SE=0.61); t(13)=2.58, p=.022 (see Figure 5). This result was later repeated in the cultural brain transplant frame, with East Asian cultures once again showing lower ratings for "still David" (M=5.17, SD=2.19, SE=0.89) than Western cultures (M=7.89, SD=1.85, SE=0.61); t(13)=2.6, p=.022 (see Figure 6). The third significant result was also in cultural frame 1 in response to ratings on whether David was "still the same person". East Asian cultures perceived David as less of a person following the brain transplant (M=2.67, SD=2.62, SE=1.06) than Western cultures (M=5.89, SD=2.73, SE=0.91); t(13)=2.27, p=.04 (see Figure 6). All other hypotheses pertaining to Eastern and Western intuitions on personal identity emerged non-significant, although for 67% of the answers (83% including equal ratings), East Asians reported lower ratings for all 3 statements (still Kim, still person, still same person) relative to Western ratings (see Figures 3, 4, 5). While this pattern was not statistically significant, it formed a notable trend. Furthermore, in computing the total SD across all variables in East Asian and Western cultures, East Asians had a total SD of 70.58, whereas Westerners had a total of 60.2.





Figures 2, 3, 4. Mean ratings of Eastern and Western cultures across all conditions (#1-8; every even number corresponding to culture frame) on attribution of names, personhood, and identity.



Figures 5, 6. Mean ratings for influence of culture on question type across neutral and cultural brain transplant conditions.

DISCUSSION

Contrary to predictions, there emerged no significant effects of culture between neutral and cultural conditions, and also along attributions of personhood, identity, and name across frames. In this discussion, I will first directly address my general hypotheses. First, I predicted that neutral conditions would not elicit any significant cultural differences. However, neutral frame 1, condition 1 (still name) was significant. Upon further extension, I predicted that in the case of significance in the neutral condition, the effect would be amplified in the cultural condition. Out of the 3 statistically significant results, only 1/3 of the results corresponded to the same statement across frames (Q 1 and Q 2: still name). In both instances, the p value was 0.22. This rejects my hypothesis that significant interactions increase in cultural conditions. This is also demonstrated by the overall absence of significant interactions, in that I predicted a notable portion of cultural conditions would be significant when neutral conditions were not (this only occurred once out of 12). I also predicted that the overall SD will be larger for Western conditions than for the East Asian conditions; this was also disproven (East Asians SD =70.58, Westerners SD = 60.2). As in the research overview, while I did not analyze for differences between statements, one can see from Figure 2, 3 ,4 that individuals do respond differently to questions of personhood, identity, and name (supporting findings from Berniunas and Dranseika 2016), but cross-culturally it appears to follow a relatively similar trend.

These results converge to suggest that contrary to suggestive research, East Asians and Westerners generally did not adhere to their norms of analytical vs. holistic (dialectic) reasoning in personal identity thought experiments. This also suggests no consistent (significant) effect of framing of cultural conditions on intuitions of personhood and identity, and no/limited effect of malleability of self vs. world, relation of self to others, etc.

This result largely undermines some widely held philosophical assumptions about the influence of culture on self-construal, suggesting that cited cultural influences/diverging models do not consistently manifest in personal identity intuitions (through thought experiments). These results provide a great starting point and context about the treatment of personal identity across cultures. Of course, more discussion is warranted, and it is important to address possible considerations/confounding factors as to why this result may have been observed. Perhaps this finding is unique to thought experiments, in that its highly

abstract qualities isolate one's judgment from culture. Prior research on identity continuity (Kung 2016), tested for cultural effects in an explicit manner, giving Indian participants a passage largely imbued with cultural cues (5 statements on changes in external relationships). It is possible that this study's cultural frame approached culture in a too weak/ineffective way, and in the process to "trigger" cultural elements, only further obscured the scenario. This could explain the lack of significant effects in cultural conditions; however, I am aware that earlier studies could also be accused of being a form of "demand task". It is also possible that responses were confounded by testing fatigue, in which I observed as the trials went on, the answers became more and more restricted to arbitrary ratings of 0 or 10. This is supported by findings that the total SD for questions 1-4 was 12.77 (all 3 statements), whereas the total SD for question 5-8 was 18.72. These polarized ratings could have been significant to the hypothesis if it was noted only in frame #4 (the scenario with inconsistencies), but this was observed throughout and across both cultures. Whether or not this is due to testing fatigue, the high SD in later questions could possibly have obstructed the effects. This ties into the results that the 3 significant effects appeared in the first two questions, perhaps suggesting that participants were more intentional with their answers in the beginning. The alternate hypothesis to the results of frame type #1 is that personal identity intuitions only differ in situations of whole-brain transplants (unlikely, but possible). A more significant observation from these results in frame #1 is that when culture does interact with self, it only does so in respect to name and identity, and not in personhood (2/3 significant findings with "name" condition; 1/3 for "same person"). This suggests that Eastern and Western cultures converge on the concept of what makes a person, but not necessarily on name and identity in certain scenarios. Furthermore, the emergence of a significant result in the cultural condition only ("still the same person") suggests that the body swap to the father's body did elicit some cultural effect on intuitions about identity. While one would still deduce from these findings that personal identity is generally resistant to cultural influence, this does highlight an instance of cultural significance.

Another factor to consider is referenced in Grossman and Na (2014)'s paper, "Research in culture and psychology: Past lessons and future challenges". They note some recurring problems in measuring social orientation and cognitive style in culture and psychology, in that self-report measures come coupled with a lack of sensitivity and misleading conclusions when examining social orientation

across cultures. This is due to vulnerability to "deprivation effects", which are, "a tendency to prefer those properties that one feels are lacking in one's cultural environment". An example by Peng is illustrated, in which Americans may report valuing humility more than Chinese, even though humility may reflect greater interdependence (characteristic of Chinese cultures). In contrast, Chinese may report valuing personal choice more than Americans, even though personal choice is a sign of independence (Grossman 2014). Thus, it is possible that factors considered indicative of one culture may have actually produced the opposite effect, with Westerners responding in the way it was anticipated East Asian cultures would. This is a possible alternative explanation to the results regarding overall SD in respect to reasoning type. Future research would benefit from ratings not derived from self-reported data.

Another possible consideration is evidence indicating that some cognitive processes are highly susceptible to cultural influence, while others are not—for instance, naive theories of mechanics and physics (Baillargeon 1995; Carey and Spelke 1994; Leslie 1982; Spelke 1988, 1990), and naive theory of mind (Asch 1952; D'Andrade 1987; Leslie, 1994; Wellman 1990). Theory of mind is described as "the ability to attribute mental states, beliefs, intents, desires, etc. to oneself and to others" (Heine 2008) and to possess the ability to empathize/recognize other's situations. In my initial exposure to this concept, I distinguished it from "the philosophy of mind" (from which self-constructs are seen to emerge)—however, it is possible that in the frame of thought experiments, philosophy of mind is also impervious to cultural influence. Despite the bounty of cultural studies observing diverging models of cultural selves, it is possible that the intuitive processes behind theory of mind and philosophy of mind converge more than anticipated—explaining the insignificant cultural effects.

Another notable consideration of this study was the participant pool. Analysis was conducted on n=15 and due to access limitations, the East Asian condition was drawn from individuals who ultimately live in the U.S., go to school in the U.S., and have ample exposure to Western ideals/culture (regardless of whether they identify/exemplify it). In other studies that observed cultural effects, samples have been drawn from the host country. For example, Chiu and Hong (1997), tested cross-cultural differences with people in Hong Kong, and Nisbett (2001) tested Japanese participants at Kyoto University, Japan. While this cultural background was considered to its best ability, it is possible this could have affected the results.

From this discussion, I indicate a couple ways in which this study could be improved. First, a larger and more representative participant pool (currently living in country) would provide more generalizability and control for the cultural components studied and hypothesized. Data showed that for 67% of the answers, East Asians reported lower ratings for all 3 statements (still Kim, still person, still same person), and 83% when less than or equal to Western ratings. None of these results were significant, however, future research could address whether better sampling would push results in the predicted direction, showing that Westerners do exhibit significantly stronger identity continuity (across personhood, name, identity) than East Asians in thought experiments. This would support research that Americans have a stronger presumption of personal identity continuity over time (Heine 2008). This replication could also potentially amplify the effect of cultural frame types on identity ratings (perhaps to a point of significance). Additionally, the flexibility of a larger pooling sample could allow for a between-subjects design, which could eliminate any confounding factors due to repeated exposure to question frames. It is possible that under different conditions, this study could observe vastly different results. However, the current findings provides an important and equally informative foundation for further research and current understanding of cross-cultural percepts of identity.

In conclusion, this study aimed to address whether East Asian and Western cultures promote different intuitions in thought experiments about self in regards to personhood and identity. Overall, this specific study's results suggest that judgments of personal identity and personhood are fairly robust across East Asian and Western cultures.

APPENDIX MATERIALS (QUESTIONNAIRE)

The following questions were asked to participants post-experiment in no particular order:

> In what countries have you lived, and for how long? If visited, for how long? 2. What languages do you speak at home? How much for each (time spent, with whom)? 3. What is your race? Ethnicity? What do you identify as (ex. American, Chinese-American)? How closely do you identify with your ethnicity? (1-10) 4. If you identify with American, for how many generations has your family lived in the U.S.? 5. If you identify with an East Asian ethnicity, how

familiar are you with Confucian values (answer 1-10)? Do you feel you incorporate these values into your everyday life/perception of the world? (if not identify, can just write N/A)

For inclusion to analysis in the East Asian condition, participants selected had to be of an East Asian racial & ethnic background (Chinese, Japanese, or Korean), rated 7 (or up) for self-identification, rate 5 (or up) for familiarity with Confucian values, and either be fluent (use at least 50% percent of the time) in their country's native language or have lived there for 5+ years. For inclusion in the Western condition, participants selected had to have lived in the U.S. their whole lives, be 3rd generation or higher, speak English 80% or more of the time, and identify 7 (or up) with their self-identification. These factors were decided in interest of best controlling for cultural influence given the study's motive and available resources, but we recognize that race and personal identity is highly complex and non-reductive.

Additional survey questions were used for simple inquisitive means: How independent of an individual would you consider yourself? (1-10) In what ways do you feel you are independent? Or not? How important would you say friends and family are to your life (in terms of your identity)? (1-10)

REFERENCES

Berniunas R., Dranseika V. 2016. "Folk Concepts of Person and Identity: A Response to Nichols and Bruno." *Philosophical Psychology* 29 (1): 96-122

Blok, S., Newman, G., & Rips, L. 2005. "Individuals and their concepts." In *Categorization inside and outside the laboratory: Essays in honor of Douglas L. Medin*, edited by W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff, 127–149. Washington, DC: American Psychological Association

Blok, S. V., Newman, G., Behr, J., & Rips, L. J. 2001. "Inferences about personal identity." In Proceedings of the 23rd Annual Conference of the Cognitive Science Society, edited by J. D. Moore & K. Stenning, 80–85. Mahwah, NJ: Erlbaum

Chiu, L. H. 1972. "A cross-cultural comparison of cognitive styles in Chinese and American children." *International Journal of Psychology* 7: 235-242.

