8-11-2022 1:00 PM

Culture and Classification: Investigating Analytic vs. Holistic Thinking Styles

Neha Khemani, *The University of Western Ontario*

Supervisor: Minda, John Paul, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Psychology

Abstract

This paper sought to explore cultural preferences for analytic and holistic thinking in classification. Experiment 1 paired the Shepard, Hovland, and Jenkins (SHJ) tasks with the Analysis-Holism scale (AHS) and a demographics questionnaire. Effects of culture on learning rates, alongside the feasibility of online data collection, were assessed. Learning difficulty differences among the six SHJ category sets were observed. Further, as predicted, higher holistic thinking correlated positively with the family resemblance task. Experiment 2 replicated the Norenzayan et al. (2002) task. Unlike in the original study, the effect of instructional condition was not significant across our full sample. Nevertheless, the non-Western sample showed higher holistic thinking in the similarity instruction condition. Moreover, our sample did not show any overwhelming preference for either analytic or holistic thinking strategies. Overall, our results are inconclusive, yet promising, and hint at some effect of culture on classification. This warrants further research in this domain.

Keywords

Classification, Category Learning, Categorization, COVIS theory, Analytic Thinking, Holistic Thinking, AHS, SHJ, Culture, Online Data Collection

Summary for Lay Audience

This paper sought to explore cultural preferences for analytic and holistic thinking in classification. The basic mechanisms of category learning are thought to be universal (Shepard, 1987), however, recent research has found that individuals from Eastern cultures tend to have a holistic processing style, whereas those in Western countries default to an analytic one (Nisbett et al., 2001; Norenzayan et al., 2002). Nevertheless, these results have not always been replicated (Murphy et al., 2017). The broad classification realm primarily includes category learning and categorization paradigms, with two main strategies widely debated: family resemblance and rule-based strategies. The family resemblance strategy uses a judgment of overall similarity, whereas rule-based responding is logical and based on a focal object with necessary and sufficient features. Experiment 1 paired the six classification sets first described by Shepard, Hovland, and Jenkins (SHJ; Shepard et al., 1961) with the Analysis-Holism scale (AHS; Choi et al., 2007) and a demographics questionnaire on an entirely online platform. These six category sets test the reliance on single feature rules, disjunctive rules, and family resemblance. Alongside testing the feasibility of online category learning data collection, the SHJ tasks were used to explore the effects of culture on the learning performance of the different category sets. Learning trends in the expected direction were observed, indicating variable learning difficulty across the SHJ sets. Also, a relationship between family resemblance responding and high holism on the AHS was observed, together with other interesting exploratory findings. Experiment 2 replicated the Norenzayan et al. (2002) task, which tested two instructional conditions. The similarity judgment condition was predicted to elicit higher family resemblance responses. Unlike the original study, our results did not show a significant effect of instructional condition. Further, only the non-Western sample showed a significant effect of condition, depicting higher holistic thinking in the similarity judgment condition as opposed to the classification condition. Moreover, neither thinking style was dominant across our sample. Overall, our results hint at cultural preferences, though the size and direction of the effect remain inconclusive. The findings motivate additional exploration of this domain by addressing limitations and future directions.

Acknowledgements

First and foremost, I would like to especially thank my supervisor, Dr. John Paul Minda. Thank you for your guidance, expertise, encouragement, and support throughout the course of this project. I would also like to thank members of my lab, the Categorization Lab, for listening, asking questions, offering feedback, and for your constant willingness to help. Further thanks to Ana Ruiz Pardo for the ongoing mentorship and assistance, and for being everyone's personal notetaker. Your effort put towards getting this project off the ground alongside hours of coding help have been so incredibly helpful. As well, thank you to Anthony Cruz for all the productive work sessions, feedback, and graphing help.

To all my friends in my cohort, thank you for all the friendship, patience, and advice. As well, thank you to all my friends at home for being my personal cheerleaders. Lastly, my most sincere gratitude is owed to my wonderful family – my parents and my brother. Your curiosity and willingness to learn more about my field of study has not gone unnoticed. Thank you for your unconditional love and support throughout my educational journey and for always believing in me. I appreciate your constant enthusiasm and encouragement towards all my endeavours.

Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

1 Introduction

1.1 Concepts And Categories

Category learning and categorization processes are essential to human and animal cognition, impacting everyday activities. Identifying an unknown animal, vegetable, or caring for a house plant all require classification decisions. Knowledge of categories is crucial to learning, remembering, and integrating novel information to ultimately inform decision making. As a result, category learning plays a central role in information processing as it promotes the expansion of existing insight to novel situations. Being a complex, yet fundamental process to making sense of the environment, a myriad of factors influence the way in which humans assign objects to categories. Many of these factors remain unexplored, consequently constricting the understanding of category learning and categorization mechanisms.

1.1.1 Category Learning and Categorization

Category learning is the process of relying on memory to improve the efficiency of assigning novel objects to appropriate groups (Ell & Zilioli, 2012). Through this, other cognitive processes that require category membership knowledge are enabled. A category is formed by a collection of related objects, though the specific object relationships can differ vastly across groups. Both in the real world and in a lab setting, categories are learned by observing novel objects, classifying these objects, and receiving feedback confirming or denying the proposed category membership. Thus, feedback is a core element of the category learning process. Another key component is categorization, which is the assignment of objects to specific groups. This action can occur with or without an attached learning objective, such that feedback may or may not be provided. Categorization studies without feedback focus on grouping strategies and object-group relationships. This is indicated by investigating the style of thinking seemingly used to group objects.

1.1.2 Theories of Category Learning

Through decades of research, various models of category learning have been proposed. Since the classical view by Smith and Medin (1981), the prototype model by Rosch (1973), Posner and Keele (1968), and Homa et al. (1973), there has been a shift from two competing single-system models of classification (Ashby et al., 2011). Research by Erickson and Krusche (1998) and Ashby et al. (1998) showed the presence of both rule induction and exemplar encoding strategies in classification. Their findings suggested that multiple category learning systems may be mediating human categorization, giving rise to the current multiple-systems models of classification. These models gained popularity because the single-system theories were unable to account for findings of multiple systems.

There is support for the notion that categorization can be based on both verbally described information and information that cannot be described verbally (Smith & Grossman, 2008; Ward & Scott, 1987; Allen & Brooks, 1991; Kemler Nelson, 1984; Zeithamova & Maddox, 2006). Further, evidence suggests that although similarity can be a driving factor for categorization, it is not always solely sufficient. This was seen in a study by Rips (1989) where classification was found to be influenced by factors other than similarity. Accordingly, it can be drawn that specific classification decisions can be motivated over others based on the instructions provided. Aligning with these findings, a multiple-systems approach of classification has been proposed (Newell et al., 2011).

Two distinct category learning mechanisms have been defined within a multiple-systems framework. The first is a rule-based, verbal process. Using this mechanism, category membership is determined by learning rules to identify necessary and sufficient features among group members. The second is a family-resemblance process in which learning is based on overall similarity. Using this process, category exemplars display a general similarity and share multiple features with one another, without a single feature that is necessary or sufficient for category membership (Rosch & Mervis, 1975). The key distinction from the rule-based system being that judgments are based on overall similarity, such as in a family, rather than an explicit rule. Also, unlike the rule-based system, the family-resemblance process is often not verbalizable.

Extending the notion of verbalizable and non-verbalizable processes, the dual-systems account highlights two independent systems of category learning: verbal and non-verbal systems. The verbal system is known as an explicit, rule-based, and declarative system that relies on the application of rules, ergo requiring working memory and attention (Minda & Miles, 2010). In contrast, the non-verbal system is an implicit, similarity-based, and procedural system that does not involve working memory or attention and is likely an unconscious process (Minda & Miles, 2010). There is reason to believe that adults can access and utilize both systems when making classification decisions (Allen & Brooks, 1991). While there exists research deciphering the precise contexts and features that mediate the relationship between the two systems, there is still much unknown in this domain. This is evident through the proposition and testing of numerous theories and models over time, most of which remain inconclusive.

1.1.3 COVIS

A prominent multiple-systems theory of category learning is the COVIS (Competition between Verbal and Implicit Systems) theory (Ashby & Ell, 2001; Ashby & Maddox, 2011). This model emphasizes two fundamental competing learning systems: the verbal system and the implicit system (Ashby & Maddox, 2011). The verbal system, also known as the explicit system, uses cognitive resources to search for, store, and apply a rule to determine category membership (Zeithamova & Maddox, 2006). This learning of explicit rules requires sufficient resources in the form of working memory, attention, and other executive functioning. The verbal system follows an active, deliberate process of hypothesis testing to devise, test, and revise the rules necessary for category membership in a strategic and goal-derived manner.

In contrast, the implicit system is procedurally based and learns non-rule-described categories using non-verbalizable, gradual, and automatic judgments of family resemblance (Ashby & Maddox, 2011). Given the lack of a verbalizable rule, non-rule-described categories can only be learned using similarity or by integrating stimulus features at a pre-decisional stage (Ashby & Ell, 2001). Like incidental learning, this system does not learn deliberately and operates outside of conscious awareness. As a result, the implicit system promotes learning more complex categories for which no verbalizable rule exists.

When learning a novel category, the COVIS model posits that both systems function simultaneously, with the more accurate system becoming the dominant strategy. However, there is a bias towards the verbal system as it is the default approach for normally functioning adults (Ashby & Maddox, 2011; Minda et al., 2008). This bias stems from the competing nature of the two systems and exists because the implicit system learns gradually over multiple exemplars, ensuring automaticity with adequate experience. This is different from the explicit system through which rules can be devised and tested immediately. Consequently, adults learning new categories will initially default to testing hypotheses and only rely on the slower, implicit system if these hypotheses are unsuccessful (Ashby & Maddox, 2011).

Feedback is a core element of category learning. Feedback upon correct categorization acts as a reward, causing the release of dopamine, which, in turn, strengthens the association between the stimulus and the categorization response (Wickens, 1990). To accurately learn a category over several exemplars, the implicit system relies on clear, timely feedback. This is because the stimulus-response connections must still be active for the association strengthening to occur (Wickens, 1990). Thus, feedback immediately following categorization is required for implicit system learning using automatic synaptic strengthening (Minda & Miles, 2010). In contrast, the timing of feedback for the verbal system is less crucial because the rule is stored in working memory until effectiveness can be evaluated (Minda & Miles, 2010). Hence, delayed feedback hinders learning when using the implicit system but not when using the explicit system. However, while the implicit system processes feedback automatically, the verbal system demands working memory, attention, and time to process the feedback. This further affirms the need for cognitive resources when using the verbal system.

In sum, the COVIS multiple-systems model of learning outlines two separate systems that operate simultaneously in competition with one another. The verbal system is the default, primary system and requires sufficient cognitive resources to learn properly. On the other hand, when the hypotheses from the verbal system are proven incorrect, the more gradual implicit system tends to take over. In this way, the dominant system depends on the category, stimuli, and task specifics such that the most accurate system leads category decisions.

1.1.4 Analytic vs Holistic Processing

On a similar note, there is evidence suggesting that category learning and categorization may be guided by two computationally distinct but coexisting cognitive strategies (Erickson & Kruschke, 1998; Folstein & Van Petten, 2004; Kemler Nelson, 1984; Nosofsky et al., 1994). Related to the dual-systems approach, this account discusses rule-based and family resemblance strategies as broader styles of processing, potentially influenced by a multitude of internal and external factors. The two distinct processing styles, analytic and holistic thinking, maintain some overlap with the verbal and non-verbal strategies.

The analytic strategy is a rule-based, formal strategy that involves decoupling an object from its surrounding context to attend to the focal object alone (Nisbett et al., 2001). Being a rule-based strategy, analytic thinking overlaps with the verbal system in that category membership is determined by rules defined by necessary and sufficient features (Nisbett et al., 2001). Applying rules requires strategic and goal directed thinking, leading analytic thinkers to deliberately test hypotheses (Minda & Miles, 2010). Accordingly, analytic thinking has been associated with intentional learning, which occurs when the goals of a task are known and makes category rules easier to learn (Kemler Nelson, 1984). This is different from a holistic reasoning style. Holistic thinking, also known as intuitive, or similarity-based classification, processes the field or context in its entirety, viewing the composition as greater than the sum of its parts (Choi et al., 2007). This view considers the relationships among objects, as well as object-field relationships, hence being a field-dependent strategy (Nisbett et al., 2001). Category membership when thinking holistically is determined using overall similarity or family resemblance judgments, much like the non-verbal system (Nisbett et al., 2001). Although aspects of family resemblance may be verbalizable, verbal abilities are not necessarily engaged when processing holistically (Minda & Miles, 2010). In contrast to the analytic style, this strategy has been linked to incidental learning, in which the task goals are not explicitly recognized, and learning is unknown to awareness. While still somewhat debated, the relative influence of these two strategies over the other may depend on stimulus structure, such as object features, and/or on the task, including task instructions, the nature of the task, or other task-related factors (Minda & Miles, 2010; Morgan & Johansen, 2020).

1.2 Culture

The multiple-systems approach of category learning and categorization is one that has been extensively studied, with a large majority of the field operating from the underlying perspective that the ability to learn and use categories is a universal aspect of biological intelligence. Category learning research has found the presence of basic category learning mechanisms across different species (Shepard, 1987), strengthening the belief that learning mechanisms are universal (Smith et al., 2012). While elements of universality exist when learning and using categories, people experience vastly different circumstances and surroundings, suggesting that human concepts may not be alike. In addition to differing cultural and social experiences, humans use language, which varies by culture, to mark behavioural equivalence classes. Language and culture are inextricably linked and can differ considerably, presenting as important factors to consider in category learning and categorization research. The assumption of ubiquitous classification is, in part, a consequence of deficient insight into category learning among diverse cultures and languages (Henrich et al., 2010). Prominent literature has focused on the narrow pool of primarily Western, Educated, Industrialized, Rich, and Democratic (WEIRD) populations (Henrich et al., 2010). Hence, classification differences that may stem from, or be related to language and/or culture have not been identified. While people across all cultures possess these two reasoning systems, the differential value placed on each system and subsequent accessibility of one system over the other may be influenced by cultural variation (Norenzayan et al., 2002).

Given that very little is known in the realm of classification and cultural variation, recent years have shown an increase in cross-cultural research contrasting categorization strategies among WEIRD and non-WEIRD populations. Findings from this research have indicated the presence of cultural variation in how categories are learned and used (Henrich et al., 2010; Norenzayan et al., 2002; Nisbett & Miyamoto, 2005; Masuda & Nisbett, 2001). In a study by Masuda & Nisbett (2001), recognition accuracy rates of Japanese participants varied significantly as a function of stimuli backgrounds, whereas American participants were unaffected by backgrounds. This supported the view that Westerners tend to focus on the focal object rather than consider object-context relationships, such as in analytic thinking. Moreover, research by Norenzayan et al. (2002) found that when making similarity-based judgments, East Asians preferred family resemblance strategies, whereas European Americans used a unidimensional rule instead. This research, along with other findings by Norenzayan and colleagues (Masuda & Nisbett, 2001; Nisbett et al., 2001; Norenzayan et al., 2002; Norenzayan & Nisbett, 2000) suggest that an influence of culture is present among category learning and categorization strategies when a culturally diverse population is studied.