Heine, S.J. 2008. *Cultural psychology*. New York: Norton.

Grossmann, Jinkyung Na. 2014. WIREs Cogn Sci 5: 1–14. doi: 10.1002/wcs.1267

Grossmann, Igor & Na, Jinkyung. 2014. "Research in culture and psychology: Past lessons and future challenges." Wiley interdisciplinary reviews. *Cognitive Science* 5. doi: 10.1002/wcs.1267.

Kitayama, S., Markus, H.R., Matsumoto, H., & Norasakkunkit, V. (1997). "Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan." *Journal of Personality and Social Psychology* 72: 1245–1267.

Markus, H.R., & Kitayama, S. 1991. "Culture and the self: Implications for cognition, emotion, and motivation." *Psychological Review* 98: 224–253.

Markus, H.R., Uchida, Y., Omoregie, H., Townsend, S., & Kitayama, S. 2006. "Going for the gold: Models of agency in Japanese and American contexts." *Psychological Science* 17: 103–112.

Parfit, D. 1971. "Personal Identity." *The Philosophical Review*: 803–827.

Parfit, D. 1984. *Reasons and persons*. Oxford, England: Oxford University Press.

Rose Markus, Hazel & Kitayama, Shinobu. 2010. "Cultures and Selves: A Cycle of Mutual Constitution." *Perspectives on Psychological Science* 5: 420-430.

Shaun Nichols & Michael Bruno. 2010. "Intuitions about personal identity: An empirical study." *Philosophical Psychology* 23 (3): 293-312.

Strohminger, N., & Nichols, S. 2014. "The essential moral self." *Cognition* 131: 159–171.

Strohminger, N., & Nichols, S. 2015. "Neurodegeneration and identity." *Psychological Science* 26: 1468–1479.

Suh, E.M. 2002. "Culture, identity consistency, and subjective well-being." *Personality and Social Psychology Bulletin* 86: 1378–1391.

Turri, J. (in press). "Personal identity and persisting as many." In *Oxford studies in experimental philosophy*, volume 2, edited by T. Lombrozo, J. Knobe & S. Nichols. New York: Oxford University Press.

Williams, B. 1970. "The self and the future." *The Philosophical Review* 79 (2): 161–180.

compos mentis

Relevance Realization in the Formation of Mental Representation: Analyzing Discrete Thought: Why Cognition Must Use Discrete Representations

Tong (Victoria) Liu
University of Toronto, St. George

ABSTRACT
This paper is to discuss the roles of computationalism and discrete representation in mental representation by focusing on the paper written by Dietrich and Markman in 2003, Discrete Thought: Why cognition must be discrete representation. This paper will develop the important role that relevance realization plays in mental representation by arguing that both computationalism and discrete representation are inadequate in explaining mental representation. This is because they cannot explain the formation and changes in functional role. They also cannot explain "meaning" in connections. Moreover, they cannot explain the relationship between discrete representation and discrimination. After establishing the crucial role relevance realization plays in mental representation, I will propose the construct of dynamic systems which will be a theoretical alternative to discrete representation.

KEYWORDS
Relevance Realization, Mental Representation, Computationalism, Dynamical System, Continuous Representation

BACKGROUND

The nature of the mind has always been one of the most important questions in philosophy. There are currently two main theories debating for the nature of the mind, one is computationalism and the other is the neural network theory.

The computational theory of mind states that the mind is a formal system, which works by encoding information in abstract propositions and manipulating the propositions in a logical and mathematical fashion. A formal system should have tokens and rules, and in a formal system, the cognitive agent manipulates the tokens following the rules (Haugeland, 1985). An example of a formal system is chess, where each chess piece is a token and how to move them around being the rule. From 1968 to 1970, an artificial intelligence computer program named SHRDLU was made to understand natural language (Gardener, Davidson & Winnograd, 1979). SHRDLU was made by implementing computationalism on a machine, but it failed in the end and its failure marked the end of the prime age for computationalism. There are many factors contributing to its failure, one of which is the inability of making meaningful connections, which will be discussed in depth later in the arguments (Vervaeke, 2017).

After the failure of good old-fashion AI (GOFAI) using computationalism, the neural network theory has emerged. Neural network theory is also called connectionism. The two significant achievements of the neural network theory are unsupervised learning and unsupervised plasticity. Unsupervised learning and plasticity enable the machine to learn and alter its own architecture to achieve its goal, which is something the computational theory of mind cannot achieve. Neural network theory also brings out the notion of a dynamic system, in fact, both unsupervised learning and plasticity are examples of dynamic systems. A dynamic system is self- organizing and time dependent. The disproportionality between input and output in a dynamic system makes the time-dependent property extremely important. An example of a dynamic system is self-organized criticality.

INTRODUCTION

Computational theory of mind has been an important theory in cognitive science, it can explain the close relationship between language and cognition (Fodor,2008). However, there are aspects of mental processes it cannot explain, such as mental representations that this paper is going to focus on. A representation

is any internal state that mediates or plays a mediating role between a system's input and output (Dietrich& Markman, 2003). In the view of the computational theory of mind, a representation serves as a data structure that can be operated on by algorithmic processes (Dietrich, 1990). It has been defined that a system has discrete representation if it has more than one representation and the representations are bounded and uniquely identifiable (Dietrich & Markman, 2003). Whereas, if a system has fuzzy boundary but still represent its environment, it has continuous representations.

In this paper, I will first argue that discrete representations are not sufficient to explain the formation and changes in functional roles of a concept. Then I will argue that discrete representations cannot give an account of meaning. At last, I will argue that discrete representation and discrimination are doubly disassociated. All the three arguments will end with the important and not-to-be-trivialized role that relevance realization plays. The thesis of this paper is to point out the crucial role of relevance realization in mental representation, and why computationalism and the discrete representation are not adequate to explain the mental representation.

ARGUMENT 1: THE FORMATION AND CHANGING OF FUNCTIONAL ROLES OF A CONCEPT

In the paper, *Discrete thought: Why Cognition Must Use Discrete Representations,* Dietrich and Markman argued that in human cognition, concepts are highly interconnected, and the functional role of a representation is determined by such connections. Since continuous representations are time dependent and always changing, they cannot be connected to each other. Therefore, they conclude that continuous representations cannot form functional roles.

There are two problems with the argument Dietrich and Markman made. The first problem concerns with how the functional role is determined. As stated above, the functional role of a concept is determined by its connections. However, the connections can go on indefinitely, because any two things can be connected directly or be connected through an intermediate concept. For example, the representation of the "titanic" can be connected with the representation of "boat", and "boat" can be connected with "sea", "sea" with "fish", "fish" with "food" etc. The connections can go on forever. The formation of a functional role faces a combinatorial explosively large web of connections. So, in order to

make applicable functional roles in a mind or a machine, the system must have a strategy to deal with the combinatorial explosion, that is, to be able to determine which connections to take into consideration when determining the functional role. In the language of computationalism, this means to draw a boundary in the connection web. So far, neither computationalism nor discrete representation explains how to draw the boundary.

The problem of "drawing the boundary" is a problem of aspect, because out of all the possible connections, we only consider a few. Aspects are caused by relevance realization, because we need to only consider the properties relevant to the problem at hand (Searle, 1993). Computationalism pre-suppose relevance realization, thus it can never explain relevance realization (Cherniak, 1986). Now it is sufficient to say that neither computationalism nor discrete representation is adequate here to explain the formation of functional roles.

The second problem is about the changing in functional roles of representations. Dietrich and Markman have stated that discrete representations are non-dynamic and are not time-dependent. Therefore, we can deduce that the functional roles of discrete representations are stable and the connections are rigid because the connections do not change spontaneously. However, the evidence has shown that functional roles are not stable. Instead, they are constantly changing. For example, imagine a baby who has never seen a table before now sees a dinner table for the first time. Before this moment, the baby did not have a representation of "table". After that, the baby's representation of "table" is "dinner table", and the functional role of a table for s/he is to "serve dinner". Later, when the baby sees a desk, the functional role for "table" would not just be "serving dinners", but also "a place to work or study". The functional role for "table" has been enriched in the baby's mental representation. Such changes in functional roles cannot be explained by discrete representations or computationalism based on the argument above. Therefore, it is very likely that the discrete representation is not adequate to account for functional roles of mental representations.

People who support discrete representation might argue that the system can in fact change the functional roles of its representation. The system can do so by breaking old connections and making new ones. Such changes can be achieved by first manipulating the algorithmic processes, that is, to have a mathematic function to determine when to break and make connections.

I would argue against their defence by pointing out that discrete representations are not time-dependent, so the system can only be changing randomly or not changing at all. If the system breaks and makes new connections by manipulating the algorithmic processes, there must be an input to trigger the activation of the algorithmic processes. And if there is such an input that can trigger the changes in connections, the system should be classified as dynamic. Activating the algorithmic process at the wrong time (i.e. when there is no input) cannot result in the intended function. Therefore, timing matters in this system. Since discrete representations are non-dynamic and not sensitive to timing, that only leaves the possibility of a system changing its connections randomly.

Randomly changing the connections would be very insufficient in a system. Given the combinatorial explosively large number of connections that can be made, the probability of forming the desired functional role is indefinitely small. So, there must be a mechanism very much like relevance realization, to choose what new connections to form in order to make the system sufficient. Therefore, we can reject the idea of random changing of connections that the counter argument brought up.

Both problems point to the solution of a dynamic system. A dynamic system can change its connections spontaneously, therefore, its self-organizing property can give an account of how the functional roles are determined.

The time-dependent and environment-dependent property of a dynamic system can break and make connections as needed. This enables the system to have constantly changing functional roles. An example of a system that can spontaneously change its structure is a small world network (Asphaug et al., 1996). A small world network can deal with the trade-off between resiliency and efficiency, and it is likely to be the way how representations are connected. The self-organizing and self-correcting property should be able to generate the optimal structure to deal with the trade off. Therefore, the connections in SWN can be made and broken as needed, which is a property that formal system do not have.

When the dynamic system breaking and forming connection as needed, "need" is a subjective term which is determined by aspect. Having an aspect enables the cognitive agent to select a set of properties that is relevant to the problem at hand. Therefore, aspects are determined by relevance realization (Searle,1993). So, relevance realization is the core property of forming representations. Since

computationalism presupposes relevance realization as stated above, it cannot explain relevance realization (Cherniak, 1986). Therefore, discrete representation can't explain the formation and changing of functional roles where a dynamic system can. Hence, the computational theory of the mind is inadequate to explain mental representation.

ARGUMENT 2: DISCRETE REPRESENTATION AND MEANING

Dietrich and Markman argued that discrete representation can explain "meaning", because meaning is derived from the connections of concepts.

The problem with Dietrich and Markman's argument is that connections in computationalism presuppose meaning. Any two random things can have an indefinite number of shared qualities, but we only consider a set of share qualities to make connections. For example, "cats" and the "solar system" shared many qualities such as they both contain carbon atoms and they are both made out of molecules. Nonetheless, we don't consider them connected directly, but rather, we consider "cat" connected with "dog" and the "solar system" with "planets". This is because we have aspects. Aspects enable us to take a small set of properties into consideration when making the connections. Aspects are caused by relevance realization (Vervaeke, 2017).

Meaning is a special kind of connections. It is how a set of properties can stand out and be connected to the problem solver. It is another form of having an aspect. Therefore, meaning is also caused by relevance realization. As given above, since computationalism pre-suppose relevance realization, it can never explain relevance (Cherniak, 1986). Now it is fair to say that meaning is not derived from discrete representations or computationalism, and I will further conclude that meaning is the cause of them.

If a machine has "meanings", it must have the ability to do relevance realization. The baby in the previous example connected "table" with himself/herself, in a sense that the table can serve him/her dinner and him/her can study on the table. Now that the baby has a meaning assigned to "tables", s/he can integrate the representation of "table" into his/her already existing knowledge network by making new meaningful connections. Without meaning, the connections can be made anywhere, "table" could be connected with "Mars". Such random connections are not helpful to problem solving, and people almost never make that kind of connections. Therefore, how the representations are relevant to the

problem and the problem solver must be considered before making connections. That is why meaning is not a product of connections, but the cause of it. The ability to choose a set of properties is relevance realization, thus I conclude that relevance realization results in meaning and connections.

Therefore, we must abandon the notion that meaning is derived from connections, or it would be saying that meaning can result in connections, and connections can once again achieve meaning. This is a circular explanation that can get us nowhere.

One of the reasons why SHRDLU failed as a strong AI is that it did not have an aspect that can generate meaningful connections that are helpful to problem solving, due to the lack of relevance realization. The connections between information is knowledge, and the process of making those connections is what we call learning. The inability to make meaningful connections results in the lack of a knowledge network and the inability of learning, thus SHRDLU has very limited functions in problem solving (Vervaeke, 2017). For example, it was only able to respond to well-defined problems and interact with the world propositionally. SHRDLU's limited capability has made it unsuccessful in interacting with the real world. So, the inability to make connections has rendered it a weak AI, as the result of the lack of meaning.

I propose unsupervised plasticity as the solution to the problem above. Unsupervised plasticity is a dynamic system that uses a recursive self-organizing process to generate self-correction, thus change the structure of the wiring by itself without presupposing meaning (Fahlman & Lebiere, 1990). Under the theoretical construct of a dynamic system, meaning is not required to make connections, because the connections are not meant to be correct the first time they are made. The connections are meant to be corrected through the self-correcting process of the system. Therefore, the possibility to use connection to explain meaning opens again without falling prey to a circular explanation. Although this is not yet achieved, it is still a plausible way of reaching the goal of meaning.

ARGUMENT 3: THE DOUBLE DISSOCIATION BETWEEN DISCRETE REPRESENTATION AND DISCRIMINATION

Dietrich and Markman have argued that in order to distinguish between two external states, S1 and S2, there must be two corresponding internal states, R1 and R2 (Dietrich& Markman, 2003). The infinite intermediate states in the

middle have to be cut in half and the two chunks must be separately categorized into the corresponding internal states. They also argued that a continuous representation cannot discriminate among external states because it is a single varying representation, it does not have distinct internal states. Since only discrete representation can classify and categorize external states, they have given the following definition: A system has discrete representation if and only if it can discriminate its inputs (Dietrich &Markman, 2003).

Dietrich and Markman's argument above is very similar to the classical theory of categorization, thus it also inherits a lot of the criticisms of the classical theory. The theory states that the mental representation or concepts are driving the categorization. One of the criticisms of classical theory is that a lot of concepts do not have clear boundaries, so the memberships of the category are hard to decide and are unstable. So, the infinite number of intermediate states between S1 and S2 cannot be simply cut in half and separately categorized, because a lot of things cannot be categorized in this way. If the internal representation R1 and R2 have fuzzy boundaries, some intermediate states will be categorized under R1 and some under R2, while also leaving some to be undecided. An example of a category with fuzzy boundary is "game", the concept "game" does not have a clear boundary (Wittgenstein, 1953). For example, some may classify politics as a game, which is how "Game of Thrones" got its name. However, not everyone agrees with this classification, as politics is not a typical game like chess.

What the original paper suggested to discriminate input by discrete representation has been shown wrong. It is sufficient to say that the definition given is invalid.

Supporters of the discrete representation might argue that clear boundaries are not always needed to produce discrete representation. It has been argued by Dietrich and Markman that discrete representations are abstract, and abstraction is created by extracting information from a continuous stream of perception. Some detailed information is inevitably lost in the abstract, thus the abstract is a unity of all the important information. In this way, the information extracted is discrete, not continuous, thus the process of abstraction can produce discrete representations.

Discrimination and discrete representation can be doubly dissociated. If Abstractions are created through extracting information from continuous stream of perception (Dietrich& Markman, 2003), the extraction of information cannot be random. If the extraction was random, the probability of creating the desired

abstract is indefinitely small, given the combinatorial explosive amount of information we perceive every day. So, if the system does not extract information randomly, it must "know" what type of information to extract first. That is to say the system must have an abstract before it can do the extraction. If a system must have an abstract to extract information, it is to say that the system has an abstract because it has an abstract. A circular explanation. Therefore, computationalism cannot explain discrimination.

Having an abstract is to have a set of properties salient to the cognitive agent, which means the cognitive agent must be able to do relevance realization before it can extract information from the continuous stream of perception. We have already seen that neither computationalism nor discrete representation can explain relevance realization, we can now establish that an abstract and a discrete representation are not causally linked. Since having an abstract is the premise of being able to discriminate, a system which has discrete representation is not necessarily capable of doing discrimination, because of the lack of abstract.

We have already established that discrete representation does not necessarily mean discrimination. Now we are going to discuss the idea that discrimination is not equivalent to discrete representation. Unsupervised learning proposed by neural network theory can serve as an alternative. In unsupervised learning, the machine is able to feedback on itself and alter the weighing to generate self corrections. One significant improvement with the unsupervised learning is that it enables a machine to learn a task without having a pre-set model of the goal state. For instance, a machine is able to learn the shape of letter "A" when provided with abundant variations. This process has been demonstrated with the wake-sleep cycle. The wake-sleep cycle is to do data-compression in the wake state, and date-particularization to generate variations in the sleep state. Through multiple cycles of wake and sleep, the machine is able to learn the shape of letter "A" eventually (Hinton et al., 1995).

The wake-sleep algorithm uses continuous representation, because the representations in each state are constantly changing. It is also a dynamic system because the system is self-correcting. Thus, it enables a machine to distinguish and discriminate the input and achieve the goal state through self-correction without having an abstract or representation first. Therefore, it is evident to say discrimination does not necessarily require discrete representation.

In conclusion, a system has discrete representation is not necessarily capable of doing discrimination, and a system that can discriminate does not necessarily need to be a discrete representation. The discrete representation and discrimination is doubly disassociated. Therefore, the definition given by Dietrich and Markman is wrong and anything derived from that definition is doubtful to be correct.

CONCLUSION

This paper challenges computationalism and discrete representation in mental representation from the perspective of formation and changes in functional roles, the meaning of connections and the relationship between discrimination and discrete representation. While we establish why discrete representation is insufficient in mental representation, we also take a step further to examine the core of mental representation, that is, relevance realization. Finally, we propose the dynamic systems as the solution to the problem computationalism and relevance realization faces. Therefore, we argued that discrete representation is not how mental representation works, and continuous and dynamic representation is the right way to go.

REFERENCES

Asphaug, E., Moore, J. M., Morrison, D., Benz, W., Nolan, M. C., & Sullivan, R. J. 1996. "Mechanical and geological effects of impact cratering on Ida." *Icarus* 120(1): 158–184.

Dietrich, E. 1990. *Computationalism*. Retrieved from http://www.tandfonline.com/doi/pdf/10.1080/02691729008578566

Dietrich, E., & Markman, A. 2003. "Discrete thoughts: why cognition must use discrete representations." *Mind & Language* 18 (1): 95–119.

Fahlman, S. E., & Lebiere, C. 1990. The cascade-correlation learning architecture. In *Advances in Neural Information Processing Systems*, 524–532. Retrieved from http://papers.nips.cc/paper/207-the-cascade-correlation-learning-architecture.pdf.

Fodor, J. 2008. *LOT2: Language of Thoughts Revisited*. Oxford: Oxford university press.

Gardner, A., Davidson, J., & Winnograd, T. 1979. *Natural Language Understanding*. Stanford: Stanford University.

Haugeland, J. 1985. *Artificial Intelligence: The Very Idea*. Cambridge, Mass: The MIT press.

Searle, J. 1993. The problem of consciousness. *Consciousness and Cognition* 2(4): 310–319.

Vervaeke, J. 2017. *Introduction to Cognitive Science*. University of Toronto.

Wittgenstein, L. 1953. *Philosophical Investigations*. Blackwell Publishing, Inc.

compos mentis

Who Decides?

Rhea Thukral
Purdue University

ABSTRACT
The most unambiguous fact of life is death. Every human on the planet can, in complete confidence, say they will die. What is not so certain is how. Those who are lucky will die of old age. But what about those who sustain severe head trauma and are now "wakefully unconscious?" It is in these cases where the determination of death becomes difficult. The most complicating of cases is when the patient is found to be in a vegetative state (VS). In a VS, the patient can, unconsciously, hear, respond to nonspecific stimuli, and exhibit spastic limb movements. For the first year of being in this state it is logical to think that the patient will recover, but any time following diminishes these chances. Those who are endowed by law to make all decisions when the patient cannot, may hold on to the hope of recuperation and refuse to unplug, effectively ignoring the patient's desires. How can we know this patient's wants if they are unconscious and no explicit record exists? It is in situations like these when better decision-making mechanisms must be implemented. Rather than waiting for tragedies to occur and then sitting in confusion when situations like these arise, we must be prepared to maneuver the unexpected effectively. Refusing to converse about death is the crux of the issue; if the federal government implemented a universal end of life form many ambiguities involved in this process will be circumvented and legal battles regarding this issue will disappear.