More specifically, it has been observed that Eastern cultures default to a holistic thinking style and Westerners default to an analytic one (Masuda & Nisbett, 2001; Nisbett et al., 2001; Norenzayan et al., 2002; Norenzayan & Nisbett, 2000). This suggests that Western thinking is governed by rules rather than similarity and is more verbal and context independent. Accordingly, Westerners are thought to favour hypothesis testing when categorizing. On the other hand, East Asians are thought to rely on more implicit, non-verbal classification that uses a family resemblance strategy to group novel items. However, it is important to note that despite reliable evidence for cultural variation in this direction, Westerners have not always shown a preference for analytic thinking. This was observed in a study carried out by Murphy et al. (2017) modeled after Norenzayan et al. (2002), which failed to replicate key results found in the original study. This goes to show that cultural variability in the ability to learn and use categories is still a relatively unexplored and inconclusive area, prompting additional cross-cultural classification research.

1.2.1 Possible Origins of Cultural Differences in Reasoning

Recent cross-cultural research has suggested that causal reasoning differences may be a product of divergent ways of organizing the world. Though a detailed account of the proposed origins of these distinct worldviews are beyond the scope of this paper, a brief discussion is warranted for context. Various theories have been proposed explaining this, however it is important to note that these are primarily speculative in nature. One such theory claims that cross-cultural differences in thinking may have philosophical origins, particularly relating back to the ancient Chinese and ancient Greeks. As distant societies, the Greek civilization gave rise to European and American civilizations while the Chinese civilization gave rise to South and East Asian civilizations. Philosophical traditions from these cultures are thought to have influenced intellectual and social practices in the West and East, primarily observed by differential principles of philosophy, mathematics, and science (Norenzayan et al., 2002; Nisbett et al., 2001).

Broadly, ancient Greek culture emphasized mathematical and computational approaches, generating rules to understand and explain the nature of objects (Nisbett et al., 2001). This led to the invention of the fields of physics, astronomy, and math - specifically geometry, where formal logic and proofs were relied on to understand underlying physical causes of objects and their behaviour (Nisbett et al., 2001). In this way, the emphasis on scientific theory and investigation among the ancient Greeks was important for the development of science based on formal logic. The reliance of formal logic supports analytic thinking in that universal rules are examined by focusing on the object itself and asserting precision when categorizing objects. On the contrary, the ancient Chinese relied on intuition, resulting in a skeptical attitude towards formal logic (Lloyd, 1990). Thus, although advancing technology was a large goal for the ancient Chinese, the driving force behind it was practicality rather than scientific theory and investigation (Nisbett et al., 2001). Additionally, the ancient Chinese were more theoretical and traditional, which led to the development of a dialectical instead of formal logic, which allows for and accepts contradictions (Lloyd, 1990). This aligns with the intuitive, less rule-based methods and attitudes of seeking a middle ground seen among holistic thinkers. Furthermore, Nisbett et al. (2001) discuss the ancient Greek encouragement for personal agency in the form of debates and individual beliefs. This was different from the collective agency fostered by the ancient Chinese, enforcing expectations of the group to guide individual behaviour (Nisbett et al., 2001). Such weight on group harmony among the ancient Chinese tracks the belief of internal relatedness of all things, which is another core premise of holistic thinking. This is also seen in traditional Chinese medicine approaches that aim to balance the flow of natural forces throughout the body (Hadingham, 1994). In complete contrast, surgery has been a longstanding key component of Western medicine, which is a method that focuses on healing a specific part of the body in isolation (Nisbett et al., 2001). These practices, approaches, and beliefs, dating back to ancient civilizations, are further discussed by Nisbett et al. (2001) and provide a basis for understanding philosophically originating cross-cultural differences in thinking and reasoning.

In addition to philosophical influences, social orientation may contribute to the culturally driven different styles of thinking. The impact of social orientation on causal reasoning centers around the idea that reasoning among Eastern cultures focuses on external causes and Westerners on internal dispositions (Nisbett et al., 2001). More broadly, collectivist cultures focus on situational and environmental influences when interacting with an object, whereas individualistic cultures look at attributes of the object itself (Nisbett et al., 2001; Choi et al., 1999; Morris & Peng, 1994). Eastern cultures follow a collectivist orientation, emphasizing group harmony and interdependence (Spencer-Rodgers & Peng, 2018; Triandis, 1989). In contrast, Westerners are considered individualistic, emphasizing personal fulfilment and autonomy in line with an independent outlook. Related to this, Easterners account for multiple interactive components beyond the object itself and believe that "everything affects everything else" (Choi et al., 1999). The distinction between internal and external causes of reasoning has been found in studies of attribution in the social domain. Chiu (1972) observed that Chinese participants were more sensitive to their environment, thus being situation-centered rather than individual-centered. Support for this was also found in a study by Ji et al. (2000) that tested American and East Asian participants on the Rod and Frame Test. The study found a higher field or context dependence among East Asians, while Americans showed more attention to the object and its relation to the self than the field. Although not fully understood yet, it is believed that the basis of these differences is rooted in a more abstract understanding of causal structure (Nisbett et al., 2001). Having a more situational view as Easterners do, it can be inferred that contextual information is of greater value when organizing the environment since it informs their understanding of external relationships and causes. In this way, it can be expected that Easterners are context, or field-dependent when attending to aspects in the environment rather than solely focusing on the object. This is different for Westerners, as they would direct their attention towards the focal object only, deeming the surrounding context as irrelevant.

Additional biases are thought to stem from these differing ways of organizing the world, showing varying levels of influence cross-culturally. The tendency for collectivist cultures to focus on situational factors when explaining behaviour gives rise to the hindsight bias, which is believed to be higher among Easterners (Nisbett et al., 2001). This has been seen in studies showing that Americans attribute behaviour largely to internal factors, such as personality traits, while Easterners focus on situational ones (Miller, 1984; Morris & Peng, 1994; Norenzayan, 1999; Masuda & Kitayama, 2004). Taken one step further, it has been argued that the fundamental attribution error is weaker among Easterners compared to Westerners (Ross, 1977; Choi et al., 1999). The fundamental attribution error, also known as the correspondence bias, is the tendency to prefer dispositional explanations of behavior over situational ones (Gilbert & Malone, 1995). In a study by Morris and Peng (1994), Chinese and American university students were asked to explain two tragedies that had recently occurred in the United States. The findings of this study showed that Chinese students preferred contextual explanations, but American students preferred dispositional ones. Choi and Markus (1998) saw similar causal attribution tendencies in a conceptual replication of this study among Korean and American participants. These findings support the differential focus on context versus objects that is believed to exist between Easterners and Westerners.

1.3 The Present Research

Cross-cultural research contrasting categorization strategies among WEIRD and non-WEIRD populations has observed cultural variation in how categories are learned and used. From the research and available evidence in this domain, there are many gaps to be filled. While it has been strongly proposed that a culture-dependent relationship between analytic and holistic thinking may exist, these claims appear to be preliminary in nature. To uncover the extent of research exploring thinking styles and culture, a scoping review was conducted. Though a full analysis is beyond the scope of this paper, some key findings are noteworthy. The aim of the review was to realize the state of the literature exploring categorization, attention, category learning, and culture to identify any relevant gaps and trends. The results indicated that higher holistic thinking was seen among non-WEIRD samples in most experiments, reflecting an effect of culture and thinking styles. Nevertheless, it was concluded that cross-cultural classification research is greatly unexplored. Thus, the review further supported the need for additional exploration of culture and classification.

The projects discussed in this paper seek to explore the possibility of cognitive universals in classification, which refers to principles and cognitive capabilities engaged in classification that are common to all intelligent life forms. A secondary aim is to investigate potential mediating or moderating variables. The realm of classification includes both category learning and categorization paradigms. In general, culture differences in simple categorization are more extensively studied than in other classification paradigms. This motivated the exploration of both category learning and categorization to gather broader insight on classification in its entirety. In landing on specific paradigms, it was critical to include ones with a high reliability of eliciting differences in learning and classifying strategies outside of cultural influences. The rationale for this was (1) to ensure that any differences in thinking styles will be reliably observed and (2) to limit the number of confounding factors when trying to study culture and thinking style differences. In line with this, methods common among North American and European samples were used, most of which were less common in other cultures given the lack of research among non-WEIRD samples. Also, the recent growth of sophisticated online data collection platforms encouraged effective testing of a more global sample. Thus, all studies collected cross-cultural data online. It is known that culture is a complex area of study with several interconnected components. This stressed the importance of narrowing the focus when diving into analytic versus holistic thinking, leading to a wider definition of culture as opposed to making language a core focus.

The overarching research question was to understand whether category learning and categorization strategies are universal across culture groups. A secondary goal stemming from this was to investigate potential predictors bringing about such an effect, whether via moderation or mediation. To tackle the category learning component, the six classification sets first used by Shepard, Hovland, and Jenkins (SHJ; Shepard et al., 1961) were coupled with the Analysis Holism scale (AHS; Choi et al., 2007) and a simple demographics questionnaire. These category sets have shown replicable rank order effects of learning rates, with visible distinctions between the easiest and most difficult sets (Smith et al., 2004). To investigate categorization, the commonly used paradigm by Norenzayan et al. (2002) was coupled with the AHS, a demographics questionnaire, the Autism Spectrum Quotient-10 (AQ-10; Baron-Cohen et al., 2001), and the Big-Five-Inventory-2-Extra-Short questionnaire (BFI-2-XS; Soto & John, 2017a). The categorization paradigm by Norenzayan et al. (2002) has been used by several researchers investigating culture and categorization; however, the results have not always replicated. This made it an interesting method to use with the underlying goal of assessing replicability. The addition of other scales to both experiments was exploratory and novel, in efforts to take a first stab at the secondary aim. Generally, we hypothesized that across both studies, East Asian participants would default to a more holistic thinking style, whereas Westerners would show an analytic one.

Chapter 2

2 Experiment 1

2.1 Introduction

Within the realm of classification, effects of culture in category learning are the least explored[1]. Current literature suggests that an effect of culture can be expected for categorization and attention studies, but not necessarily for category learning paradigms. This is because there is a higher level of uncertainty surrounding culture and category learning. With the aim of diving into this relatively unexplored area, this study served to observe processing style differences and investigate whether these differences are driven largely by culture. Another aim of this study was to examine the possibility of measuring these effects using an online task, with the objective to eventually scale the project globally.

The present study intended to investigate the effect of task demands on category learning strategies. A category learning task consisting of six category sets first described by Shepard, Hovland, and Jenkins—referred to as "SHJ" in this paper was used (Shepard et al., 1961). These SHJ category sets test the reliance on single feature rules, disjunctive rules, and family resemblance strategies. These tasks have been used to explore learning in children, adults, older adults, and even non-human primates, indicating high task reliability (Lewandowsky, 2011; Minda et al., 2008; Rabi & Minda, 2016; Smith et al., 2004). The category sets vary in both complexity and optimal strategy, such that single feature rules should be learned the easiest, multi-feature rules second, family resemblance and nonlinear rules next, and incoherent categories should be the most difficult to learn. In addition to the six SHJ category sets, the Analysis-Holism Scale (AHS) by Choi et al. (2007) was administered. This 24-item scale was used to measure the four major components of analytic-holistic thinking: causality, attitude towards contradictions, perception of change, and locus of attention. While this is a

---

[1] This finding is from a scoping review that was conducted to gauge the breadth of research on classification and culture. Though the review is recent (last search conducted in 2021), it has not been published.

novel scale, it is broadly believed that these four components guide one's overall style of thinking, such that East Asians score higher on the AHS (higher holistic thinking) and Westerners score lower on the AHS (higher analytic thinking) (Choi et al., 2007).

To explore category learning, thinking styles, and the effects of culture across populations, this study paired the SHJ tasks with the AHS and a demographics questionnaire in an entirely online platform. Using this task, the goal was to explore the effects of culture on participants' learning rates. Based on the findings by Shepard et al. (1961), it was hypothesized that performance differences will be observed such that performance will decrease with increasing category task difficulty. Thus, as the optimal strategy to categorize progresses in difficulty from a single feature rule to an XOR rule to a family resemblance strategy and more complex rules, it is predicted that overall performance should decrease. Accordingly, Type I (unidimensional) is expected to be the easiest task to complete, followed by Type II, then any of Types III, IV and V (predicted to be of equal difficulty), and lastly, Type VI (see Table 1). Type VI is expected to be the most difficult and should show the lowest performance. It is also hypothesized that AHS scores will predict performance on the SHJ category learning tasks. The expectation is that trends will be in accordance with findings of Eastern cultures showing a holistic processing style and Western cultures defaulting to an analytic one (Nisbett et al., 2001; Norenzayan et al., 2002). Correspondingly, no differences among culture groups are expected when categories can be learned by a single strategy. However, when concepts can be learned in more than one way, subtle preferences for holistic approaches in non-Western samples are anticipated. That is, more holistic thinking, seen by high AHS scores, should be related to better performance on the family resemblance task (i.e., Type IV), whereas analytic thinking, seen by low AHS scores, should be related to better performance on rule-based tasks (i.e., Type II).

Table 1: Task, Type, sub-Type, and stimulus number by category set.

Task	Type	Subtype	Category A Stimuli	Category B Stimuli	Task	Type	Subtype	Category A Stimuli	Category B Stimuli
1	I	1	1 2 3 4	5 6 7 8	19	IV	1	1 2 4 6	3 5 7 8
2	I	2	1 2 5 6	3 4 7 8	20	IV	2	1 3 4 7	2 5 6 8
3	I	3	1 3 5 7	2 4 6 8	21	IV	3	2 3 4 8	1 5 6 7
4	I	4	5 6 7 8	1 2 3 4	22	IV	4	3 5 7 8	1 2 4 6
5	I	5	3 4 7 8	1 2 5 6	23	IV	5	2 5 6 8	1 3 4 7
6	I	6	2 4 6 8	1 3 5 7	24	IV	6	1 5 6 7	2 3 4 8
7	II	1	1 2 7 8	3 4 5 6	25	V	1	1 2 3 8	4 5 6 7
8	II	2	1 3 6 8	2 4 5 7	26	V	2	1 2 4 7	3 5 6 8
9	II	3	1 4 5 8	2 3 6 7	27	V	3	2 5 7 8	1 3 4 6
10	II	4	3 4 5 6	1 2 7 8	28	V	4	4 5 6 7	1 2 3 8
11	II	5	2 4 5 7	1 3 6 8	29	V	5	3 5 6 8	1 2 4 7
12	II	6	2 3 6 7	1 4 5 8	30	V	6	1 3 4 6	2 5 7 8
13	III	1	1 2 3 6	4 5 7 8	31	VI	1	1 4 6 7	2 3 5 8
14	III	2	2 5 6 7	1 3 4 8	32	VI	2	2 3 5 8	1 4 6 7
15	III	3	2 3 4 7	1 5 6 8	33	VI	3	1 4 6 7	2 3 5 8
16	III	4	4 5 7 8	1 2 3 6	34	VI	4	2 3 5 8	1 4 6 7
17	III	5	1 3 4 8	2 5 6 7	35	VI	5	1 4 6 7	2 3 5 8
18	III	6	1 5 6 8	2 3 4 7	36	VI	6	2 3 5 8	1 4 6 7

2.2 Methods

2.2.1 Participants

A total of 823 individuals participated in the study, of which 200 were recruited using Prolific and 623 were undergraduate students at Western University. The Western University undergraduate sample had 487 participants complete the study online and 136 complete the study in-person. All participants were randomly drawn from their respective subject pool. Undergraduate students were compensated with course credit in an undergraduate psychology course and Prolific participants were compensated with 3.75 GBP.