KEYWORDS
Death, End of Life Policy, Vegetative State, Patient Considerations

BACKGROUND

Death is no longer a black and white situation; it has turned into a continuum where consciousness is on one end and brain death (BD) is at the other. Between consciousness and BD are two other states, the coma state and vegetative state (VS). The line between vegetativeness and death can only be clearly defined by the surrogate who can choose to leave the patient on life support until BD or cardiovascular death occurs. BD, as defined by legal communities and the American Medical Association (AMA), is a loss of brain function which is equivalent to cardiopulmonary death according to the Uniform Determination of Death Act of 1980 (UDDA). Unfortunately, not all communities agree with the medical and legal communities; some believe death occurs when the heart ceases which does not occur at the same time as brain death and can continue beating while the patient is on life support. A patient is not completely dead until life support has been removed as the body can continue to function without the brain if the patient is assisted (Sade 2011).

The technology which brought the question of death into a more complicated light was the organ support system. This new technology changed patient care drastically when it was introduced in the 1950's (Machado 2010). No longer was the patient dead or alive, they were in a limbo state. The machine allowed physicians to prolong death until a feasible treatment could be found, but it also forced the medical, ethical, and legal communities to redefine what a human is, and to determine the ethics of shutting off a patient's life support. Organ support technologies showed that a patient with a non-functional brain still has functional organs which lead to the question, is this person alive? It is now widely accepted that a patient who is determined to be BD is completely dead when previously the definition of death was the cessation of breathing and heartbeat (cardiovascular death) (UDDA 1980).

A distinction must be made between all forms of consciousness and unconsciousness. A coma is the umbrella term for a complete loss of consciousness. Between a coma and BD lie several other stages, but the state of main concern in this paper is the VS. A VS is defined, paradoxically, as wakeful unawareness; a patient in a VS is awake and can be aroused but is completely unaware of who they are or their environment; there is also evidence that patients in a VS have the ability to hear conversations. In other words, the patient is capable of organic functioning but, in the opinion of this author, lacks all the other characteristics

of humanism. Patients in a VS can spontaneously respond to different stimuli, for example a foul smell might elicit facial grimacing and other physiological responses, spastic movement of limbs, patients in a VS may even frown or smile (and in rarer cases weep or laugh). These responses cannot be confused with a regaining of consciousness as this is the paradoxical nature of the VS. A coma, unlike a VS, is characterized by the inability to force arousal and a complete lack of consciousness; there is no number of stimuli which could cause a patient in a coma to elicit a response.

In all cases, death decisions are put to the patient and the surrogate; the physician's only job is to provide information so the patient and surrogate may make an informed verdict. When the patient is no longer conscious, all information goes to the surrogate who is then to decide what to do next. There are only very specific reasons, according to the AMA, in which the physician can step in and question the surrogate's wishes:

- The patient and the surrogate cannot come to an agreement
- A surrogate cannot be found or the one listed is not willing and the patient is not in a condition to make decisions
- In the physicians best professional judgment, the surrogate is violating the patient's clearly expressed values, goals for care, or treatment preferences or the choice made is not in the best interest of the patient

In these cases, the physician must consult an ethics board before taking further steps in providing care. Additionally, the final point listed is highly susceptible to subjective analysis especially in cases where the patient has not explicitly expressed their values, goals for care, or treatment preferences (AMA Principles of Medical Ethics 2016).

The World Health Organization (WHO) states there is no uniform international consensus on the clinical criteria of brain death. In 2006, several countries, including Canada, published a consensus on the recommended method to determine BD: Computer Tomography Cerebral Angiography (CTA). In 2008, the Academy of Medical Royals College in the United Kingdom published that the preferred method should be a clinical diagnosis of BD and not neuroimaging.

However, the Academy did acknowledge that in certain cases clinical diagnoses cannot be performed and in other cases, neuroimaging increases certainty in BD diagnosis but no recommendations were given in these cases (Dineen et. al. 2012). There are several neuroimaging techniques a physician can choose to determine BD but there is only one clinical method, the apnea test. The apnea test has three subsets:

1. The patient must lack all evidence of responsiveness (coma)

2. Clear absence of brainstem reflexes (follows eye tracking motions which should not be changing as the head is put into different positions)

3. An absence of a breathing drive (Apnea)

The apnea test is usually repeated twice to ensure reliable results or is paired with other testing to improve reliability. If the patient is found to be BD the appropriate documentation is completed and an organ procurement organization is contacted if the patient was an organ donor (as dictated by federal and state laws). In cases where the patient is not found BD, the surrogate has one of two options which will be discussed later (American Academy of Neurology 2010).

A CTA measures blood flow through the blood vessels in the brain. To determine BD there are three different types of evaluation point scales; 10-point, 7-point, and 4-point. These scales refer to different numbers of intercranial (IC) blood vessels; the 7-point and 4-point scales measure vessels in either the posterior or anterior portions and the 10-point scale measures vessels in both the posterior and anterior portions. The common criterion between each of the three scales is a lack of opacification in the middle cerebral arteries and the deep cerebral veins. In other words, if the middle cerebral arteries and deep cerebral veins are not seen in the angiogram this is a sign of BD. Unfortunately, CTA scans are difficult read because it tracks blood through the blood vessels and therefore different forms of blood pools (e.g. hemorrhage) can obscure the reading and consequently confound the assessment of IC filling in BD patients (Dineen et. al. 2012).

Prior to complete cessation of IC blood flow, the IC pressure (ICP) increases dramatically which is the most important mechanism of BD. Given this mechanism

of BD, transcranial doppler ultrasonography (TCD) has been recommended to assess blood flow in suspected BD patients as it is non-invasive. TCD measures blood flow velocity and direction in proximal areas of the brain. Neurologists and internists are given training to perform, read, and interpret TCD scans which, unfortunately, means the results are operator-dependent and can lead to confounding results. BD patients exhibit a reverberating flow and/or systolic spikes which can be seen in a multi-depth TCD scan. A TCD or a CTA can be used in tandem with an apnea test to either determine or confirm BD (Machado 2010).

If the patient is not found to be BD the family can choose to wait, depending how long the patient has been in a VS or they can remove the patient from the ventilator. Patients who have been in a VS from one year or more are considered unlikely to rise from the state (Sade 2012). Because surrogates have the power of decision making, they can pick any of these two options which, if physicians have determined the chances of the patient recovering are not likely, means an untreatable patient is receiving treatment. If surrogates choose to leave the patient on organ support in the hospital, a patient who will probabilistically never recover is now occupying a bed that the hospital could be using for other critical care patients. This also means the family is supporting a patient who will likely never recover from their state and spending thousands of dollars doing so.

The American Academy of Neurology (AAN) guidelines state that organ support is to be discontinued once BD has been declared (AAN 2010). However, because the patient's surrogate is given the power to determine when organ support should be shut off, some people choose to continue to leave their family member on life support following determination of BD. Some states in the United States allow family members accommodations if they object to BD on religious, personal or other basis. The state of New York requires hospitals to have a set of guidelines in place for surrogates who object to the definition of BD. The Reasonable Accommodation clause in the Guidelines for Determining Death (2011), requires hospitals to give surrogates the option for continued artificial respiration for a set period during which surrogates are also given access to other resources such as members of clergy, ethics committees, bereavement counselors, and others to address concerns, objections, to determine next steps, or to come to terms with the situation. New York makes the distinction between those who object to the definition of BD and those who have difficulties understanding what has happened to their loved one and the legislation states that hospitals should

give both groups of people access to the same resources. When physicians start the process of determining BD, the surrogate is only required to be informed or an effort to contact the surrogate must be made; consent to perform BD determination testing is not required in the state of New York.

Each state has separate legislation regarding BD; California, New York, Illinois, and New Jersey all have some form of a reasonable accommodation clause in the legislation in cases where the surrogate disagrees with the definition of BD. New Jersey allows families to keep the patient on organ support until cardiopulmonary death occurs, which means the family would be supporting a patient who is legally and medically deceased. BD means the patient is never "going to wake up" as is the colloquial phrase. Patients who are in a suspected vegetative state (VS) have some potential of recovering some if not all their cognitive functioning. However, those in VS are not guaranteed to recover from their state and have the potential to slip further down the death continuum or stay in a persistent VS for the rest of their life. In these types of situations, the patient is not legally dead and forces many questions:

1. Do we have the right to turn off organ support even if the patient is not legally dead?

2. If the patient does recover from their VS what will their quality of life be?

3. Will their life be diminished complete dependence on their family members?

4. At which point does a person cease being a person?

5. How can we know if the patient would want to be kept alive if we have no record of their wishes?

DO WE HAVE THE RIGHT TO TURN OFF ORGAN SUPPORT EVEN IF THE PATIENT IS NOT LEGALLY DEAD?

It is medically accepted that once a patient has been in a VS for over one year there is little to no chance the patient will ever regain consciousness. This does not mean, however, that patients never arise from their VS. Researchers in France were able to awaken a patient who had been in a VS for 15 years by stimulating his vagus nerve which is responsible for several conscious functions. The patient is now capable of responding to simple commands, and perceived threats all after one month of vagal stimulation (Begley 2017).

This case brings up the second question: *If the patient does recover from their VS, what will their quality of their life be?* In the case of the French patient, he will be under constant scrutiny for the rest of his life because now he is a case study for his physicians. Additionally, we are unaware to what extent this stimulation will last. If the stimulation does wear off, will he revert into a VS, will this stimulation cause his condition to worsen and cause BD? Because of these questions, the French patient has lost nearly all his freedoms. His consistent need to be monitored, to relearn all motor movements, he may even need consistent vagal nerve stimulation; this monitoring may force the patient to stay in the hospital for most of his life. Is it ethical for researchers to experiment on humans before having a basic understanding of how this stimulation has an impact on other animals? A literature search found no further research into VS deep brain stimulation in other animals. The FDA requires that drugs are tested on animals before applying the treatment to humans, the same must be applied to this type of testing. This treatment could have a profound impact on this patient's life and the lack of research in this topic on other organisms is disconcerting. For example, deep brain stimulation in Parkinson's disease patients was first tested in monkeys before the treatment was given to humans and the same must be applied to VS deep brain stimulation (The Deep-Brain Stimulation for Parkinson's Disease Study Group 2001).

To further answer the main question in this section, in 1990, Terri Schiavo suffered a cardiac arrest causing severe hypoxia (lack of oxygen) and subsequent brain damage. It was determined that Terri was in a persistent VS and there was little to no chance she would ever recover. Given this information, Terri's husband wished to end her life citing that Terri would not want to be kept alive in such a state. Conversely, Terri's parents were not convinced of the evidence provided to

them by her physicians and wished to maintain her organ support. Both parties had such a severe disagreement regarding her care that this case was argued in front of the United States Supreme Court who ruled in favor of Terri's husband. Her feeding tube was to be removed, thereby removing all support to Terri and killing her. Autopsies later demonstrated Terri's brain had atrophied to the point where no amount of treatment or therapy could have reversed the damage. If Terri's parents had won their legal battle, her feeding tube never removed, and she remained supported by hospital resources, the hospital would be allocating resources to a patient who was never going to recover. Until Terri had been unplugged she was utilizing resources, physician time, and a bed in the Intensive Care Unit all of which could have been given to a patient whose condition would improve. It is not someone's given right to remove organ support, no one has the right to remove what is keeping another alive, but it is important to understand that by allocating resources to someone whose wishes we do not know or whose prognosis is not one of recovery we are costing other patients the right to have access to adequate care (this is an idea we will revisit).

IF THE PATIENT DOES RECOVER AND THEIR LIFE IS DIMINISHED TO COMPLETE DEPENDENCE ON FAMILY MEMBERS, WHAT THEN?

In a hypothetical situation, a woman has been in a VS for almost two years now and spontaneously one day she arises from her VS while her brother is sitting by her bedside. This woman is completely aware of who she is and her environment. She can respond to basic questions, but not verbally, she can only respond by blinking her eyes. Her neurologist informs the family members there is little chance she will be able to recover any other motor function and will require constant care. The family understands and moves the patient to hospice care. From the time at which the patient woke up to the time she will die, the family will not remember her how she was before her traumatic accident but what she is like now; frozen and incapable of basic human qualities like expressing intelligible thought. In every way but the medical diagnosis this patient is a vegetable. If the patient knew she would end up in this state, would she want to have been kept on life support? Every patient has the right to die with dignity; the right to die before severe complications cause other family members to intervene and take care of them. No one wants to die, it is why talking about death is essentially taboo. Humans know what lays at the end, but because humanity refuses to accept that

end, the subject is difficult to approach. This avoidance is understandable, but when patients are left in a state like this hypothetical patient, it is necessary to force conversations before family members choose to spend thousands of dollars and countless hours taking care of a loved one who will no longer be capable of taking care of themselves.

Valentine's day 2010, Sam Goddard suffered two massive strokes during a soccer match. Sam was a healthy, young accountant who was just weeks away from marriage and following these two strokes he was reduced to a man completely dependent on his fiancé and his family. Sam's physicians had determined he would never recover from his VS; Sam's fiancé disagreed and refused to let him go. She eventually found the sleeping medication, Stilnox which rose Sam from his VS. This drug worked miracles for Sam and his family but he can no longer go an hour without this drug. Sam's family has turned into his caretakers; Sam's fiancé can no longer love him as a spouse but as someone she must look after. Yes, Sam is happy to be alive, but his life has gone from a healthy active man to complete debilitation. Is this the life anyone would want their loved one to live? It is inspiring to see a patient with a poor prognosis recover, but everyone deserves to live and die with dignity. But the opinion of this author does not matter, what matters is that Sam is happy to be alive and above all it must always be the decision of the patient who the treatment is being applied to.

IS A PERSON WHO IS NO LONGER CAPABLE OF SELF-CARE, INTELLIGIBLE THOUGHT OR ANYTHING ELSE THAT INTRINSICALLY MAKES THEM A PERSON, A PERSON?

For years philosophers have been attempting to define what a person is and what they have concluded is that a person is capable of what any other person is capable of; intelligible thought, reason, emotions, self-awareness, and the lists go on (Farah and Heberlein 2007). The problem with attempting to define personhood is that what makes us human is not explicit, it is implicit and is verified by the people around us. The issue comes when someone is no longer acting as a person. The people around this human still want to validate his or her personhood because they know this human was a person once. Once a person's recovery stagnates to the point where they require constant care, cannot display the intrinsic properties of personhood, this person is now a human, an empty shell. A person is capable of more than being able to respond to basic necessity

questions; a dog can respond to basic need questions when trained. Once a person can no longer be defined as a person objectively and intuitively, they are a human. This is a difficult concept to grasp and seems wrong, but once a person loses all implicit facets of personhood they are now a human not a person.

HOW CAN WE KNOW IF A PATIENT WANTS TO BE KEPT ALIVE IF WE HAVE NO RECORD OF THEIR WISHES?

In cases where the patient is comatose on arrival to the hospital and then progresses into a VS for more than one year, caretakers likely do not have a record of this patient's wishes. It is the surrogate's decision to keep this patient alive; would this patient have wanted to be kept alive for this long without signs of recovery? This is the main reason why such conversations must be forced. Without knowing, the surrogate could be keeping their loved one alive when they don't want to be kept alive in such a state. These rules can be easily implemented in adult care, unfortunately children are much more complicated especially when cases of unexpected tragedy occur. In the state of Virginia there was a case of a toddler who choked on a popcorn kernel and was subsequently placed onto life support. Mirranda Grace Lawson's physicians wanted to perform an apnea test to determine BD, but the parents refused the testing stating that they thought the test would harm their daughter. Following this, Mirranda's parents gave a handwritten letter to the doctors citing their Christian beliefs do not allow for such testing to occur, as taking their child off life support is considered murder. The ensuing legal battle ruled in the hospital's favor, but the parents were allowed pay a $30,000 bond to prevent the hospital from performing the test while the parents appealed to the Virginia Supreme Court. This appeal was later retracted when Mirranda died of cardiovascular arrest. What must be considered is that the Virginia Commonwealth University Health System Pediatric Intensive Care Unit has only 14 hospital beds; not only did Mirranda's presence cause other intensive care patients to be turned away, but her care was costing the hospital $10,000 a day. This brings up two important points. First, if a patient is suspected legally dead, does the hospital have the right to perform the apnea test without the surrogate's consent? Second, does the life of the many other patients outweigh the wishes of one surrogate?

IF A PATIENT IS SUSPECTED LEGALLY DEAD DOES THE HOSPITAL HAVE THE RIGHT TO PERFORM THE APNEA TEST WITHOUT THE SURROGATE'S CONSENT?

In the case of children, if the surrogate is of sound mind, families have the right to refuse BD testing. In July of 2016, Allen Callaway was brought in to St. Vincent's Healthcare (SVH) in Montana. Following his admission he experienced a brain herniation, which lead Allen's physicians to suspect BD and the need to perform clinical testing; SVH sought consent from Allen's mother who subsequently agreed. The evaluation was found to be consistent with BD and to produce reliable results the physicians must perform the apnea test a second time. When Allen's doctors informed his mother, she refused and filed a lawsuit to prevent any further testing. The court found that performing medical procedures on a child requires parental consent citing personal autonomy, privacy, and protection. In this case, because the patient was a minor, Montana ruled in the parent's favor but if every physician sought the permission of the surrogate to perform apnea testing in every case there would be a blockage of death determination. Families do not want to learn that their loved is gone and this has implications not just within the family but would impact other patients who need Intensive Care Unit support (Lewis and Pope 2017).

This statement answers the next question; do the lives of the other patients outweigh the wishes of the one surrogate? Yes. It is difficult for people to comprehend death, but it is not something that must stop other people from receiving the care they need.

POLICY SUGGESTION

Death is never an easy subject to consider objectively. No one wants to consider dying despite knowing we will one day. Given these gray areas, there is a necessity to implement a system and a new law which can, to some effect, circumvent these ambiguous issues. Additionally, more research funding needs to go into studying death, the implications of being in the different stages of unconsciousness, and the most reliable way to diagnose BD. Furthermore, the public needs to become more educated in the different stages of death so when the time comes, all people are capable of making intelligible decisions.

All people have the right to die with dignity, but there are circumstances in which people are not informed of impending death. In cases like this, contingency

plans are best put into place. Like the organ donation card, people should be asked to fill out a form to determine end of life decisions before unfortunate events occur. It would be naïve to think this form would not be met with severe backlash, but the organ donation card implies the same concept. At the time of death, a human's organs would be procured; this form simply goes one step further and defines what the person wants at the end of their life. In cases where the patient is comatose on arrival and the surrogate does not know the patient's wishes, this form can become a rubric of sorts to guide the representative. With this form, people who wish to be left on organ support will be and those who wish to be removed off life support after a given period will be. Asking people to fill out this form would be difficult not only because forcing people to have a conversation they have avoided most of their life is uncomfortable, but also because it is impossible to reach the entirety of the United States population. The best way to circumvent both issues is to ask people to fill out these forms when they are obtaining or renewing a driver's license which is when the question of organ donation arises. Even though there are people who do not procure a driver's license, this is the best method to maximize form fulfillment. Most people procure their first drivers license at age 16, and despite this appearing to be very early to some people to start making such decisions, it appears that at any age a conversation about death is too early. After the form is filled out for the first time it would be updated every time the driver's license is to be renewed, which varies from state to state and is age dependent. This form would detail all the patient's wishes and desires in the event the patient becomes comatose and progresses into a VS; essentially this form would become the person's representative and the surrogate would follow this paper when making decisions regarding their care. Additionally, a federal law needs to be implemented regarding death. Currently, each state has the power to decide what sorts of responsibilities and privileges the surrogates and physicians can have. The federal government needs to create an umbrella law that all states and hospitals must abide by. In doing so, states with differing BD laws will not have people from outside states asking for care that their previous hospital could not provide due to stricter BD laws. This law would need to include clauses regarding reasonable accommodation and children. Death is an emotional process and for this federal law to be effective, people must be given time to process and understand what is occurring and states that allow for reasonable accommodation understand this process. Furthermore, pediatric

care is different than adult care. If the form goes into effect children under the age of 16 will not have forms, it is in these cases where the parents must make the decisions for their children. These two policies must be directed to Congress as they are the ones who can implement new laws and make changes to existing programs. They also direct the budget which would require changes to hire new federal employees to aid people in filling out these forms.

Our complete lack of understanding in what occurs during the limbo stages of death is preventing physicians from treating patients so they may return to normal consciousness. More research needs to be funneled into treatment of brain revival so that cases of VS's are no longer issues. To create more research, more grants need to be directed towards VS studies. The National Institutes of Health creates the grants and Congress creates the budget for the NIH to fund the varying types of research occurring all over the country.

Furthermore, the public needs to be better educated on this issue; most people are unfamiliar with the basic definitions of a coma, VS, and BD. This lack of understanding causes mass confusion when asked to make serious decisions and can be avoided with additional education. One of the jobs of the Department of Health and Human Services (HHS) is to provide education to the American people regarding various health hazards. It is the job of the HHS to create awareness for various medical conditions and determination of death must become one of those issues. The education policy must be directed to the HHS and to Congress. Congress would need to provide more funding to the HHS to allow for widespread education and research.