The final sample consisted of 702 adults (*M* = 20.55, *SD* = 5.81, *range* = 17.00 - 62.00) after removing those with technical errors, incomplete responses, mis-matched IDs, and failed attention checks. The final sample had a total of 250 males, 446 females, four self-identifying individuals, and two that preferred not to say. Age, English proficiency scores, and AHS scores of the overall sample are depicted in Table 2.

Table 2: Descriptive Statistics

	N	***M***	***SD***	***Min***	***Max***
Age	702	20.55	5.81	17	62
English Proficiency	702	2.87	0.35	1.00	3.00
AHS Score	702	4.87	0.49	2.75	6.54

Most participants stated English (*n*=427) as their first language. Other languages recorded were Mandarin (*n*=45), Spanish (*n*=40), Portuguese (*n*=22), French (*n*=7), and Arabic (*n*=9). An additional 148 individuals stated 'Other' and four preferred not to say. Most participants were living in North America (*n*=512), with the rest in Europe (*n*=119), Africa (*n*=29), Asia (*n*=13), Central America (*n*=5), South America (*n*=5), Caribbean Islands (*n*=2), or Other (*n*=15). Two participants preferred not to state their region. Figure 1 shows the spread of participants by region across the world. This plot was created using the location coordinates in which participants completed the study. Note that though a spread across Europe can be seen, the highest concentration of points was in Canada, likely due to the large Western University undergraduate participant pool.

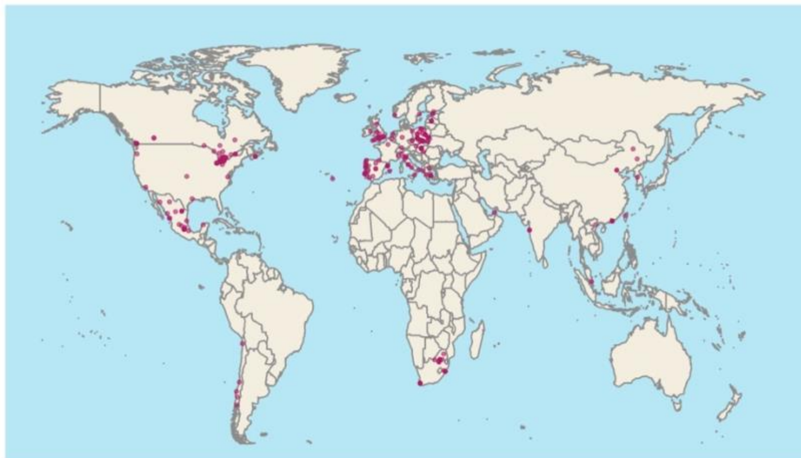


Figure 1: Participant spread by region. Each point represents a participant location.

2.2.2 Materials

Stimuli. The stimuli were shapes that varied on three binary dimensions of shape (square or triangle), size (large or small), and color (white or black) (Figure 2). With three dimensions, a total of eight stimuli were created. To clearly differentiate along the dimension of size, the large shapes were about twice as big as the small ones (Figure 2). All shapes were presented on a grey background to preserve the dimension of color. The stimuli were sorted into six Types of category sets based on the six classification sets first described by Shepard et al. (1961; SHJ category sets). Each category set had eight stimuli and two categories, for a total of four stimuli in each category (Category 1 versus Category 2).

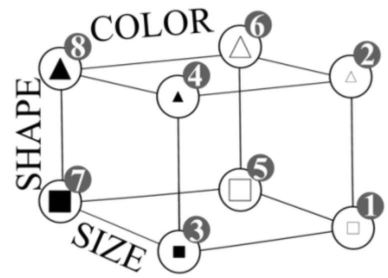


Figure 2: Three binary dimensions creating a total of eight stimuli.

For each of the six SHJ category sets, different aspects of the stimuli were relevant to learning and categorizing the shapes. Type I was a single-dimensional category set, requiring knowledge of only *one* of the three dimensions. Type II was a disjunctive rule (XOR) category set that required knowledge of *two* of the three features. Types III, IV, V, and VI required knowledge of *all three* dimensions to accurately predict category membership. Note that Type IV was a family resemblance category set in which category members shared most features with one another instead of any single feature being necessary for category membership. Thus, all features were relevant to learn the Type IV set. Although all three dimensions are relevant for Types III, IV, and V, some (but not all eight) of the stimuli can be properly classified by knowing the values on just two of the three dimensions.

Further, each Type was divided into six subtypes to counterbalance the relevant dimension and the set that belonged to category 1 or 2. For example, the six Type I problems had a different dimension be relevant and alternated the set that was in category 1 and 2. As a result, each of the six Types had six sub-Types for a total of 36 tasks (or SHJ category sets) used (see Table 1).

Questionnaires. Two questionnaires were used: a general demographics questionnaire and the Analysis-Holism scale (AHS) by Choi et al. (2007). The demographics questions gathered participants' main language, gender, age, English proficiency, and region of current residence. The AHS consisted of 24 items on a 7-point Likert scale (1 = *strongly disagree,* 7 = *strongly agree*) that covered the four major components of analytic-holistic thinking: causality, attitude towards contradictions, perception of change, and locus of attention.

Sample items from each factor respectively as are follows: "*everything in the universe is related to one another*", "*it is more desirable to take the middle ground than to go to extremes*", "*current situations can change at any time*", and "*the whole, rather than its parts, should be considered to understand a phenomenon.*" Holistic thinking is thought to be context-dependent, such that object-object and object-field relationships are considered. This is primarily gauged by items under causality and locus of attention. Extending this way of thinking, holistic thinkers are believed to have a less individualistic and more collectivist mindset, rooted in the idea that elements in the universe are interconnected (Choi et al., 2007). This is different from analytic thinking, which views objects independently and logically (Choi et al., 2007). As a result of this, continuous change is expected among holistic thinkers due to the constant interaction among interconnected elements. Additionally, holistic thinkers believe in finding a 'middle-ground' when faced with contradicting views, whereas a logical, more analytic approach suggests that opposing views *cannot* co-exist (Choi et al., 2007). With the understanding of these four major factors, the AHS aims to broadly measure individuals' mindset under common styles of thinking.

2.2.3 Procedure

The experiment consisted of two components completed in a single testing session: a category learning task and a survey. The category learning task was conducted on Pavlovia using a between-subjects design. To minimize participant confusion, the study was conducted on Qualtrics, with the Pavlovia task embedded. The task began by randomly assigning participants to one of the 36 different versions of the SHJ classification tasks. Each trial began with a target (one second), followed by the stimulus and a prompt for the participant to select either "Category 1" or "Category 2". The prompt lasted three seconds. If no response was given, a reminder message was shown: "Please respond more quickly." Once participants pressed the corresponding key (i.e., either "q" for "Category 1" or "p" for "Category 2"), feedback indicating either "correct" or "incorrect" was shown on the screen for one second, marking the end of the trial. The task consisted of ten blocks of eight trials each, resulting in a total of 80 trials. There were no breaks between blocks. For each block, the eight stimuli were randomly ordered. The task took 15 minutes to complete on average, though the duration could vary by difficulty of the SHJ task Type.

Once the task was completed, participants returned to Qualtrics to complete the survey component. First, participants answered demographic questions followed by an attention check question. Next, participants completed the AHS followed by a second attention check question. This marked the end of the survey component. The survey took approximately ten minutes to complete.

2.3 Results

2.3.1 Percent of Correct Responses

Performance by Task Type. For the following analyses, an average across the 6 different counterbalancing versions of each type was computed. The first analysis looked at participants' performance on the six classification tasks (Types 1-VI). This analysis was straightforward and calculated the number of correct categorization scores by task Type. Performance was collapsed across all blocks for each category Type and accuracy was calculated as a percent of correct responses. This was done by tallying the number of correct responses across all blocks and denoting the final value as a percentage. Figure 3 depicts a downward trend whereby Type I shows the best performance and Type VI shows the worst. A one-way analysis of variance (ANOVA) was conducted to examine the effect of category Type (I, II, III, IV, V, VI) on performance scores collapsed across all blocks. The results revealed a significant main effect of category Type $F(5, 696) = 149.6$, $p < .001$, indicating that performance significantly differed by the assigned SHJ task Type. A Tukey post-hoc test revealed that Types I and VI showed significant mean differences against all other Types ($p < 0.01$). Performance on Type II was significantly better than performance on Types I, V, and VI ($p < 0.01$), but not Types III and IV. Performance difference between Types III, IV, and V were not statistically significant, demonstrating that performance did not differ significantly between these three types.

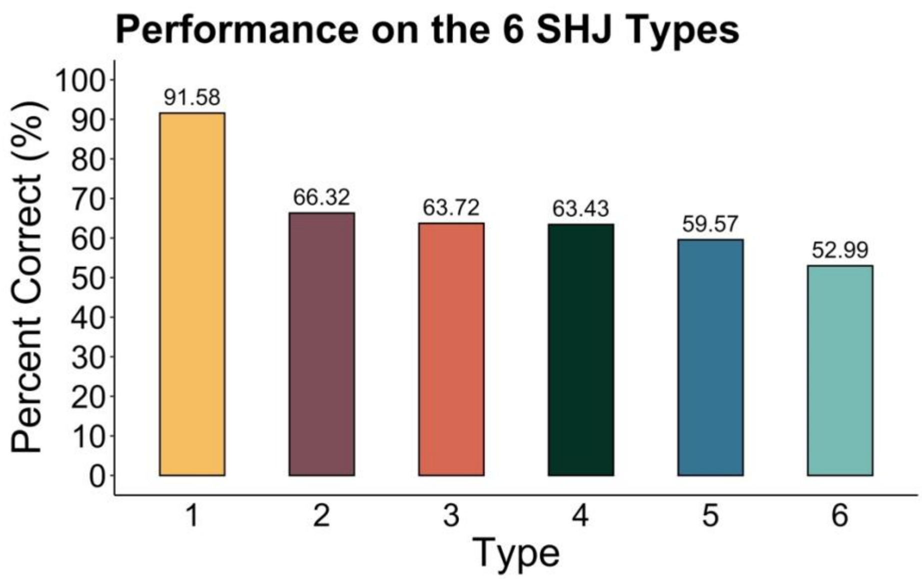


Figure 3: Performance on the six SHJ tasks collapsed across blocks.

Performance across blocks. Performance across blocks was analyzed by calculating the learning accuracy during each block for each SHJ Type for a total of ten blocks. Once again, accuracy was defined as a percent of correct categorization responses. The percent accuracy scores across blocks were grouped by SHJ task Type. High correct categorization responses relative to incorrect ones indicate that the participant learned the task, whereas higher incorrect responses demonstrate that the participant did not learn the task. Participants were expected to perform at a lower rate in the earlier blocks since the task would have yet to be learned. As more blocks were completed, participants had more exposure and received more feedback on the learning task. Successful learning of the category set was demonstrated by an upward trend across blocks, exhibiting higher percent accuracy over time. The learning curves in Figure 4 align with those of Shepard et al. (1961), with Type I being the easiest and fastest to learn and Type VI being the most difficult. It can also be seen that learning differences among Types II, III, IV, and V are more apparent in later blocks relative to the early blocks. As can be observed in Figure 4, Types III, IV, and V were similar in difficulty as predicted. Hence, this analysis indicates that participants learned the separate SHJ tasks as anticipated, with Type I being the easiest, followed by Type II, then Types III, IV, V, and lastly, Type VI being the most difficult.

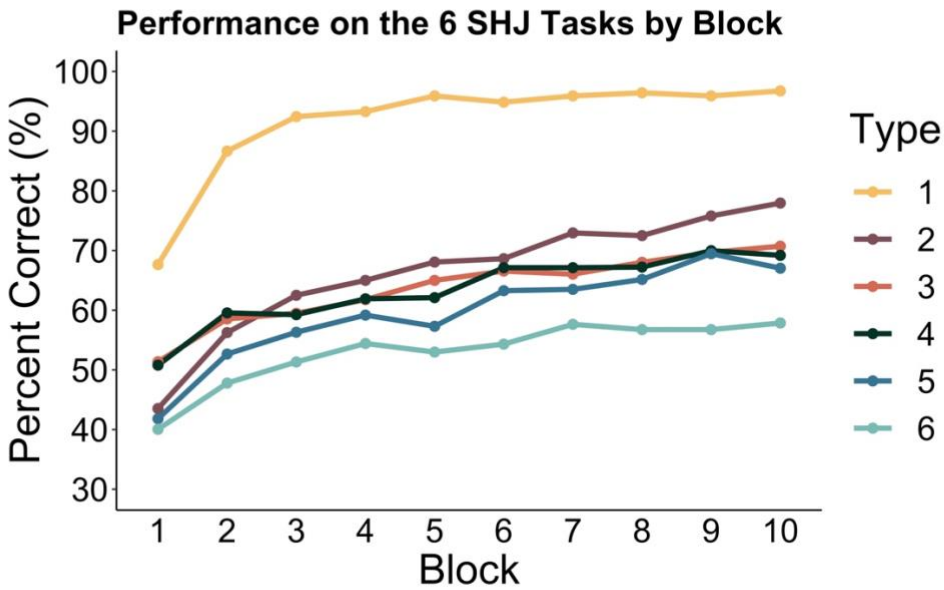


Figure 4: Performance across the six SHJ task Types by block.

2.3.2 Correlational Analyses

Task Type and AHS Score. To examine the relationship between category learning performance on the SHJ tasks and AHS scores, several Pearson correlations were conducted. This was done by grouping the percent accuracy of learning performance by task Type across all blocks and correlating those scores with corresponding mean AHS scores. Before mean AHS scores were calculated, six items were reverse scored in accordance with analyses by Choi et al. (2007). Then, an average score of all 24 AHS questions was calculated, resulting in a single AHS value for each participant. Higher AHS scores reflected a stronger tendency for holistic thinking. Six correlational analyses were carried out (Table 3; Figure 5) to inspect for a relationship between the performance on the task Type and AHS score. Each analysis used a different dataset, corresponding to each of the six task Types. The results show that the relationship between performance on the Type IV SHJ set and AHS score ($r = 0.31$, $p < .001$) was significant. Although a moderate correlation, this association being in the positive direction signifies that as predicted, holistic thinking was correlated with Type IV task performance. No other significant correlations were observed between the SHJ task Types and AHS scores. Note that although non-significant, the Type I ($r = -0.01$, *ns*) and Type II (r

= -0.06, *ns*) category sets showed a negative association with AHS scores. This trend, albeit small, was in the expected direction. Interestingly, it can also be observed that Type V (r = -0.14, ns) showed an unanticipated non-significant, negative trend as well. The internal consistency of these two items was also calculated, with a Cronbach Alpha value of $\alpha = .01$.

Table 3: Correlation analysis of task Type and AHS Score

Task Type	**N**	***r***	***p*-value**
Type I	119	-0.01	0.89
Type II	110	-0.06	0.50
Type III	120	0.09	0.35
Type IV	127	0.31	0.00
Type V	113	-0.14	0.15
Type VI	113	0.04	0.70

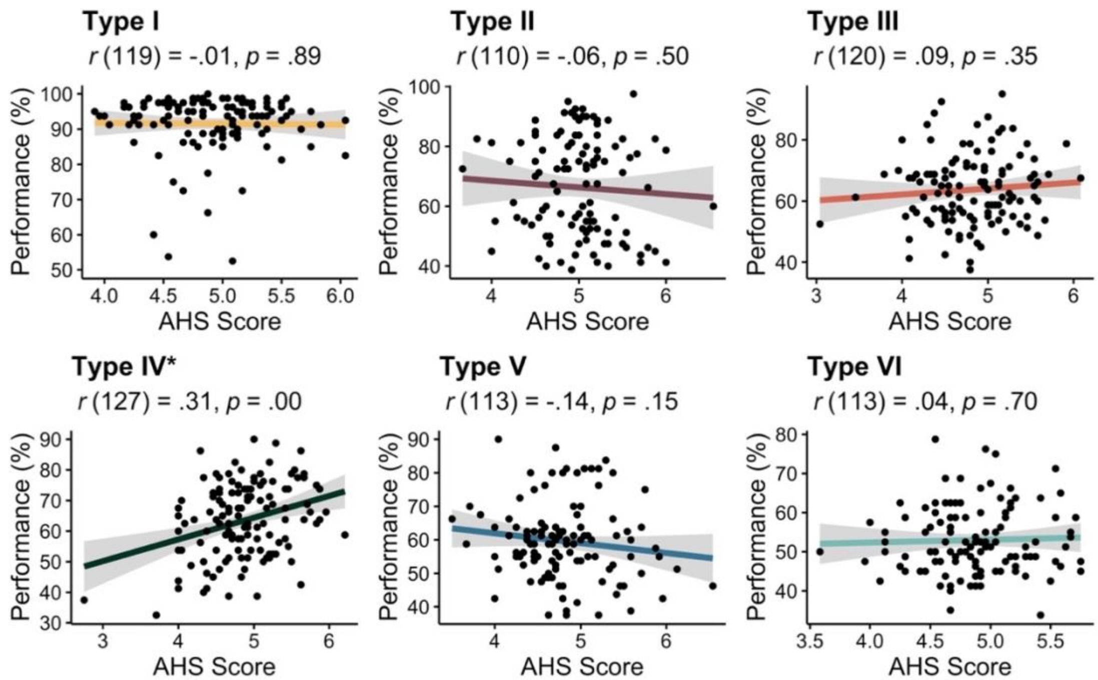


Figure 5: Correlational analysis plots by task Type.

2.3.3 Exploratory Analyses

To survey trends related to the region in which participants were residing, secondary correlational analyses were carried out. Despite the imbalanced regional distribution, sufficient data was collected from North America, South America, Europe, Africa, and Asia. Due to insufficient data from Australia, Pacific Islands, and the Caribbean Islands, these data were not included in these secondary analyses. The region data was coded to assign numbers 1-11 in the order of presentation on the survey (*North America=1, Central America=2, South America=3, Europe=4, Africa=5, Asia=6, Australia=7, Pacific Islander=8, Caribbean Islands=9, Other=10, Prefer not to say=11*).

Multiple separate regression analyses were conducted to predict category learning performance based on the region participants were living in. A regression model using the full dataset, not separated by region or task Type, was run with the predictor of region on AHS scores. Note that data corresponding to "Other" and "Prefer not to say" were not included in this analysis. Interestingly, this model was significant, indicating that AHS scores were generally predicted by the region in which participants were living in ($F(1, 683) = 8.09$,

$p < .00$). Since each of the SHJ category sets was uniquely structured to optimize different categorizing strategies, individual regression analyses were carried out for each category Type. Each model used a different dataset, corresponding to each of the six task Types. This was done to account for the finding of learning accuracy being differentially predicted by the task Type. All six models were carried out as standard regression models with the predictor of region and outcome variable of AHS scores. The only difference among models was the dataset used, which corresponded to a specific SHJ category Type (I-VI). Of the six regression models, only the Type IV model showed significance, such that for the family resemblance task, the region participants lived was a significant predictor of AHS scores ($F(1, 125) = 6.15$, $p < .05$). A summary of AHS scores by region is shown in Table 4.

Two correlational analyses were also explored, each using different variables to group the data. The first analysis inspected for any relationship between performance on the tasks and region of residence, with the data grouped by task Type (I-VI). Accordingly, six analyses, one for each category Type, were carried out to observe any correlation between task performance (represented as percent accuracy) and region. The findings indicated that only Type IV performance showed a relationship with region ($r = -0.17$, $p = .05$). No other task Type showed a significant relationship of performance with region (*ns*).

The second analysis grouped the data by region, forming 11 groups in total – one for each of the nine regions listed in the questionnaire, and two for "Other" and "Prefer not to say". Given insufficient data for three regions, "Other", and "Prefer not to say", six correlations were carried out. Each analysis looked for a relationship between AHS scores and task performance (represented as percent accuracy). A significant relationship between AHS scores and task performance was found for participants in Europe ($r = 0.20$, $p < .05$) and Asia ($r = 0.74$, $p < .01$) only. Both European and Asian participants showed a higher overall task performance across all Types with higher AHS scores (more holistic thinking). Though this result depicts that AHS scores and task performance may be related, this result is also indicative of variance within the results of each region. Generally, the AHS scores corresponding to each region seem to be in the mid-to-high holism range. Thus, varying levels of holism may exist within a given region. Moreover, task performance, since it was not grouped by task Type, also shows a broad range. This is expected as performance varies across Types as seen in the rank-order effects of learning the SHJ task Types.

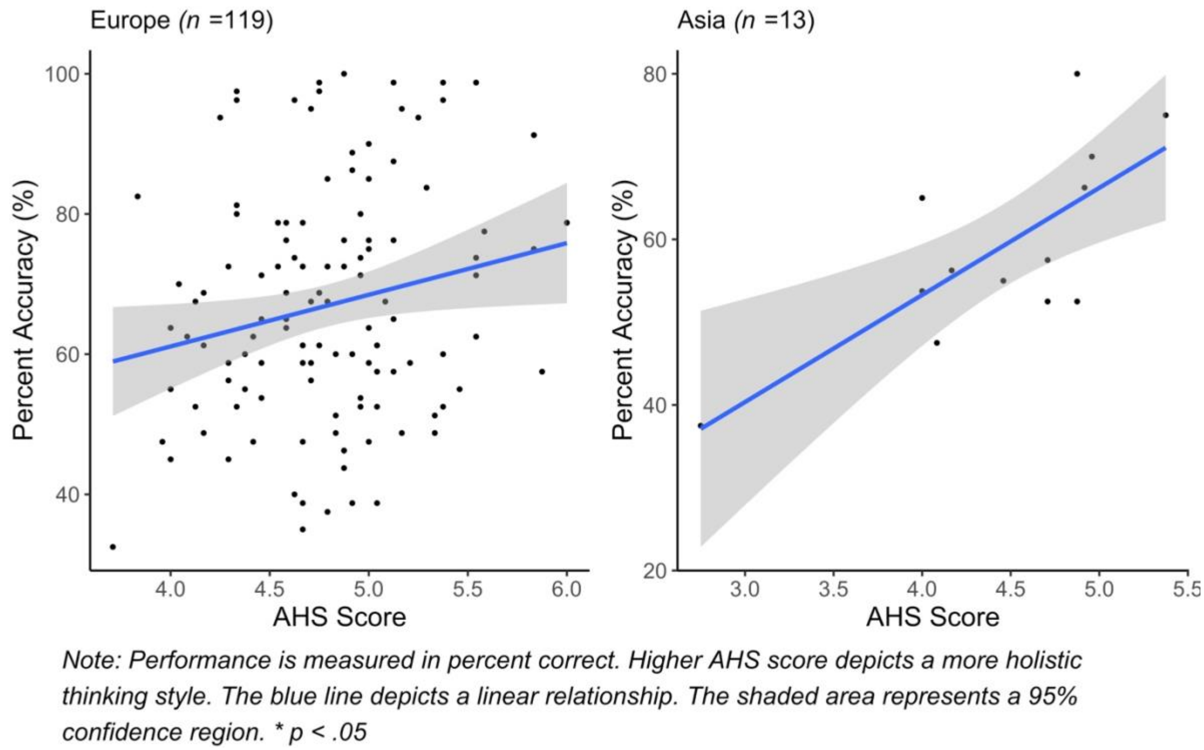


Figure 6: Scatterplots showing percent accuracy and AHS scores for Europe and Asia.

Table 4: AHS scores by region

	N	*M*	*SD*	*Min*	*Max*
North America	512	4.90	0.48	3.46	6.54
Europe	119	4.79	0.45	3.71	6.00
Asia	13	4.45	0.67	2.75	5.38
Africa	29	5.08	0.44	4.33	5.92
South America	5	4.83	0.57	4.29	5.79
Central America	5	4.53	0.43	4.17	5.25

Caribbean Islands	2	3.83	1.12	3.04	4.63

2.4 Discussion

The present study examined the relationship between AHS scores and performance on six Types of category learning tasks, adapted from Shepard et al. (1961). The six types consisted of a unidimensional rule task (Type I), a complex disjunctive rule-based task (Type II), nonlinear rule tasks (Types III and V), a family resemblance task (Type IV), and an incoherent category task (Type VI). The AHS scale by Choi et al. (2007) was administered to provide additional insight on analytic versus holistic thinking, with higher scores indicating more holistic thinking. The secondary aim of this study was a feasibility assessment for online categorization data collection. The widely used SHJ category learning tasks were ideal for this study given the expected rank-order effects from Types I–VI. The task followed a between-subjects design in which participants completed the category learning task and questionnaires in a single session on a computer. Data was collected from participants both in person and online using multiple recruitment platforms. Learning accuracy on the SHJ task was calculated as a percentage of correct categorization responses tallied across all ten blocks. It is worth discussing that the subject pool was primarily university students living in North America rather than a more culturally and regionally diverse sample. This was attributed to the early piloting stages of this study in that initial data collection was geared towards simply collecting reliable data. For this reason, exploring cultural differences was limited to the data collected from participants recruited using Prolific. However, given the between-subjects design, the size of the Prolific sample was insufficient to analyze independently. Thus, AHS scores were imperative to compensate for this limitation as insight into thinking styles was obtained using AHS scores as the discriminator instead of region.

Looking at performance differences in learning the SHJ category sets, the results of this preliminary study show a similar trend to that of Shepard et al. (1961). As expected, in order of difficulty, Type I was the easiest to learn, followed by Type II, then Types III–V, and lastly, Type VI was the most challenging one to learn. Types III, IV, and V were expected to be of equal difficulty due to the high similarity in the category structures as discussed in the methods. The COVIS proposition of the procedural implicit system predicts that complex category sets are learned more gradually and automatically. Figure 4 supports this, as learning more complex rules, in contrast to a unidimensional one, is observed to be a gradual process. The unidimensional rule was easily learned in a shorter amount of time in contrast to the more complex and non-linear rule tasks. Holistic thinking considers context and relationships among features to capture objects in their entirety as opposed to a context-independent outlook. Thus, it can be drawn that holistic thinking should be more beneficial when learning complex category sets. This is a product of the multiple features and relationships that must be accounted for when learning complex category sets. In this way, these findings indicate that unique optimal strategies predict accuracy on most SHJ task Types, while also exhibiting feasible data collection using an entirely online task. Furthermore, results of this study tentatively suggest that category learning and classification preferences among different language and culture groups across locations, assuming appropriate translation of material, can be studied using this methodology. Therefore, appropriate scalability of these methods to address the 'WEIRD problem' in this research domain seems promising.

The correlational analyses found holistic thinking to be significantly related to learning the Type IV SHJ category set. This relationship in the positive direction was anticipated as the Type IV set is optimally learned using a family resemblance strategy. Holistic thinking, also known as similarity-based thinking, focuses on relationships between objects and overall similarity rather than object focal features alone. Being field-dependent, this style of thinking considers the context surrounding focal features, prompting intuitive judgments instead of solely relying on formal reasoning. Holistic thinking is comparable to the family resemblance strategy as the overall similarity of category members is perceived to make an intuitive judgment. Using a family resemblance strategy requires gradual learning over multiple exemplars wherein there is no one set of necessary and sufficient features identifiable for category membership. This resembles incidental learning seen in holistic thinking. Moreover, small negative associations were seen for AHS scores and the Types I, II, and V category tasks. Though these trends were non-significant, it seems higher accuracy on these SHJ category sets was predicted by lower AHS scores, indicating an advantage for analytic thinkers on these tasks. Types I and II were predicted to trend in this direction owing to the optimal strategy for Type I being a single-feature rule and Type II being a complex disjunctive rule, both of which align well with an analytic thinking style. However, this trend for Type V was unexpected. This is because the Type V set involves all three dimensions for optimal categorization, which should favour holistic thinking. A plausible explanation for this is that since some stimuli, although not all eight, can be properly categorized using the values on just two of the three dimensions, a similar strategy to the disjunctive rule set (Type II), may have been used. This strategy would not have been optimal; yet it would be functional. More data from a broader range of participants may allow for more variability in AHS scores and consequently, potentially a stronger test of this relationship.

The AHS scores did not correlate as strongly as expected with task performance. One explanation for this is the existence of other competing correlates contributing to task performance. Despite the sparse region datapoints, exploratory analyses were conducted looking at relationships between where participants were living, task performance, and AHS scores. It was hypothesized that an effect of culture showing a relationship between thinking style and region would be found. Higher AHS scores were expected from non-WEIRD samples, indicating higher holistic thinking among these populations (Choi et al., 2007). Inspecting the AHS scores by region (Table 4) shows similar mean scores across WEIRD and non-WEIRD regions. This indicates that despite the varying range, our sample was more holistic overall. That being said, the regression model showed that region significantly predicted AHS scores across all task Types, providing support for the AHS scale. However, when the data was grouped by SHJ task Type, this effect only appeared in the Type IV dataset. Hence, it is likely that this finding in the Type IV dataset may have driven the significant finding of the overall, ungrouped dataset. Type IV task performance was also significantly correlated with region. It is notable that region was a significant predictor of AHS scores given that task performance on the family resemblance task, Type IV, was significantly correlated with AHS scores as well. Related to this, region might partly be driving task performance on this SHJ category set. Nonetheless, it is important to recognize the caveat of unbalanced sample sizes whereby the large North American sample could be skewing these findings. Lastly, despite AHS scores significantly correlating with task performance across all category sets for European and Asian populations, this should also be cautiously considered due to the small sample sizes recruited from these regions. Furthermore, grouping all of Europe together can be misleading since Europe comprises a range of WEIRD and non-WEIRD regions. Overall, these results warrant additional thorough and unbiased analyses before drawing any major conclusions.