Additionally, the international medical community must come to an agreement on what the best method to test for BD is and the most reliable interpretation of these scans. In the case of a CTA, which point system is best used to determine the status of the patient's brain. In TCD scanning, uniform training must be given to the internists and neurologists performing the scans so that there is a uniform measure of a VS and BD. This policy can only be directed towards the WHO, who creates the international medical standards for physicians. To determine the best methods for measuring VS and BD more research needs to be conducted, funded by the WHO.

CONCLUSION

With newer life-support technologies emerging humans can no longer starkly define life and death. What used to be a clear line, has turned into a confusing and stepwise process which, in some cases, can revert and lead to recovery but often leave patients in a consistent state of limbo.

The bulk of the decision-making lays in the hands of the patient chosen surrogate; the physician can only provide information to the surrogate to make informed choices. Because the surrogate makes most of the choices for the patient they are representing, the surrogates' biases are inserted into their decision-making; religion, morals, personal experiences all come together to allow the surrogate to make a choice about another person's life. These representatives are meant to channel the patients' wishes and coordinate with the physician in accordance with these wishes. Unfortunately, most people do not plan for such events to occur and therefore do not express their wishes in the event of such a situation arising. A new system must be implemented to force people to have these conversations and to avoid situations in which the surrogate projects their wishes onto the patient when the patient cannot make those decisions. Additionally, more research needs to be done so that in the future we may be able to treat people in a VS. Death is never something people choose to discuss, but it is something that requires further education so that humanity becomes less fearful of the topic and so we can move towards less complicated ethical standards.

REFERENCES

A. 2016, July 08. "Family fights to keep toddler who choked on popcorn on life support." Retrieved November 19 2017.

A. 2010. "*Update: Determining Brain Death in Adults"* [Pamphlet]. St. Paul, MN: American Academy of Neurology.

Shea, Caitlin. 2017, June 12. "Love and other drugs: A family's journey to help their son after devastating strokes." *Australian Story*.

Begley, S. S. 2017, September 25. "Brain Stimulation Partly Awakens Patient after 15 Years in Vegetative State." Retrieved November 19 2017.

Caplan, A. 2015, March 31. "Ten Years After Terri Schiavo, Death Debates Still Divide Us." *Bioethicist*.

Code of medical ethics. 2016. Chicago, IL: American Medical Association.

Department of Health. (n.d.). Retrieved November 19 2017.

Eder, K. E., Haussen, D. C., Searls, D. E., & Henninger, N. 2012. "Teaching NeuroImages: Reverberating TCD flow pattern in brain death." *Neurology* 79 (9).

Farah, M. J., & Heberlein, A. S. 2007. "Personhood and Neuroscience: Naturalizing or Nihilating?" *The American Journal of Bioethics* 7 (1): 37-48.

Guridi, J., Rodriguez-Oroz, M. C., & Obeso, J. A. 2008. "Deep Brain Stimulation of the Globus Pallidus Pars Interna and Subthalamic Nucleus in Parkinsons Disease: Pros and Cons." *Deep Brain Stimulation in Neurological and Psychiatric Disorders,* 277-289.

Laureys, S., Owen, A. M., & Schiff, N. D. 2004. "Brain function in coma, vegetative state, and related disorders." *The Lancet Neurology* 3 (9): 537-546.

Lewis, A., & Greer, D. 2017. "POINT: Should Informed Consent Be Required for Apnea Testing in Patients With Suspected Brain Death? No." *Chest* 152 (4): 700-702.

Lewis, A., Varelas, P., & Greer, D. 2015. "Prolonging Support After Brain Death: When Families Ask for More." *Neurocritical Care* 24 (3): 481-487.

Lewis, A., & Pope, T. M. 2017. "Physician Power to Declare Death by Neurologic Criteria Threatened." *Neurocritical Care* 26 (3): 446-449.

Machado, C. 2010. "Diagnosis of Brain Death." *Neurology International* 2 (1).

Muramoto, O. 2016. "Informed consent for the diagnosis of brain death: a conceptual argument." *Philosophy, Ethics, and Humanities in Medicine* 11 (1).

Sawicki, M., Bohatyrewicz, R., Safranow, K., Walecka, A., Walecki, J., Rowinski, O., Wojczal, J. 2014. "Computed tomographic angiography criteria in the diagnosis of brain death—comparison of sensitivity and interobserver reliability of different evaluation scales." *Neuroradiology* 56 (8): 609-620.

Sade, R. M. 2011. "Brain Death, Cardiac Death, and the Dead Donor Rule." *Journal of the South Carolina Medical Association* 107 (4): 146-149.

"Schiavo autopsy shows irreversible brain damage." 2005, June 16. Retrieved January 12 2018, from http://www.nbcnews.com/id/8225637/ns/us_news/t/schiavo-autopsyshows-irreversiblebrain-damage/#.WlZHPKhKvic

Taylor, T., Dineen, R. A., Gardiner, D. C., Buss, C. H., Howatson, A., & Pace, N. L. 2014. "Computed tomography (CT) angiography for confirmation of the clinical diagnosis of brain death." *Cochrane Database of Systematic Reviews*.

Williams, T. 2005, June 15. "Schiavo's Brain Was Severely Deteriorated, Autopsy Says." NY: New York Times.

compos mentis

Human Intuition in Folk Mereology

Yufan Yang
Northwestern University

ACKNOWLEDGMENTS
I would like to thank Dr. Lance J. Rips for the insightful comments and advice during the writing of the manuscript and for piquing my interest in experimental metaphysics.

ABSTRACT
Mereology, the study of parts and when parts are considered an object, was brought to the forefront of the metaphysical field following the proposal of Peter van Inwagen's special composition question. The special composition question was proposed to determine if there was a definitive answer regarding the conditions that lead to mereological composition, or composition that defines a sum of parts as an object. While there have been many answers to the special composition question, each answer has been criticized on the grounds that they violate basic human intuition, especially intuition of the folk—non-philosophers. In Rose and Schaffer (2017), it is argued that answers to the SPQ should be liberated from having to accommodate folk mereology—folk intuition on mereological composition—due to folk mereology being tainted with teleological thinking and adhering to teleologically restricted composition. However, experimental design flaws and the failure to acknowledge the creative intentions account damages the results and claims made in Rose and Schaffer (2017). This publication presents a similarly-designed study to that of Rose and Schaffer (2017) but with revisions that incorporated the creative intentions account and fixes to other such criticisms. Utilizing questionnaires that asked participants mereological composition questions following the presentation of scenarios that featured conditions such as 'added function' or 'added creative intent,' the results of this study suggest that folk mereology is influenced more heavily by creative intent restricted composition rather than teleologically restricted composition. Though the results of this study need to be substantiated by an identical, larger-scale study, it is discussed that, should the results of this study were to become validated, folk mereology cannot be debased on account of teleological influences.

KEYWORDS
Special Composition Question, Folk Mereology, Teleology, Teleologically Restricted Composition, Creative Intentions Account, Creative Intent Restricted Composition, Experimental Metaphysics

INTRODUCTION

The conditions that determine mereological composition have been a leading metaphysical topic of debate for many years. Mereology can be defined as the study of when parts are considered a whole, or an object. A definitive answer as to when humans believe mereological composition—composition where a sum of parts is considered an object—occurs can lead to broad implications regarding how humans generally perceive the material world. Such debate on what constitutes an object was first brought to the forefront of the philosophical field by the 1990 book *Material Beings*, in which Peter van Inwagen proposes the "Special Composition Question" (SPQ). The SPQ asks, "Under what conditions does plurality, or a sum of parts, X, become a single object, Y?" (van Inwagen 1990, 20) and has garnered a spectrum of answers. *Universalism* answers the SPQ by claiming that a sum, X, composes a single object under every condition. *Nihilism*, on the other hand, claims that a sum, X, does not compose a single object under any condition (van Inwagen 1990, 72-75). These two views establish the polar ends of the mereological spectrum, and all other answers to the SPQ fall in between these two. Van Inwagen himself composes an answer to the SPQ, theorizing that composition occurs only if a sum, X, constitutes a life (van Inwagen 1990, 81-97). However, virtually every answer to the SPQ has been deemed unsatisfactory by many philosophers on the grounds that each of these answers violate the intuition of the folk, or non-philosophers (Rose and Schaffer 2017). The question inquiring whether or not answers to the SPQ need to accommodate folk intuitions, or if folk intuitions even have any validity at all, is discussed in-depth in Rose and Schaffer (2017). In Rose and Schaffer (2017), it is argued that folk intuition regarding mereological sums, or folk mereology, is inconsequential to metaphysical debates about whether sums of parts form wholes and thus, should liberate any answers to the SPQ from having to accommodate it.

FOLK MEREOLOGY IS TELEOLOGICAL

Rose and Schaffer claim that folk intuition is inconsequential in metaphysical debates due to it—folk intuition—being steeped with teleological bias as a result of teleological thinking (147, 2017). Teleological thinking is the thinking that existence occurs as a result of function; for example, the thinking that a knife exists because it can cut things is teleological in nature. Teleological thinking is not inherently flawed. Dink and Rips (2017) state that "teleological explanations

are naturally suited" (Dinks and Rips 2017, 207) for the realm of people and artifacts, as there is validity to teleological explanations for artifacts, given that many artifacts are created when there is a need for a certain task to be done. Still, teleological thinking can be problematic when it extends past the artifact realm. This is highlighted by the idea of 'promiscuous teleology,' which provides the backbone of Rose and Schaffer's argument.

'Promiscuous teleology', as presented in Rose and Schaffer, is the idea that humans have a natural tendency to view all things in nature as existing because of a function (143, 2017). Studies have found that children will prefer outlandish teleological explanations such as rocks are pointy so that animals wouldn't sit on them and smash them' versus mechanistic explanations rooted in science and facts (Kelemen 1999, 1443-1444). While less is often expected from children in regards to an understanding of the physical laws that govern the world, work in the psychological field has found that adults, though more selective in their use of teleology, also have an extreme bias for teleological explanations. The teleological bias found in childhood persists through adulthood and is merely "masked" (Lombrozo, Kelemen, and Zaitchik 2007, 999-100). One notable example of the underlying persistence of teleological thinking in adults can be found in Kelemen et al. (2013), which reported that even physical scientists, who firmly stand by the mechanistic explanations of the world, defaulted to teleological explanations such as, 'trees produce oxygen so that animals can breathe,' when they were cognitively pressured (Kelemen et al. 2013, 1079).

Rose and Schaffer (2017) support the theory of 'promiscuous teleology,' stating that the idea's rejection of mechanistic explanations and the idea's presence in folk intuition makes folk intuition unscientific. Furthermore, Rose and Schaffer (2017), expand upon the idea of 'promiscuous teleology' by stating that humans, especially the folk, not only display 'promiscuous teleology,' but are also 'teleomentalists' (Rose and Schaffer 2017, 144-145), ones who heavily use teleology to explain their surrounding world (Allen et al.1995, 13-14 and Heider et al. 1944). Due to the inclusion of 'promiscuous teleology' in folk intuition, as well as the folk being teleomentalists, Rose and Schaffer hypothesized that adding a function to a sum of parts will increase the likelihood that the folk will perceive the sum as an object (2017, 148).

There were four studies performed in Rose and Schaffer (2017) to test this hypothesis. The studies were consistent in their overall structure, as in each study,

participants were presented with a scenario that featured a sum of things and asked—on a 1-7 scale, 7 being that they completely agree—if the sum composed a single object. Each scenario featured two variations, one in which the sum of things performed no function, and one in which the sum of things did perform a function.

In Study 1, the scenario featured two politicians shaking hands and tasked the participants with choosing from one of two viewpoints when presented with each variation of the scenario: the viewpoint that the politicians did not form a single object during the handshake or the viewpoint that the politicians did form a single object during the handshake. Present in the with-function (WF) variation of the scenario and absent from the non-function (NF) variation of the scenario was the condition that the image of the politicians shaking hands was used as a model for a statue called "Unity". The results of this study revealed that, when presented with the WF scenario, the participants expressed far greater agreeability towards the view that the handshake was an object than when the participants were presented with the NF scenario. The NF scenario garnered a mean agreeability score of 2.48 (SD=1.84), meaning that the participants generally thought that the handshake between the politicians did not form a single object, while the WF scenario garnered a mean agreeability score of 4.86 (SD=1.60), meaning that the participants on average leaned towards the view that the politicians did form a single object. A switch in intuition was observed, with the participants on average changing their answers from "the sum is not an object" to the "the sum is an object" in the presence of the added function (Rose and Schaffer 2017, 148-151).

In Study 2, the scenario featured two researchers who had found a new substance called "Gollywags" and glued two of them together; participants were tasked with choosing between the viewpoint that the glued Gollywag sum formed a new, single object or the viewpoint that the glued Gollywag sum did not form a new object—that it was two separate Gollywags glued together—when presented with each variation of the scenario. Present only in the WF variation of the scenario was the condition that the glued Gollywag sum formed the "Gollywag-Supporter," and that the supporter helped alleviate back soreness in one of the researchers after he placed it on his chair. The results of this study revealed that, when presented with the WF scenario, the participants expressed greater agreeability towards the view that the Gollywag sum was a distinct object than when the participants were presented with the NF scenario. The NF scenario

netted a mean agreeability score of 3.85 (SD=1.94), meaning that participants were generally unsure, but on average leaned slightly towards the viewpoint that the Gollywag sum was not a new, single object. The WF scenario netted a mean agreeability score of 5.15 (SD=1.99), showing that participants on average did lean towards the Gollywag sum being a new, single object. The results of this study did not have the drastic switch in intuition found in Study 1, but a noticeable difference was still observed (Rose and Schaffer 2017, 151-153).

In Study 3, the scenario featured researchers who glued two mice together, and tasked participants with choosing between the viewpoint that the glued-mice sum was a new, single object or the viewpoint that the glued-mice sum did not form a new object—that it was just two separate mice glued together—when presented with each variation of the scenario. Present only in the WF variation of the scenario was the condition that the two mice were glued together side-to-side and formed "the mini-bomb detector". Rose and Schaffer, as an additional variable, wanted to see the effect labeling a sum would have on folk mereology, hence the inclusion of the name "the mini-bomb detector". The results of this study revealed that, when presented with the WF scenario, the participants expressed greater agreeability towards the view that the glued-mice sum was a distinct object than when the participants were presented with the NF scenario. The NF scenario had a mean agreeability score of 3.0 (SD=1.73), while the WF scenario had a mean agreeability score of 4.7 (SD=1.15). Once again, this study revealed a switch in the intuitions of the participants when a function was added, similar to Study 1. Rose and Schaffer performed other variations of this study without the label, "the mini-bomb detector," and concluded that they found no detectable differences between the no-label conditions and the with-label conditions (Rose and Schaffer 2017, 153-155).

In Study 4, the scenario featured an avalanche that scattered a bunch of rocks in front of a man's lawn. Participants were tasked with choosing between the viewpoint that the arrangement of rocks was a new, single object or the viewpoint that the arrangement of rocks did not form a new object—that it was just an arrangement of individual rocks—when presented with each variation of the scenario. There were two WF variations in this study, one in which the added function was accorded, and one in which the added function was designed. The accorded function variation stated that the man decided to leave the rocks on his lawn 'because the rocks made his lawn look beautiful,' while the designed function

variation stated that the man decided to arrange the rocks into a rock garden to make his lawn more beautiful. The results of this study revealed that, when presented with either WF scenario, the participants expressed greater agreeability towards the view that the rock arrangement was a distinct object than when the participants were presented with the NF scenario. Additionally, participants expressed greater agreeability towards the view that the rock arrangement was a distinct object in the designed function variation than the accorded function variation. The NF scenario had a mean agreeability score of 3.05 (SD=1.34), the accorded function scenario had a mean agreeability score of 5.05 (SD=1.77), and the designed function scenario had a mean agreeability score of 5.84 (SD=1.52). As was also observed in Studies 1 and 3, the participants' intuition concerning mereological composition, on average, became flipped when a function was added. (Rose and Schaffer 2017, 155-58).

Following the results of these studies, Rose and Schaffer (2017) stated that teleological thinking deeply pervades folk mereology. They did not, however, go as far as saying that teleology is the only factor that drives folk mereology, and instead adopted the condition of *teleologically restricted composition*, which states that teleology plays a significant role, but not the only role, in determining when the folk perceive a sum of parts as an object (Rose and Schaffer 2017, 160). Ultimately, due to experimentally demonstrating the influence teleology has on folk mereology, Rose and Schaffer claim that all answers to the SPQ can be liberated from having to accommodate folk intuition and cannot be debased because of it (2017, 175).

CRITICISM FOR ROSE AND SCHAFFER (2017)

There are several problems with how the studies in Rose and Schaffer (2017) were conducted, undermining the validity of the results and claims. Two clear problems in Rose and Schaffer (2017) is the use of guiding language and the inconsistent accounting of labeling. Rose and Schaffer only directly acknowledged the presence of a label in Study 3 (the "mini-bomb detector"), which had accommodations and further variations to account for the possible impact a label could have on the results (Rose and Schaffer 2017, 153-155). However, this was not done for Studies 1 and 2, despite the WF scenarios in both also featuring a label for the sum of parts (the statue "Unity," the "Gollywag-Supporter"). This inconsistency can be considered guiding language, language that pushes the

participants towards an answer preferred by the experimenters, and its presence discredits Studies 1 and 2, as it is possible the inclusion of labels skewed the results obtained in those studies. Another example of guiding language can be found in Study 4, with the unnecessary use of the phrase "compose a new object" in each of the WF variations. In the designed function variation, rather than describing the sum of parts with a sentence such as, "he has created an arrangement of rocks that makes his front lawn beautiful," which effectively conveys that the arrangement of rocks performs a function (make his lawn beautiful), Rose and Schaffer instead used the following sentence: "he thus thinks that the arrangement of rocks compose a new object, namely an object that makes his front lawn beautiful" (Rose and Schaffer 2017, 156). A variation of this sentence was also used in the accorded function variation, leading to an over-clarification of the added function in both WF variations that forcefully suggested to participants that a new object had been composed, rather than letting the participants' intuitions lead them to that conclusion organically.

Perhaps the most damaging criticism to the validity of the claims and results in Rose and Schaffer (2017) was featured in Korman and Carmichael (2017). Korman and Carmichael (2017) introduced the *creative intentions* account, which states that humans are more likely to view a sum as an object, if the sum is the "product of intentions" to make something of a specific kind (Korman and Carmichael 2017, 189-190). This account is clearly different than what was proposed in Rose and Schaffer (2017)—promiscuous teleology and teleologically restricted composition—and was a variable not considered in those studies, as in each of the WF scenarios in Rose and Schaffer (2017), creative intent was present: the politicians and sculptor intended to make a model in Study 1, the researcher intended to make a 'Gollywag-supporter' in Study 2, the researcher intended to make the "mini-bomb detector" in Study 3, and the homeowner intended to make a rock garden in Study 4. In conjunction, creative intent was also noticeably absent from each of the NF scenarios: every NF scenario was worded in a way that suggested the sum in question was formed randomly with no underlying intent, such as a description of the sum in Study 4 as "a bunch of rocks strewn across the lawn" (Rose and Schaffer 2017, 157). Due to the lack of creative intent in each of the NF scenarios and the inclusion of creative intent in each of the WF scenarios, it is possible that creative intent was a major, if not causal, factor for the large increase in agreeability observed between the NF and WF scenarios in

Rose and Schaffer (2017). As Rose and Schaffer did not manage to isolate function as the only changing variable between the NF and WF scenarios, the results and both claims made by Rose and Schaffer (2017)—the folk view mereology through *teleologically restricted composition* and answers to the SPQ can be liberated from having to accommodate folk intuitions—should be considered invalid until it can be experimentally supported that creative intent has an insignificant effect on folk mereology, in comparison to function.

Thus, to verify the results and claims in Rose and Schaffer (2017), I designed and conducted a study with similar NF and WF studies as those presented in Rose and Schaffer (2017), but with creative intent taken into account. Other criticisms to the validity of the Rose and Schaffer's results were also taken into account in the experimental design.

EXPERIMENTAL DESIGN

The experiment was a two-by-two study with four distinct scenarios. Participants for the study, recruited randomly by invitations at a university library, were undergraduate students at Northwestern University with little to no background in philosophy. The experiment was performed in-person with each participant filling out four questionnaires, one for each of the four distinct scenarios. The general structure of each scenario was consistent, as each scenario featured a 7-year old girl named Amy, who had stacked her stuffed animals together, forming a stuffed animal pile. The conditions regarding whether Amy intentionally formed the stuffed animal pile and whether the stuffed animal pile performed any purpose were varied across scenarios. Participants were given an unlimited time to read through each scenario, and after finishing the reading of each scenario, were asked to answer two true-false comprehension questions. While the same true-false comprehension questions were asked after each scenario, the correct answers to the comprehension questions differed amongst scenarios. Each scenario culminated in the asking of the following composition question, which revealed if the participant believed mereological composition had occurred: "Do you think that the stuffed animal pile forms a new, distinct object or is it just a stack of a number of distinct objects (stuffed animals)?" Before answering the composition question, participants were told ahead of time that they had five seconds to answer it. Since the experiment was performed in-person, the participants were reassured that the time limit was not implemented

to put them under duress, but rather to elicit their first response being the answer. At the very end of each scenario, participants were asked—on a scale of 1-7, with 7 being the most confident—how confident they were in their answer to the composition question.