Another explanation for the strength of associations between AHS scores and task performance can be attributed to a majority Western participant pool. Given the high number of North Americans completing the task, our sample does not consist of a fair distribution of thinking styles. While it would be expected that the high North American population would likely favour analytic thinking, this is not the case. Nonetheless, AHS scores showed a significant relationship with region across all task Types, revealing that scores were sensitive to region differences. Further, it is important to note that with a more diverse population, the AHS scale by Choi et al. (2007) can be properly evaluated for accuracy in predicting analytic versus holistic thinking styles across culture groups. This goes to say that supplementary analyses validating this scale are required to understand unexpected trends. In short, although we are designing and testing a method to address the WEIRD problem, our data are still clearly susceptible to this same bias. More work is needed, and we view these results as provisional but encouraging.

Finally, a reason that the results did not entirely show expected trends could be the insufficient variability on the very construct that we wish to address and study, effects of culture. With the between-subjects design, the final sample had an average of 117 participants per group. This does indicate adequate power, though a more distributed population, both culturally and age-wise, may better show the predicted effects.

In conclusion, this study successfully gathered data on an entirely online categorization task and showed expected trends in learning the optimal strategy to accurately categorize six different category learning tasks. The preliminary analyses show that the correlation between task Type and AHS scores for the family resemblance task (Type IV) is significant, and the rule-based tasks (Types I and II) are trending in the predicted directions. With future data collection, we hope to gain more insight into the category learning task types and thinking styles, as well as the relationships between culture and analytic versus holistic thinking. Going forward, further analyses to validate the AHS, collecting data from a more diverse sample to better assess effects of culture and age, and exploring effects of language are recommended.

Chapter 3

3 Experiment 2

3.1 Introduction

Categorization is an essential cognitive process of grouping into classes based on distinguishing characteristics and features, such as seen in basic level and spontaneous classification (Kemler Nelson, 1984; Rosch & Mervis, 1975). Most commonly, categorization research involves sorting and/or assigning tasks. Sorting tasks present participants with items to be divided into the best and most natural groups, whereas assigning tasks present participants with a target item to be placed in one of two or more groups (Smith et al., 1999). Although previously focused on WEIRD populations, categorization research has more recently included culturally diverse populations to explore effects of culture. That said, the findings have revealed mixed results. Some studies have found East Asians to be more holistic (Choi et al., 2007; Norenzayan et al., 2002; Ji et al., 2004), whilst others have found Westerners and East Asians to be equally holistic (de Oliveira & Nisbett, 2017), and a few have even hinted at thinking style differences within Western subpopulations (Knight & Nisbett, 2007; Varnum et al., 2008). Motivated by equivocal findings, the primary purpose of this study was to analyze categorization performance and thinking styles on a well-known task. This would allow for assessing replicability and to observe any new findings. A supplementary goal was to broadly investigate the reliability and validity of the AHS by Choi et al. (2007) in examining analytic versus holistic thinking tendencies.

Cross-cultural categorization research commonly makes comparisons between East Asian and American populations using differing paradigm specifics ranging from the nature of the task, particular instructions, conditions, and stimuli type – i.e., words versus objects. It has been argued that the relative influence of similarity- versus rule-based strategies over the other can be driven by both the nature of the task and the precise instructions presented (Evans & Over, 1996; Norenzayan et al., 2002). Seeing as forced-choice tasks are frequently cited in the literature, this research aimed to test the replicability of the widely cited rule-versus family resemblance-based classification and similarity judgments task by Norenzayan et al. (2002). Outcomes from Norenzayan et al. (2002) show that across conditions, participants preferred to categorize based on a rule. However, a culture by response-type interaction was only found in the similarity condition, seen as East Asians tended to use family resemblance strategies rather than rule-based ones. Though, it is important to note that these results have not always replicated (Murphy et al., 2017).

Norenzayan et al. (2002) examined similarity versus categorization judgments using a category structure borrowed from Kemler Nelson (1984). Rather than a free-sorting task, participants in this study assigned a target object to one of two categories that could be defined by either a unidimensional rule or by overall similarity. Two groups of four members with four binary features each were created for a total of eight items. For one of the four features, all members in the first group had one version and all members in the second group had the other version of the same feature. The other three features were more variable among members to allow for a general overall similarity within each group. For example, in the flower stimulus set (see Figure 7), members in group one have straight stems, whereas group two members have curved stems. The rest of the features (i.e., petal type, presence of a leaf, and circle in the head) varied such that three of the four members of each group had one version of these features. This increased similarity within group members and dissimilarity between group members. The target stimuli were specifically constructed to ensure that responses driven by rule or family resemblance criteria led to different decisions, with no *correct* way to categorize (Norenzayan et al., 2002). Thus, if participants relied on a unidimensional rule, categorization would be based on a single feature shared by the target object and all category members. Alternatively, relying on family resemblance would prompt them to categorize based on overall similarity among the target object and category members with no shared necessary and sufficient feature (Norenzayan et al., 2002; Murphy et al., 2017). To test whether task instructions evoke differing strategies, participants were randomly assigned to one of two conditions: a similarity judgment condition or a categorization condition. Each condition differed solely in the task instructions, keeping all other components the same. The similarity judgment condition explicitly asked participants which group the target was “most similar to”, encouraging a family resemblance strategy. This was different from the categorization condition, in which participants were asked which group the target “belongs to”.

The current consensus regarding cultural differences in categorization holds that East Asians tend to think holistically whereas Westerners default to an analytic thinking style (Masuda & Nisbett, 2001; Nisbett et al., 2001; Norenzayan et al., 2002; Norenzayan & Nisbett, 2000). Accordingly, it is hypothesized that East Asians will make family resemblance judgments more than Westerners, who will classify based on rules more. Holistic thinking focuses on overall similarity, relationships between objects, and a context-dependent view, which is in line with the family resemblance strategy. This is different from analytic thinkers who are anticipated to show rule-based categorization. As well, holistic thinking is expected among East Asians, whereas analytic thinking is expected among Westerners (Nisbett et al., 2001; Norenzayan et al., 2002; Norenzayan & Heine, 2005). The expected effects of task instructions and reaction times follow the findings by Norenzayan et al. (2002), demonstrating that higher holistic thinking and faster response times will be seen in the similarity judgment condition compared to the classification condition. Faster response times are expected in the similarity judgment condition as family resemblance judgments become automatic over time. This is a product of requiring less cognitive resources when making within category similarity judgments given the emphasis on holistic thinking. Conversely, the classification condition is expected to elicit an analytic thinking style requiring a discriminative rule to be applied each time. Applying a rule in this way requires more cognitive resources, thus using formal logic rather than an automatic judgment. It is also expected that an effect of culture will be seen in the similarity judgment condition (Norenzayan et al., 2002; Saalbach & Imai, 2007). To assess the reliability of the AHS scale, scores will be analyzed as a predictor of task strategy. It is hypothesized that higher scores on the AHS will predict higher family resemblance categorization and lower scores will show more rule-based responding (Choi et al., 2007).

A secondary aim of this research is to explore potential mediating and/or moderating variables and their role(s) in differing categorization strategy use. While the broader question of universal category learning strategies across culture groups remains unknown, very little research exists exploring other contributing factors. This called for the inclusion of exploratory measures to assess individual differences in personality and processing styles. Specifically, the Autism-Spectrum Quotient-10 (AQ-10; Allison et al., 2012), and the Big-

Five Inventory-2-Extra-Short Form (BFI-2-XS; Soto & John, 2017a) were administered alongside some demographic questions.

Autism-Spectrum disorder (ASD) has been associated with highly focused attention, seeing as individuals with autism have difficulty disengaging attention from a stimulus or task (Klinger & Dawson, 2001). This presents challenges in prototype abstraction, which is a process that requires attention to multiple features of an object rather than any single feature. Different from rule-based categorization whereby a single feature is the focus, prototype abstraction parallels the more holistic family resemblance strategy. Accordingly, is it expected that a high ASD score, denoted by high mean AQ-10 score, will correlate negatively with holistic thinking (Klinger & Dawson, 2001; Mercado et al., 2020). Among the big five personality domains, agreeableness and conscientiousness are particularly of interest. Agreeableness has been found to predict a style of thinking highlighted by Choi et al. (2001) as one of holistic thinkers. This is seen by efforts to minimize interpersonal conflict (Graziano et al., 1996; Jensen-Campbell & Graziano, 2001) and maintain intragroup cooperation (Graziano et al., 1997). Agreeableness has also been associated with orienting to object-field relationships, thereby viewing the whole as greater than the sum of its parts (Ashton & Lee, 2001; Digman, 1997; Wiggins & Trapnell, 1997). On the contrary, conscientiousness has been linked to high attention and effort to detail and to task-relevant activities, prompting more feature-oriented, rule-based responding (Gellatly, 1996; Alaei et al., 2014). Consequently, it is hypothesized that higher scores for agreeableness and lower scores for conscientiousness on the BFI-2-XS will correlate positively with holistic thinking.

To summarize, this research hopes to test the replicability of a widely cited task by Norenzayan et al. (2002), while also investigating analytic versus holistic thinking using the AHS and other supplementary explanations of differences in thinking styles.

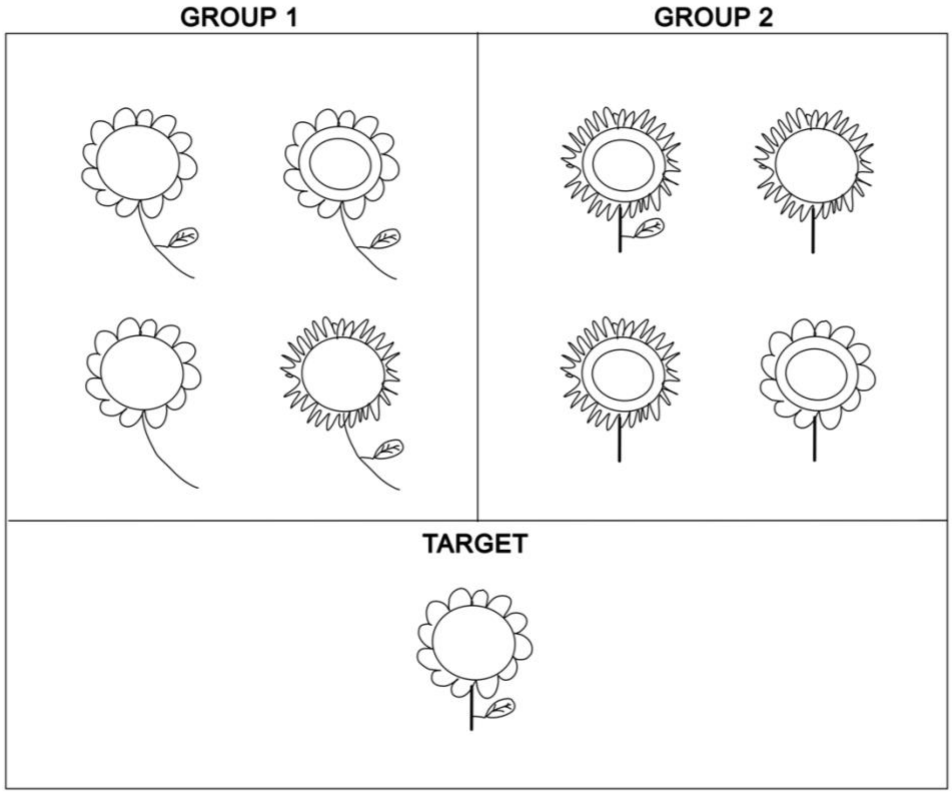


Figure 7: An example of stimuli used in the task. Each one of the features here corresponds to one of the binary values in Table 5. The target item is to be assigned to one of the two groups. For example, the four binary features here are: straight vs. curved stem, spiked vs. curved petals, the presence or absence of a leaf, and the presence or absence of a circle in the head. The target perfectly matches one group on one dimension (all of Group 2 have straight stems) and imperfectly matches the other group on three dimensions (most of Group 1 has curved petals, leaves, and no circle in the head). This allows for the contradiction of a perfect rule against multiple matching features.

3.2 Methods

3.2.1 Participants

All participants were recruited and tested online. Two recruitment platforms were used, for a total of 311 participants across platforms. 123 of these were undergraduate students at Western University and were compensated with course credit in an undergraduate psychology course. Fluency in English was the only prerequisite for the undergraduate psychology students. No other pre-screening or exclusion criteria were used for this population. 179 participants were recruited on Prolific and compensated in GBP. The Prolific sample was pre-screened for fluency in English and balanced to collect equal numbers of male and female participants.

The final sample consisted of 296 adults (*M* = 25.19, *SD* = 9.40, *range* = 18.00 - 65.00) after cleaning the data for failed attention checks, mis-matched IDs, and other incomplete fields. The final sample had a total of 129 males, 163 females, and four self-identifying individuals. Other descriptive statistics, including English proficiency, AHS scores, AQ-10 scores, and BFI-2-XS scores can be observed in Table 8.

The most common first language among the sample was English (*n*=156). Other languages recorded were Hindi (*n*=76), Spanish (*n*=23), Portuguese (*n*=18), Mandarin (*n*=7), Arabic (*n*=4), French (*n*=4), and Russian (*n*=2). Two individuals stated 'Other' and four preferred not to say. Most participants were living in North America (*n*=128), with the rest distributed around Asia (*n*=82), Europe (*n*=52), Africa (*n*=22), Central America (*n*=3), and Other (*n*=9).

3.2.2 Materials

Stimuli. The stimuli used were taken from the original study carried out by Norenzayan and colleagues (2002) and re-constructed to achieve a higher quality digital image. Of the 20 stimulus sets, each set had an abstract structure constructed from a set of four binary features as shown in Table 5. The set showed a category pair (Group 1 and Group 2) side-by-side comprised of four items each, and a clearly labelled target object at the bottom to be classified in either Group 1 or Group 2. Groups were labelled numerically, avoiding any category labels. A total of ten category pairs were used and two alternative target objects were created for each category pair. To counterbalance the design, the ten category pairs were presented with only one of the two alternative targets shown with each pair. This resulted in a total of 20 images presented to participants.

Each category pair corresponded to a specific type of item (see Figure 7). Note that the target objects were created to allow for classification in either group depending on the strategy used. One of the four binary features in each group defined the group, which enabled the use of a rule based on a necessary and sufficient feature. To facilitate family resemblance classifications, the other three features within each group took on different values. In this way, an overall similarity among items was observable since each of the four items in each group contained three of the four binary values (Table 5).

Table 5: Stimuli category structure breakdown.

Group 1	Group 2	Target objects
0000	1111	0111
0100	1011	1000
0010	1101	
0001	1110	

Questionnaires. Four questionnaires were used: a general demographics questionnaire, the AHS (Choi et al., 2007), the Autism-Spectrum Quotient-10 (AQ-10; Allison et al., 2012), and the Big-Five Inventory-2-Extra-Short Form (BFI-2-XS; Soto & John, 2017a).

The demographics questions gathered participants' main language, bilingualism, English proficiency, age, gender, and region of current residence.

The AHS consists of 24 items on a seven-point Likert scale (1 = *strongly disagree,* 7 = *strongly agree*) that cover the four major components of analytic-holistic thinking: causality, attitude towards contradictions, perception of change, and locus of attention. Higher scores on the scale indicate higher holistic thinking.

The AQ-10 and BFI-2-XS scales are shortened versions of their original, lengthier questionnaires. These served as secondary analyses for exploratory purposes. The shortened versions were used to counteract participant fatigue, frustration, and careless responding.

The AQ-10 is a quick 10-item questionnaire, shortened from the original 50-item Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001). The AQ is structured around five subdomains of social interaction, communication, attention to detail, attention switching, and imagination – all of which are characteristic of adults with Autism Spectrum Conditions (ASC). The AQ-10 retains the reliability and predictive validity of the full 50-item inventory in that those with ASC are more likely to score higher on the AQ-10 (Allison et al., 2012). The ten items included on the AQ-10 are the most discriminating items on all versions of the AQ, with two items from each of the five subdomains (Allison et al., 2012). Each of the ten items is rated on a four-point Likert scale (1= *definitely agree*, 4= *definitely disagree*), with a score of six or higher signifying a potential ASC and requiring further diagnoses.

The original, 60-item Big-Five-Inventory-2 questionnaire (BFI-2; Soto & John, 2017b) assesses the Big Five personality domains and 15 facets: Extraversion (with facets of Sociability, Assertiveness, and Energy Level), Agreeableness (with facets of Compassion, Respectfulness, and Trust), Conscientiousness (with facets of Organization, Productiveness, and Responsibility), Negative Emotionality (with facets of Anxiety, Depression, and Emotional Volatility), and Open-Mindedness (with facets of Intellectual Curiosity, Aesthetic Sensitivity, and Creative Imagination) (Soto & John, 2017b). The BFI-2-XS consists of just three items per domain, for a total of 15 questions rated on a five-point Likert scale (1 = *strongly disagree,* 5 = *strongly agree*), while still retaining most of the validity and reliability of the BFI-2 (Soto & John, 2017a). Given the conciseness of the 15-item BFI-2-XS, this version is better suited to assess personality only at the level of the Big Five domains, rather than facet-level traits. This fits within the scope of this study to consider the impacts of the Big Five personality domains on thinking styles (Soto & John, 2017b). Each of the 15 items follows the initial prompt being "*I am someone who…*". Sample items from the domains of Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Open-

Mindedness respectively as are follows: "*Is dominant, acts as a leader*", *"Is compassionate, has a soft heart", "Is reliable, can always be counted on", "Worries a lot", "Is fascinated by art, music, or literature."*

3.2.3 Procedure

The experiment involved two components to be completed in a single session: a categorization task and a survey, both of which were completed on a computer. Instructions and materials for all components were provided in English.

The categorization task was carried out on Pavlovia and followed a between-subjects design, such that participants were randomly assigned to either the similarity judgment or the classification condition. The stimuli remained the same regardless of the condition, with the only difference being the specific instructions provided. Participants were asked to choose, depending on the condition, which group the target item 'belongs to' (classification condition) or is 'most similar to' (similarity judgment condition). Following this, participants completed the forced-choice categorization task. The task began by presenting one stimulus set at a time, each depicting two groups and a target item to be categorized.

Each trial began with a fixation cross (one second), followed by a single stimulus set and a prompt for the participant to select either "Group 1" or "Group 2" by pressing either the "q" or "p" key respectively. If no response was given, a reminder message was shown: "Please respond more quickly." The task consisted of 20 stimulus sets presented to participants in a random order, with no breaks in between stimuli. As soon as a response was made, the program automatically showed the next stimuli. The task took less than 10 minutes to complete. The dependent measure was the percentage of rule or family resemblance responses per participant, averaged across the 20 trials. Reaction times (RT) were also recorded.

The surveys were conducted on Qualtrics following the categorization task. First, participants answered demographic questions followed by an attention check question. Next, participants completed the AHS followed by a second attention check question, the AQ-10 scale, and a third attention check question. Finally, participants completed the BFI-2-XS followed by a final attention check question. This marked the end of the survey component. The survey took approximately 15 minutes to complete.

3.3 Results

3.3.1 Task Scoring

Responses on the categorization task were indicated by specific keys representing either an analytic or holistic strategy. To quantify the responses, the percent of holistic, or family resemblance, responses was calculated for each participant. This was done by calculating the mean number of holistic responses across the 20 trials and multiplying this value by 100 to represent a percentage for each participant.

3.3.2 Condition Differences

Across All Groups. An independent samples *t*-test was conducted to compare the holistic responses on the task, quantified as a percentage, in the similarity and classification conditions. The results indicated a near-significant difference in the percent of holistic responses by instructional group; $t(294)=-1.82$, $p = .069$. Despite higher holism seen in the similarity condition ($M=51.32$, $SD=17.25$) compared to the classification condition ($M=47.55$, $SD=18.29$), these differences were not significant.

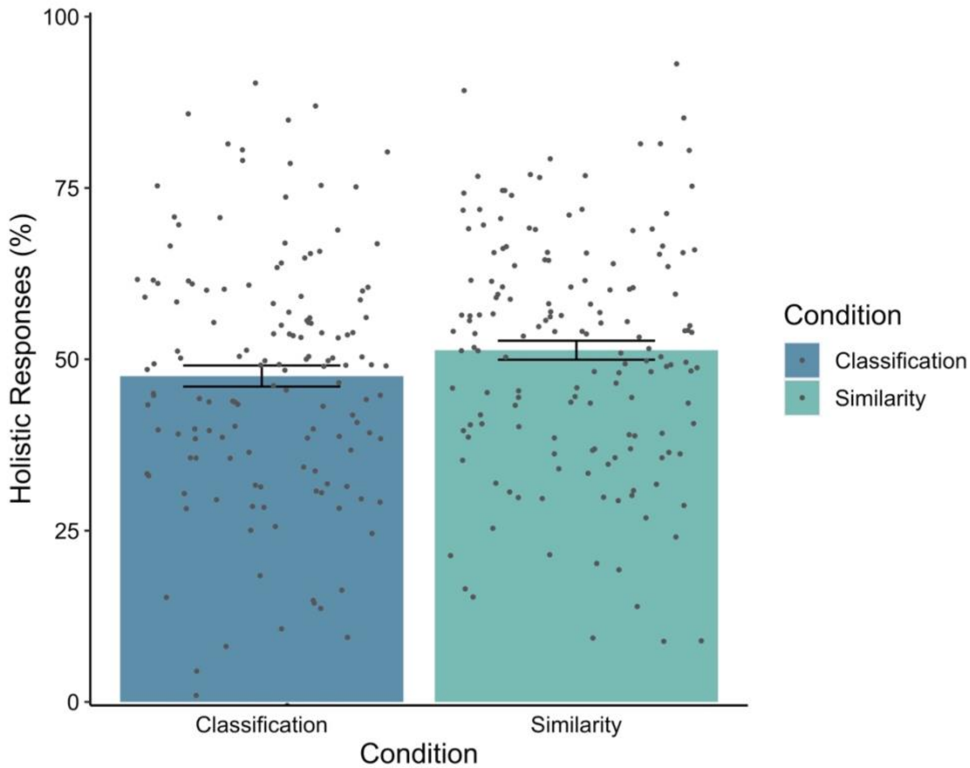


Figure 8: A bar plot showing holistic responding in the classification versus similarity conditions across the full sample.

Grouped by Culture. The sample was grouped into WEIRD and non-WEIRD populations based on region of residence. All regions other than Europe and North America were classified as non-WEIRD samples, with Europe and North America comprising the entire WEIRD group. An independent samples *t*-test was conducted within each group to compare the percentage of holistic responses in the similarity versus classification conditions.

The results indicated that holistic responding significantly differed between the similarity and classification conditions ($t(178)=-2.77$, $p < .01$) among the WEIRD sample ($n=180$). In this population, the classification condition elicited 45.4% family resemblance responses compared to 52.3% in the similarity instruction condition. This was different among the non-WEIRD sample ($n=116$), which showed no difference between groups ($t(114)=0.27$, $p = .79$). This sample showed 49.7% family resemblance responses in the classification condition compared to 50.6% in the similarity judgment condition.

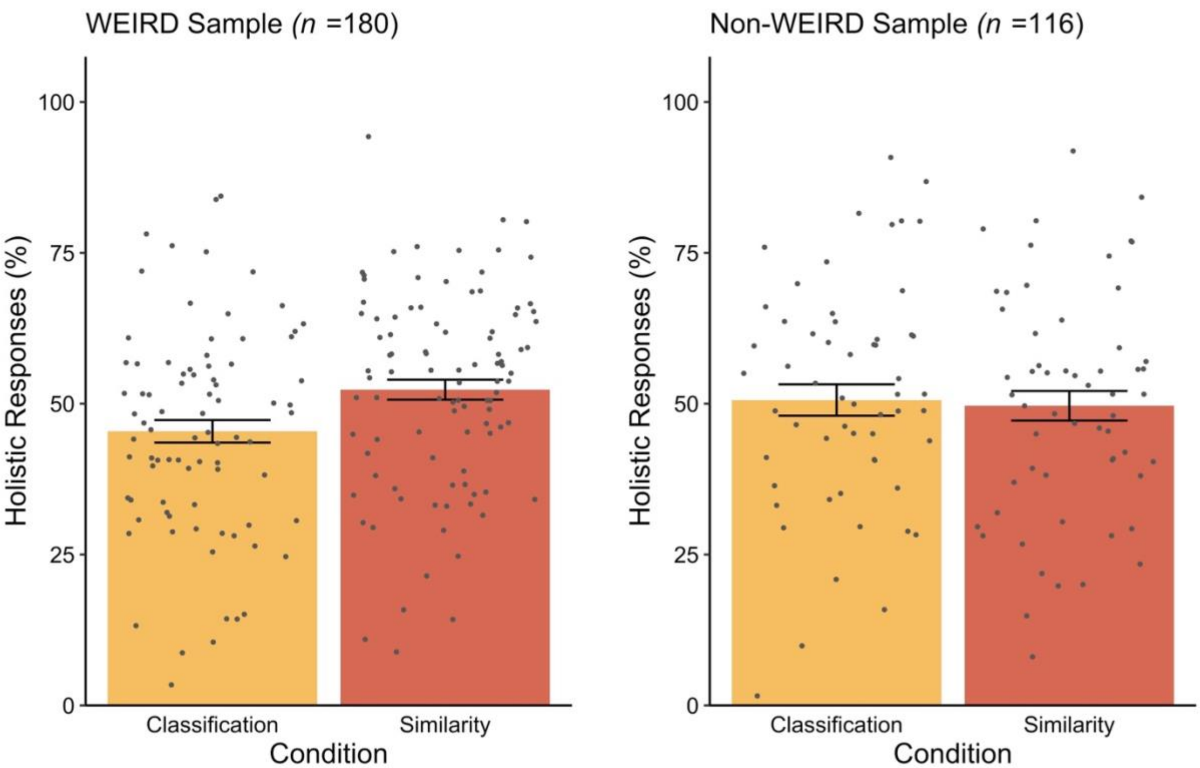


Figure 9: Percentage of holistic responding in each instructional condition (classification vs. similarity) among WEIRD and non-WEIRD samples.

3.3.3 Regression

A multiple regression analysis was used to test if the percentage of holistic responses was significantly predicted by the independent variables of condition (similarity versus classification), AHS score, AQ-10 score, and each of the big five domains (extraversion, agreeableness, conscientiousness, negative emotionality, and open mindedness). These eight predictors were entered into the model. The results of the regression (Table 6) indicated that the instructional condition and the predictor of conscientiousness were significant ($p < .10$) at predicting holistic responses. The overall model, though, was not significant and the predictors were only able to explain 4% of the variance ($R^2 = .04$, $F(2,278)=1.46$, $p=.17$).

Another multiple regression analysis was conducted with the addition of three more predictors. These predictors were language, gender, and whether participants lived in a WEIRD or non-WEIRD region. This model was also not significant and only predicted 6%

of the variance (R^2 = .06, $F(21,265)$=.78, p=.75). The results indicated that only conscientiousness significantly predicted holistic responses at the p <.10 level.

Table 6: Multiple regression model 1

Model/Predictors	*b*	*CI 95% (Lower, Upper)*	*t-value*	*p-value*
Model 1				
Intercept	66.518	36.847, 96.190	4.413	.000
Condition	3.486	-0.589, 7.562	1.684	.093
AHS Score	0.482	-3.838, 4.801	0.219	.827
AQ Score	-1.323	-7.732, 5.087	-0.406	.685
EV	0.923	-1.555, 3.400	0.733	.464
AG	-1.105	-3.756, 1.547	-0.820	.413
CS	-2.400	-5.195, 0.398	-1.688	.093
NE	-1.043	-3.210, 1.124	-0.947	.344
OM	-2.400	-5.239, 0.440	-1.663	.097

Note. *p < .05, **p < .01. b = unstandardized regression coefficient. **Sample size = 296**

3.3.4 ANOVA

A two-way analysis of variance was conducted to inspect whether an interaction between region of current residence and instructional condition impacted the dependent variable of holistic responses. The response indicated that no significant effect of interaction was found ($F(5, 284) = 1.99$, $p < .08$).

Another two-way analysis of variance was conducted with reaction times as the dependent variable looking at culture by instruction condition (classification versus similarity judgment). The results indicated a main effect of reaction time by culture group (WEIRD vs.

non-WEIRD; $F(1, 292) = 17.49$, $p < .00$), such that the non-WEIRD population responded quicker across conditions. The interaction was nonsignificant.

3.3.5 AHS Scores by Region

Table 7: AHS Scores by Region

	N	*M*	*SD*	*Min*	*Max*
North America	128	4.84	0.47	3.67	6.21
Europe	52	4.96	0.47	4.13	5.96
Asia	82	5.01	0.52	3.92	6.17
Africa	22	5.21	0.54	4.25	6.67
Central America	3	5.13	0.29	4.79	5.29

3.3.6 Scales and Other Descriptive Statistics

Table 8: Descriptive Statistics

	N	*M*	*SD*	*Min*	*Max*
Age	296	25.19	9.40	18	65
English Proficiency	296	2.91	0.29	2.00	3.00
AHS Score	296	4.94	0.50	3.67	6.67
AQ-10 Score	296	2.71	0.33	1.80	4.00
Agreeableness	296	3.53	0.81	1.67	5.00
Conscientiousness	296	3.20	0.81	1.33	5.00

Extraversion	296	2.91	0.86	1.00	5.00
Negative Emotionality	296	3.22	1.04	1.00	5.00
Open Mindedness	296	3.69	0.74	1.33	5.00

3.3.7 Correlation Matrix

A correlation matrix (Figure 10) was computed depicting correlations among AHS scores, AQ-10 scores, the five BFI-2-XS domains, and percent holism from the task (dependent variable). The five BFI-2-XS domains are agreeableness (AG), conscientiousness (CS), extraversion (EV), negative emotionality (NE), and open mindedness (OM). Internal reliability was also calculated between the eight items. The Cronbach Alpha value of $\alpha = -.03$ indicates that the eight measures used were not closely related.