The purpose of the experiment being delivered in person, as well as the inclusion of the two comprehension questions, was to ensure that the participants understood the conditions of each scenario before they answered the comprehension question. Participants were allowed to ask for elaboration on the comprehension questions, so long as it did not directly give away the answers to them. The inclusion of a time limit was to encourage participants to give their first answer, which reflects their initial mereological intuition, and the participants were reassured that the time limit was not to stress them out, so as to not add unintended extra conditions, such as cognitive duress. The comprehension question being a two-option question instead of a numerical Likert scale was to avoid inconsequential results, such as the '3.85 on a scale of 7' result observed in the results for the NF scenario in Study 2 of Rose and Schaffer (2017), which in Rose and Schaffer's own words, meant that the participant believed "neither was right" (Rose and Schaffer 2017, 178). The two-option question format can also more conclusively reveal a flip in intuition regarding mereology, compared to attempting to interpret a flip in results in a dependent study with numerical data. The confidence rating was implemented in order to help discern whether participants were just purely guessing in their responses or if they felt at least some confidence in how in their responses, which gives further insight into whether or not they were using their intuition to respond.

In Scenario A, the scenario conditions were NF and no-creative intent (NCI):

Scenario A

Amy is a 7-year-old girl who is filled with boundless energy and loves to play with her stuffed animals. One day, she randomly stacks all of her stuffed animals together, forming a stuffed animal pile. Later in the day, one of Amy's parents asks her what the purpose was for forming the stuffed animal pile, to which Amy just simply shrugs and says honestly, "No reason, and pile doesn't really do anything either."

> **Comprehension Check 1**: Amy stacks her stuffed animals together with the intent of making a stuffed animal pile.

A). True

B). False

Comprehension Check 2: Amy's stuffed animal pile performs a function.

A). True

B). False

****

Do you think that the stuffed animal pile forms a new, distinct object or is it just a stack of a number of distinct objects (stuffed animals)? (5 seconds to answer)

A). New, distinct object

B). Stack of single, distinct objects (stuffed animals)

As the questions, both the comprehensions questions and the composition question, asked in Scenarios B-D were the same as the ones asked in Scenario A, they were omitted from this paper, so as to not be repetitive. Participants, however, did physically see the questions for Scenarios B-D on the questionnaires they answered. In Scenario B, the scenario conditions were WF and NCI:

Scenario B

Amy is a 7-year-old girl who is filled with boundless energy and loves to play with her stuffed animals. One day, she randomly stacks all of her stuffed animals together, forming a stuffed animal pile. Later the next week, one of Amy's parents asks her what the purpose was for forming the stuffed animal pile, to which Amy

states honestly, "No reason, but I've found out that I feel safer when it's around, keeping my nightmares about the Boogeyman away."

In Scenario C, the scenario conditions were NF and with-creative intent (WCI):

Scenario C

Amy is a 7-year-old girl who is filled with boundless energy and loves to play with her stuffed animals. One day, she decides to stack all of her stuffed animals together. Upon doing so, Amy triumphantly declares, "The stuffed animal pile I've been trying to make is finally complete!" Later in the day, one of Amy's parents asks her what purpose the stuffed animal pile was for, to which Amy just simply shrugs and states honestly, "Nothing."

In Scenario D, the scenario conditions were WF and WCI:

Scenario D

Amy is a 7-year-old girl who is filled with boundless energy and loves to play with her stuffed animals. One day, she decides to stack all of her stuffed animals together. Upon doing so, Amy triumphantly declares, "The stuffed animal pile that I've been trying to make is finally complete!" Later in the day, one of Amy's parents asks her what purpose the stuffed animal pile was for, to which Amy states honestly that she's been having nightmares about the Boogeyman and that the stuffed animal pile has helped keep him away.

RESULTS AND DISCUSSION

Responses to the composition question were only recorded for participants who answered both comprehension questions correctly in each scenario. Only ten participants passed this screening, and due to the small sample size, the results of this study must be substantiated by a larger scale, identically-designed study for any concrete claims to be made. However, for the sake of interpretations, the following results and discussion section will hypothetically explore the ramifications that such results, if validated, would have on the folk mereology debate.

Scenario A had the highest amount of responses that said the stuffed animal pile was a sum of parts—90 percent—and the lowest amount of responses that said that the stuffed animal pile was a new, distinct object—10 percent. This result was expected, as Scenario A served as the control and provided a baseline scenario with neither creative intent or function; the responses in Scenario A could

be compared to responses in other scenarios to confirm if function and/or creative intent caused a change in the participants' agreeability towards the stuffed animal pile being a new, distinct object.

In both Scenarios B and C, which were the 'function alone' condition and the 'creative intent alone' condition respectively, the number of participants that believed the stuffed animal pile was a new, distinct object was greater than in Scenario A. While this provided evidence that both conditions have an impact on folk mereology, the addition of creative intent in Scenario C proved to have the greater effect. When function was the added condition in Scenario B, less than half of the participants—40 percent—viewed the stuffed animal pile as an object, compared to the majority of participants—60 percent—answering that the stuffed animal pile was an object when creative intent was the added condition in Scenario C. The results of these scenarios contradict the argument of *teleologically restricted composition* presented in Rose and Schaffer (2017), as the addition of creative intent was more influential on folk mereology than the addition of function. For *teleologically restricted composition* to be true, teleology must play a 'significant role' in determining mereological composition, where 'significant' is implied to be 'dominant' compared to other possible factors of influence. This condition was clearly not met by the results of Scenarios B and C. Since the claim of folk mereology conforming to *teleologically restricted composition* is synonymous with the claim that answers to the SPQ do not need to accommodate folk intuitions, both claims presented in Rose and Schaffer (2017) are severely damaged by the (hypothetical) findings in Scenarios B and C.

The findings in Scenario D provide further damage to the claims presented in Rose and Schaffer (2017). Scenario D was designed to observe if the positive effects on folk mereology—increased agreeability to the view that the sum is an object—due to the addition of creative intent and function would be additive. The structure of Scenario D mirrored the WF scenarios in Rose and Schaffer (2017), in order to also determine if the WF results obtained in Rose and Schaffer (2017) were due to the inclusion of both creative intent and function. The results of Scenario D revealed that the positive effects on folk mereology from both added conditions—function and creative intent—were indeed additive, as the vast majority of participants—90 percent—viewed the stuffed animal pile as a new, distinct object, which was a much greater percentage than those obtained in the scenarios that featured either added condition alone, Scenarios B and C.

These results were similar to the results obtained by the WF scenarios in Rose and Schaffer (2017), especially the results from the WF scenario in Study 2, as when presented with Scenario D, participants also overwhelmingly favored the view that the sum of parts was a new, distinct object. As the structure of Scenario D was identical to the WF scenarios in Rose and Schaffer (2017), but with the presence of creative intent acknowledged, this is evidence that the huge flips in intuition observed between the NF and WF studies in Rose and Schaffer (2017) were due to the combined effects of both added creative intent and function, contrary to what Rose and Schaffer claimed, which was that the addition of function alone led to the intuition flips.

The results of these scenarios support the claim that folk mereology is driven by *creative intent restricted composition*, rather than *teleologically restricted composition* and discredits the results obtained in Rose and Schaffer (2017). However, the antithesis to the conclusions drawn in Rose and Schaffer (2017), the conclusion that folk mereology can be used to reject answers to the SPQ, cannot be made. Further work exploring whether or not creative intent is flawed and unscientific when implemented in spheres outside of the artifact realm, like how teleological thinking is flawed and unscientific beyond the artifact realm, must be conducted before such a conclusion can be made.

CONCLUSION

Folk mereology, according to Rose and Schaffer (2017), is driven by teleological thinking, and thus, cannot be used as a reason for rejecting answers to the SPQ. Incorporating criticism from Korman and Carmichael (2017) and taking into account other flaws present in the experimental design of Rose and Schaffer (2017), the study proposed and carried out in this paper found results that contradicted the claims and conclusions presented in Rose and Schaffer (2017). It has been discussed that, should the results of these studies been substantiated by an identically designed, larger scale study , Rose and Schaffer's claims about folk mereology being teleologically driven would be disproven: instead of *teleologically restricted composition* driving folk mereology, *creative intent restricted composition* is what drives folk mereology. The results of this paper warrant further work to be done in the form of a similarly designed large-scale study with proper error analysis techniques, in order to add validity to the findings of this paper. Human intuition on what constitutes an object should not

be discounted in metaphysical debates on the basis of teleological thinking, as further understanding and much work is needed before such conclusions can be made.

REFERENCES

Allen, Collin, and Marc Bekoff. 1995. "Function, Natural Design, and Animal Behavior: Philosophical and Ethological Consideration." In *Perspectives in Ethology, Vol. 11" Behavioral Design*, edited by Nicholas S. Thompson, 1-47. New York: Plenum Press.

Dink, Jacob W., and Lance J. Rips. 2017. "Folk Teleology and its Implications." In *Experimental Metaphysics*, edited by David Rose, 207-236. London: Bloomsbury Press.

Heider, Fritz, and Marianne Simmel. 1944. "An Experimental Study of Apparent Behavior." *The American Journal of Psychology* 57 (2): 243-259.

Kelemen, Deborah. 1999. "The scope of teleological thinking in preschool children." *Cognition* 70 (3): 241-272.

Kelemen, Deborah, Joshua Rottman, and Rebecca Seston. 2013. "Professional physical scientists display tenacious teleological tendencies: purpose-based reasoning as a cognitive default." *Journal of Experimental Psychology: General* 142 (4): 1074-1083.

Korman, Daniel Z., and Chad Carmichael. 2017. " What Do the Folk Think about Composition Does it Matter?" In *Experimental Metaphysics*, edited by David Rose, 187-206. London: Bloomsbury Press.

Lombrozo, Tania, Deborah Kelemen, and Deborah Zaitchik. 2007. "Inferring design: evidence of a preference for teleological explanations in patients with Alzheimer's disease." *Psychological Science* 18 (11): 999-1006.

Rose, David, and Jonathan Schaffer. 2017. "Folk Mereology is Teleological." In *Experimental Metaphysics*, edited by David Rose, 135-186. London: Bloomsbury Press.

van Inwagen, Peter. 1990. *Material Beings.* Ithaca: Cornell University Press.

cognethic.org