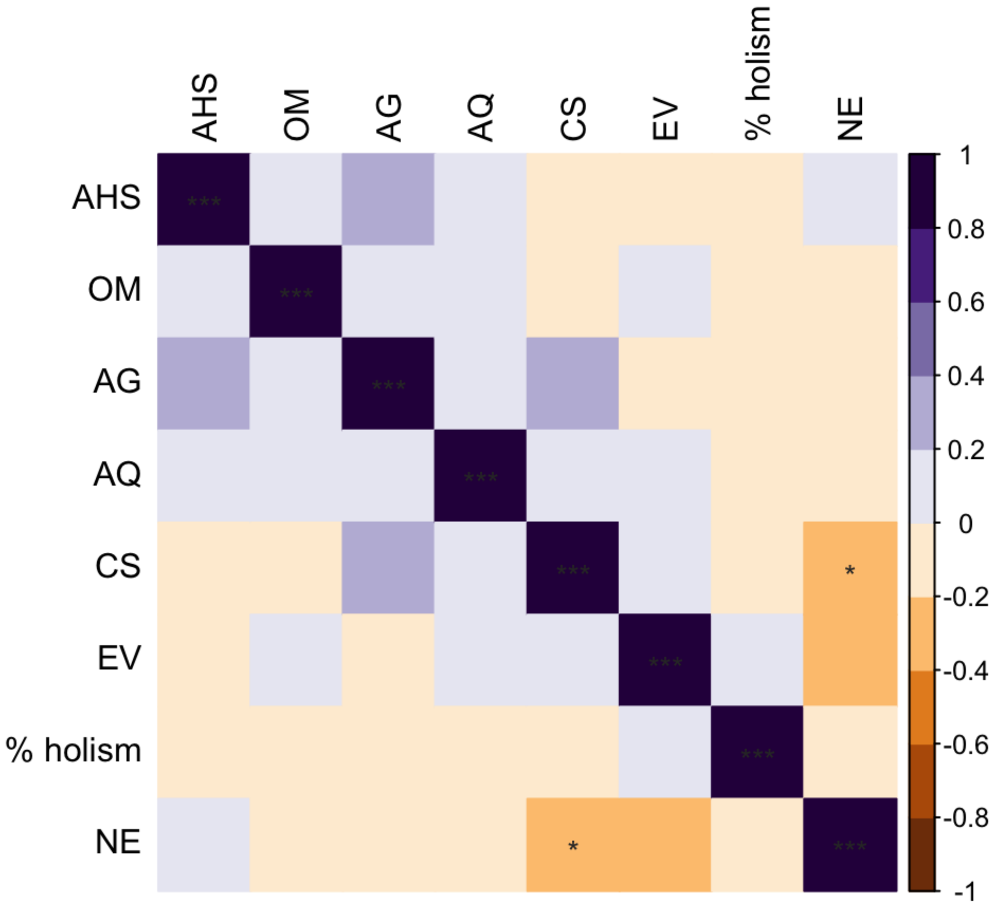


Figure 10: Correlation matrix showing all predictors and the dependent variable.

3.4 Discussion

This study aimed to replicate the task originally carried out by Norenzayan et al. (2002). This task has been widely used in recent categorization and culture literature. Although Norenzayan and colleagues (2002) found both an effect of instructional condition and an effect of culture, these effects have not always replicated (Murphy et al., 2017). Consequently, the primary goal of this research was to investigate the effects of the two instructional conditions administered in the task, as well as explore any effects of culture. The task consisted of 20 trials of a target object presented alongside two groups of four objects each. Participants were asked to categorize the target object as being in either Group 1 or Group 2. The stimuli were constructed such that neither group would be incorrect, but rather would reflect different categorization strategies, namely an analytic or a holistic strategy. The task followed a between-subjects design that randomly assigned participants to either the similarity or the classification condition. Keeping all else the same, the two conditions differed only in the task instructions. The similarity condition instructed participants to classify the target object based on which group it was the "most similar to", whereas the classification condition instructed participants to classify the target object based on which group it "belonged to". Being a relatively unexplored domain, factors other than culture that may give rise to a specific thinking style have seldom been explored. Thus, a secondary aim of this research was to test additional predictors in hopes to better account for the variance observed in analytic versus holistic responses to the task. This led to the inclusion of exploratory measures such as the AHS scale, AQ-10 scale, and BFI-2-XS scale.

The study was completed in a single session and all data was collected online. Participants' responses were quantified as a percentage of holistic responses across all trials. The subject pool varied between university students in North America and participants recruited on Prolific, thus being a moderately diverse sample. Though the original study had close to equal groups of East Asian and North American participants, it is worth noting that this was not achieved in this replication. Rather, the majority of participants were from North America, with Asia and Europe following. This may have limited the comparison of groups based on cultural differences, since both Europe and North America are typically considered WEIRD populations. That being said, our overall sample, including North Americans and

Europeans, showed higher holism than expected on the AHS scale (Table 7). This contrasted our predictions - suggesting that European Westerners may not be as analytic as suggested in previous literature. While the AHS was incorporated as an accompanying culture measure for insight into thinking styles, it is still considered a reasonably novel scale.

The results reveal that participants across conditions did not vary greatly on classifying based on a unidimensional rule and family resemblance. This contradicts the findings by Norenzayan et al. (2002) where participants were found to overwhelmingly prefer the unidimensional rule across conditions (M = 67% vs. M = 33%). However, our findings align with the results by Murphy et al. (2017) showing similar rates of rule-to-family resemblance categorization. Moreover, Figure 8 indicates that a range of analytic and holistic responses were observed, rather than large clusters of one or the other. Also aligning with Murphy et al. (2017), the classification task instructions did not elicit dominant rule-based responses that were seen by Norenzayan and colleagues (2002). When inspecting the differences between the similarity and classification conditions on holistic responses, higher holistic responses were seen in the similarity condition (M = 51.32%) across culture groups. This trend was expected, as similarity instructions were predicted to encourage holistic responding. The classification instruction condition, on the other hand, showed higher rule responding (M = 47.55%) in the overall sample. Though trending in the expected directions, these categorization response differences between instructional groups were not significant. Nonetheless, despite this null result, we cannot confidently claim that task instructions had no effect on the differences in categorization responses.

Two ANOVAs were conducted, one looking at holistic responding and the other using reaction times as the dependent variable. Both ANOVAs found no effect of an interaction between culture and condition. A main effect of reaction time and culture was found, indicating that non-WEIRD samples were faster to respond across conditions. This result was not detected by Norenzayan et al. (2002), and instead a main effect of reaction time by similarity judgment condition was found. This discrepancy is justified since a significant effect of instructional condition was not found in our sample.

Following analyses done by both Murphy et al. (2017) and Norenzayan et al. (2002), each instruction condition was further analyzed within culture groups. To do this, the sample was grouped by culture (WEIRD vs. non-WEIRD). Then, response differences within the classification and similarity conditions were inspected for each group. Comparing WEIRD versus non-WEIRD samples in each instruction condition, only WEIRD samples showed an effect of task instruction, with the similarity condition eliciting 52.3% holistic responding compared to 45.4% in the classification condition. WEIRD samples showed slightly higher rule based responding in the classification condition and higher family resemblance judgments in the similarity condition. This was different from the non-WEIRD sample, for which holistic responding was alike across conditions. Unsurprisingly, the non-WEIRD sample showed higher family resemblance responding in the classification condition (49.7%) than WEIRD participants (45.4%). Accordingly, it can be concluded that the task instructions had a higher effect on WEIRD samples. This was not the case for non-WEIRD samples, who were not as impacted by task instruction and equally used analytic and holistic strategies in categorizing. This demonstrates that the task instructions differentially impacted the strategy used among diverse culture groups. It is noteworthy to include that the percentage of holistic responses among the non-WEIRD and WEIRD populations showed a similar range, thus distributed along the spectrum in both culture groups (Figure 9). However, though both groups showed a spread of responding, the non-WEIRD population generally seem to show higher holistic responses from the baseline (Figure 9). These results are thought-provoking, especially given the divergence from both Norenzayan et al. (2002) and Murphy et al. (2017). It appears our population did not vary too much, showing a range of analytic and holistic responding regardless of condition and culture group. That said, it is difficult to make a claim about defaulting thinking styles by a particular culture group. Some limitations to consider are the uneven distribution of culture groups, the study requiring English fluency and comprehension, and the lack of considering language effects.

These findings contradict those by Norenzayan et al. (2002), where East Asian samples were found to prefer holistic responses for similarity-based judgments and American samples showed higher analytic responses across both conditions. Findings by Norenzayan et al. (2002) exhibited that although European Westerners responded analytically in the classification condition, they also responded analytically in the similarity judgment condition (*M*= 69%). On the contrary they found East Asians responded more holistically in the similarity judgment condition (*M*= 59%), while showing higher analytic responding in the classification condition. Not only did we see no effect of task instruction, but we also found no preference for holistic responses across conditions for non-WEIRD samples. Further, WEIRD samples did not show any strong preference for rule-based responses, and actually displayed higher family resemblance responding in the similarity instruction condition compared to non-WEIRD samples. Though Murphy et al. (2017) found similar rates of family resemblance to rule responding across conditions, their results also contradict with ours since no effect of instruction was found among their North American sample. On the contrary, their Korean samples showed higher sensitivity to the task instruction, which is the opposite of our findings.

From these comparisons, the question of whether WEIRD subjects are always rule-based and logical is raised. While Norenzayan et al. (2002) supported this claim, our results suggest that this may not be the case. Our results also suggest that there may not be much of a culture difference, but more likely one of task instruction instead. This prompts further research into the specific paradigms, stimuli, and task instructions utilized when investigating cross-cultural differences in thinking styles. Additionally, the finding that WEIRD samples were more impacted by task instruction stands out. This is because both Murphy et al. (2017) and Norenzayan and colleagues (2002) found the opposite, which warrants a discussion of conceivable reasons for this divergence.

Some probable reasons for the different findings are as follows. First, our task was completed entirely online, with no experimenter supervision or guidance. However, considering that only participants who completed all attention checks accurately were included in the analysis, it is unlikely that this would drastically impact the results. Second, our sample included participants beyond just North America and East Asia. Despite North American and East Asian participants being included, participants from other parts of Asia, Europe, Africa, and Central America were also sampled. Ergo, participants were grouped into WEIRD and non-WEIRD groups, based on a broad sorting of their region of residence. From this, the WEIRD group included both Europe and North America, and the remaining regions were placed in the non-WEIRD group. This was not a precisely determined process, which yields to possible error in distinguishing WEIRD from non-WEIRD populations. It is also likely that extending the sample beyond just East Asian and North American cultures may have enabled the different results observed. Furthermore, unequal spreads of participants were collected from different regions, with North America being in the majority. A larger, more balanced sample would allow for individual cultures to be analyzed against one another in an unbiased way, which may alter the results. Furthermore, given that the difference between conditions showed the expected trend while nearing significance, it is possible that the study was underpowered or lacking a more culturally representative sample. In this way, a larger, more representative sample, such as seen in the research by both Norenzayan et al. (2002) and Murphy et al. (2017) could elicit significant group differences. Lastly, the study required English fluency since it was conducted entirely in English. Thus, having not considered any effects of language may have also impacted the results.

Finally, the regression models indicate that much of the variance explaining holistic responses was not accounted for with the predictors used. This may be a result of the seemingly lack of variance in responding to the task. Upon inspection of the range of the data points, participants did not vary drastically in holistic responding by culture, stimuli, or by task instruction. However, a more representative sample may be required to account for variance by WEIRD or non-WEIRD cultural influences. An alternate explanation is that other unaccounted factors may be at play in studies showing differences in categorization strategy responses.

Future directions include collecting a larger, more balanced sample to allow for stronger analyses by precise culture/region. Additionally, including more specific options for culture in the survey, as opposed to nonspecific selections of “Asia”, “Europe”, and “Africa”, would increase accuracy of culture judgments. This is because the varying factors underlying WEIRD demographics can differ greatly over larger areas. Additionally, when considering populations as WEIRD and non-WEIRD groups, future studies should gauge factors beyond culture such as levels of education, industrialization, wealth, and democracy. Reasoning for this lies in the belief that elements beyond culture may the influence thinking styles of diverse populations. Also, seeing as holistic versus analytic response rates were nearly even across cultures and conditions, implementing methods allowing participants to explain their categorization rationale following the task can protect against guessing. Lastly, investigating other manipulations and task instruction effects such as predictor variables of impulsivity and/or time measures may show interesting observations. These would offer insight into the mindset from which participants are making categorization decisions to inspect whether quick judgments versus slower ones encourage different thinking styles.

In conclusion, the results of this study for the most part did not replicate those of Norenzayan et al. (2002). Though some similarities were found among our results and those of Murphy et al. (2017), some key differences were also revealed. As a result, more work in this domain is required to uncover and understand the diverging results observed.

Chapter 4

4 Overall Discussion

4.1 Discussion

The domain of classification has been extensively studied over decades, leading to the proposal, testing, and revelation of various models and principles of classification. It is widely believed that the basic principles of classification are universal, in that the fundamental mechanisms by which categories are learned and used are shared by all humans (Smith et al., 2012). However, the claim of universal classification centers on research that has narrowly focused on WEIRD populations. As a result, recent cross-cultural classification research has contested this commonly held belief, positing that classification and thinking styles are dependent on culture and/or language factors (Henrich et al., 2010; Norenzayan et al., 2002; Nisbett & Miyamoto, 2005; Masuda & Nisbett, 2001). Given that Westerners are the predominantly studied population, the lack of insights from East Asian cultures, among other non-WEIRD populations, warrants our exploration of classification universals.

Classification is a broad research area and consists of category learning and categorization as major components. Though widely studied among Western populations, the influence of culture on ways in which humans learn and use categories remains reasonably unexplored. Classification literature has compared two distinct styles of thinking: analytic versus holistic thinking. Analytic thinking follows formal logic, is rule-based and context-independent, and focuses on a focal object/feature. On the contrary, holistic thinking is similarity-based, context-dependent, and focuses on object-feature relationships rather than a focal feature. Two classification studies were conducted to investigate the relationship between culture and thinking styles. The first study examined category learning using the SHJ category sets and the AHS. The aim of this study was to assess for differences in thinking as a novel category set is learned, and to observe variances as the optimal strategy to learn categories changes. Experiment 2 replicated the categorization task by Norenzayan et al. (2002) with the addition of exploratory scales, namely, the AHS, BFI-2-XS, and the AQ-10. Alongside exploring potential mediator and mediating variables, the aim of this study was to observe any correlations between culture, categorization strategies, and task instructions. Data for both studies was collected online, though experiment 1 was also administered in person. The demographics questions were kept mostly the same for both projects and the same platforms to carry out the task and the questionnaire components were used across studies.

4.1.1 The Tasks

This research looked at family resemblance and rule-based responding to gain insight into analytic versus holistic thinking styles. Each study inspected distinct aspects of family resemblance and rule-based responding, which was highlighted by the different classification task types. Experiment 1 was a learning task, thus feedback was provided immediately upon categorization to indicate whether a correct or incorrect decision had been made. The SHJ stimuli had a predetermined category membership based on the specific structure of the particular category set. Category membership was determined by learning the optimal strategy underlying a specific category structure, which could be achieved as participants completed the task. The optimal strategy varied for each of the six task Types, with rule-based and family resemblance strategies being among the six. Participant responses were observed across all blocks to recognize the learning trajectory for their assigned task Type. The second task, which was based on the study by Norenzayan et al. (2002), was vastly different. This study implemented a forced-choice categorization paradigm without a learning component. Feedback upon categorization was not provided in this study, which was in line with the nature of the task and the stimuli structure. The categorization response correctness of the stimuli used here could not be evaluated because the target item had no pre-determined category membership by design. Instead, the group that the target was placed in depended on the strategy used, with one group eliciting a rule-based strategy and the other prompting family resemblance responding.

In this way, each study was designed to explore different aspects of family resemblance and rule-based responding. This allowed for breadth in examining the broad area of culture and classification, providing added insight into any potential effects. The range covered by these preliminary tasks was beneficial for assessing the value in exploring culture and classification in its entirety, as it offered general directions for both the present and for future exploration on a larger scale.

4.1.2 Summary of Findings

The results of experiment 1 were promising and generally showed predicted effects. Expected rank-order effects were observed in experiment 1, indicating that Types I through VI varied in task difficulty with Type I being the easiest and Type VI the most difficult. Thus, it can be drawn that the optimal strategy to learn each category differed as a function of task complexity. In addition, a significant correlation was found between the Type IV category set and AHS scores. Since the optimal strategy to learn the Type IV set was a family resemblance one, this indicated that higher AHS scores (higher holism) were related to the family resemblance strategy. On the contrary, the rule-based tasks (Types I and II) showed a trend in the negative direction with AHS scores (low holism). Although this trend requires further investigation, it aligned with our predictions and was encouraging. The exploratory regression analyses showed that AHS scores were significantly predicted by region of residence, particularly for Type IV. This served as validation for the AHS scale, though more research is warranted to confirm this claim. Finally, task performance by region was also explored but these findings were tough to comment on given the small sample sizes recruited outside of North America.

The results of experiment 2 were interesting despite the observations not matching the expected trends. The effect of instructional condition was nearing significance, indicating that though not definitive, a noteworthy effect seems promising. Since task instructions typically vary across classification studies, the results suggest that this may be a confounding variable requiring more attention. Accordingly, exploring diverse task instructions in depth with the added study of culture is recommended. Further, it is possible that language plays a role in effects of instructional differences, which is another insightful area worth exploring. Another fascinating result was the range of holistic and analytic responding across all populations tested in each instructional condition. This was different from Murphy et al. (2017) and Norenzayan et al. (2002), whereby there was no single dominant thinking style among our participants, but rather a close-to-even split of analytic and holistic thinking.

However, Murphy et al. (2017) found a similar range of responding differences between Asian and American subjects in the similarity instructional condition. This begs the question of whether thinking style differences are dependent on culture and to what extent, and if other factors may have a considerable impact. While Murphy et al. (2017) and Norenzayan et al. (2002) found that Asian but not American subjects saw different styles of thinking as a function of instructional condition, the opposite was seen here. With the caveat being that our samples were grouped broadly into WEIRD versus non-WEIRD populations, our WEIRD (American and European) samples were observed to show differential thinking as a function of instruction. Moreover, no interaction of culture and instructional condition was found, which aligned with findings by other researchers. It was interesting that the regression model, though insignificant, predicted very little variance. This prompts an investigation of whether the variance observed followed a trend at all, or if responses were more random. Lastly, because the model did not fit much of the variance seen in the results, many of the exploratory scales did not depict any effects worth discussing.

4.1.3 Conclusions

The specific goals of each study varied quite a bit, while the broader research question remained the same. Consistent family resemblance effects were observed in experiment 1, showing a relationship between learning and thinking style. An effect of region was also found on performance and on AHS scores, which was interesting. Although findings from experiment 2 were less definitive, the results encouraged further insight into this domain. Overall, across studies it can be drawn that a strong effect of culture may not be present when learning and using categories. However, classification performance does seem to be impacted by thinking style and potentially other related factors. Consequently, the effect of culture on classification responses cannot be ruled out. Further, while we cannot confidently claim whether an effect of culture does or does not exist, it seems the effect of thinking style may occur based on task circumstances. This is drawn as only some category set Types showed an effect of thinking style on task performance, namely the family resemblance category structure. Also, the categorization task showed a more likely effect of instructional condition than culture, which is noteworthy. Though both paradigms warrant further research, no overpowering effects of culture were seen throughout. Overall, the results indicated that thinking styles impacted some aspect of classification performance, thus suggesting that there may exist a relationship between language, culture, and geography. The strength, direction, and other predictors involved in these relationships are still unknown, however, further exploration is warranted. We draw that culture, region, language, and other related factors are worth exploring when interpreting the results of classification studies given that thinking style, which relates to performance, may be impacted by these influences. Although the literature has indicated that thinking styles may be driven by these influences, this finding was not as prominent in our results.

Furthermore, the AHS scale measures one's general tendency to think analytically or holistically based on the four overarching factors of causality, perception of change, attention, and locus of control. Despite being largely North American, our sample was not an analytic one – with AHS scores falling between mid to high holism. Nonetheless, the relationship between performance and region when exploring the effect of the instructional manipulation stood out. Though this effect was not strong, it was significant, thus drawing future attention to this. The AHS scale was also administered in both studies, however, the lack of a largely representative sample makes it difficult to confidently validate the scale. In sum, the results across studies are preliminary yet encouraging.

4.2 Limitations & Future Directions

Limitations across both studies include gathering a more balanced and culturally diverse sample, recruiting a larger sample, and accounting for language effects. A goal of this research was to compare culture groups, which was made difficult given the uneven spread of the diverse groups collected. North Americans were the majority culture sampled, with more than double the number of participants than any other culture. To ensure sufficient group sizes and power when dividing groups into WEIRD versus non-WEIRD comparisons, culture groups were combined. Having a more balanced, culturally diverse sample would allow for even, adequately sized groups to directly compare European Westerners to East Asian populations. Also, given that both studies followed a between-subjects design, collecting a larger sample would be beneficial to increase the statistical power. Another limitation was the imbalanced online to in person data collection ratio. Although the results showed expected trends, which indicated that reliable data can be collected online, we were unable to compare these groups directly. This was a result of there being substantially more data collected online than in person. Online data collection presents limitations of its own, such as the inability to explicitly enforce the task instructions. Moreover, with the great accessibility of translator applications, it is unknown whether participants viewed task instructions in English. It is important to note, nonetheless, that both experiments required participants to be fluent in English to understand the task instructions and complete the questionnaires effectively. This presents a limitation when collecting cross-cultural data, since the English proficient pool of individuals may indicate a sub-population within the broader sample of that culture. If that is the case, then this sub-population may present different results than otherwise observed by those not fluent in English. To account for the way in which language shapes thinking, it would be interesting to collect data in participants' native languages. Lastly, it is important to consider whether our tasks were sensitive in capturing effects of culture and thinking styles. Experiment 2 used has been widely administered by other researchers and has shown effects of culture in previous research. Thus, despite some mixed results, there is reason to believe that thinking styles can be assessed through this task. Conversely, the SHJ task has not been previously used to assess effects of culture, nevertheless it has been effective in evaluating learning performance of varying category structures.

Despite the overarching research question being largely unanswered, our results motivate a deeper dive into the domain of culture and classification. The preliminary nature of the two studies discussed suggests that more promising insights may become apparent by addressing the limitations faced. In addition, experiment 1 was conducted on a smaller scale to assess the feasibility of collecting reliable cross-cultural data. Having concluded that cross-cultural data can be reliably collected online, an appropriate next step would be to run this study on a larger scale, using the necessary translations and collecting data from a larger, balanced cross-cultural sample. Other limitations include accounting for language effects, collecting a larger, more diverse sample, and narrowing the focus on East Asian versus American cultures.

The WEIRD population encompasses Western, educated, industrialized, rich and democratic societies, of which only Western samples have been a key focus. Beyond culture, education level, socioeconomic status, political orientations, etc. are core aspects of society that may contribute to certain thinking styles. These components, grouped under the umbrella term *WEIRD*, should be considered for the impact on the perception of objects within the environment. It is also believed that aspects of language influence categorization, particularly regarding relational groupings (Markman & Hutchinson, 1984). Given the interconnectedness of language and culture, isolating one from the other is incredibly challenging. Accordingly, studying the effects of language and culture on classification with the addition of related predictors such as bilingualism, immigration, and acculturation would be valuable. Predictors beyond these should also be explored. For example, COVIS discusses that rule-based responding is immediate, thus being quicker than the more gradual similarity-based responding. In this way, the strategy used when making classification decisions may be impacted by the amount of time one has when making judgments. Thus, exploring a manipulation of time when completing classification tasks could be fascinating.

Another related, yet underexplored area of categorization is that of thematic and taxonomic concepts. Conventionally, categorization has been defined taxonomically, such that items within a category share core features and some functional, biological, or perceptual properties (Lin & Murphy, 2001; Medin & Smith, 1984). Simply put, taxonomies represent a hierarchical system of concepts separated by varying specificity. In this system, the higher in the taxonomy, the broader the requirements are for class inclusion in the category (Levitin, 2002). Taxonomic categorization follows a logical, rule-based structure requiring necessary and/or sufficient features. This is comparable to analytic thinking in that grouping is not context-dependent, but rather relies on defining features for category inclusion. However, it has been argued that this method of categorization fails to capture the depth, richness, and breadth of human concepts and categorization (Lin & Murphy, 2001). Thus, it has become more commonplace to investigate categorization with regards to both taxonomic and thematic relations. Thematic categories consist of items grouped conceptually, without any internal or physical resemblance or without similar functions (Lin & Murphy, 2001). Instead, items thematically grouped together may hold complementary relations, requiring the consideration of context and external relations, resembling holistic thinking. It is important to note that thematically grouped items are often not of the same kind, whereas taxonomically grouped ones are always of the same kind (Markman & Hutchinson, 1984). For example, while a taxonomic category would consist of different species of dogs, or more broadly, animals, a thematic category may consist of a dog and leash – two functionally related, complementary, entities.

Forced-choice tasks are commonly used to investigate thematic versus taxonomic category formation, whereby participants are given a target to pair with one of two options. These are similar to categorization paradigms, however, in this case the two options would differ in being taxonomically or thematically related to the target item. Just like investigating analytic versus holistic thinking, task instruction, stimuli, and other paradigm-related factors may influence the strategy used. Recently, variations among groups have been attributed to culture differences. Although the distinction between thematic versus taxonomic categorizations have been explored more among children, Chiu (1972) proposed that a cultural basis may underlie these differences. Though related to the general understanding of analytic versus holistic thinking, the exploration of taxonomic and thematic relations was beyond the scope of this paper. As an interesting and unexplored area of study, it is recommended that future research explores this domain further, potentially tying it together with classification studies.

4.3 Conclusion

Experiment 1 found expected trends in category learning and the findings were consistent with our predictions pertaining to culture and family resemblance responding. Also, interesting exploratory trends were seen, particularly regarding the effects of region and task performance. The second study, replicating Norenzayan et al. (2002), showed some expected trends, but did not always align with the original study results. Though unexpected trends were not found, the results seemed inconclusive with regards to the direction and specifics of how culture plays into thinking styles. Taken together, it can be concluded that culture does impact classification task performance. Nevertheless, more research is needed to understand the underlying functions and relationships among culture, geography, language, and classification.

References

Alaei, M. M., Ahmadi, M., & Zadeh, N. S. (2014). The Impact of Rater's Personality Traits on Holistic and Analytic Scores: Does Genre Make any Difference too? *Procedia - Social and Behavioral Sciences*, *98*, 1240–1248.

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology. General*, *120*(1), 3–19.

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward Brief "Red Flags" for Autism Screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist in 1,000 Cases and 3,000 Controls. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(2), 202–212.e7.

Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*(3), 442–481.

Ashby, F. G., & Ell, S. W. (2001). The neurobiology of human category learning. *Trends in Cognitive Sciences*, *5*(5), 204–210.

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147.

Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality. *European Journal of Personality*, *15*(5), 327–353.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Chiu, L.-H. (1972). A Cross-Cultural Comparison of Cognitive Styles in Chinese and American Children. *International Journal of Psychology: Journal International de Psychologie*, *7*(4), 235–242.

Choi, I., Koo, M., & Jong An Choi. (2007). Individual differences in analytic versus holistic thinking. *Personality & Social Psychology Bulletin*, *33*(5), 691–705.

Choi, I., Nisbett, R. E., & Norenzayan, A. (n.d.). Causal attribution across cultures: Variation and universality. *Psychological Bulletin*, *125*(1), 47–63.

Choi, & Markus. (1998). Implicit theories and causal attribution East and West. *Unpublished Manuscript, University of Michigan*.

de Oliveira, S., & Nisbett, R. E. (2017). Beyond East and West: Cognitive Style in Latin America. *Journal of Cross-Cultural Psychology*, *48*(10), 1554–1577.

Digman, J. M. (1997). Higher-order factors of the Big Five. In *Journal of Personality and Social Psychology* (Vol. 73, Issue 6, pp. 1246–1256). https://doi.org/10.1037/0022-3514.73.6.1246

Ell, S., & Zilioli, M. (2012). Categorical Learning. *Seel, Norbert M.(ed.), Encyclopedia of the Sciences of Learning, Springer US,* 509-512.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology. General*, *127*(2), 107–140.

Evans, J. S. B. T., & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction. *Psychological Review*, *103*(2), 356–363.

Folstein, J. R., & Van Petten, C. (2004). Multidimensional rule, unidimensional rule, and similarity strategies in categorization: event-related brain potential correlates. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(5), 1026–1044.

Gellatly, I. R. (1996). Conscientiousness and task performance: Test of cognitive process model. *The Journal of Applied Psychology*, *81*(5), 474–482.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*(1), 21–38.

Graziano, W. G., Hair, E. C., & Finch, J. F. (1997). Competitiveness mediates the link between personality and group performance. *Journal of Personality and Social Psychology*, *73*(6), 1394–1408.

Graziano, W. G., Jensen-Campbell, L. A., & Hair, E. C. (1996). Perceiving interpersonal conflict and reacting to it: the case for agreeableness. *Journal of Personality and Social Psychology*, *70*(4), 820–835.

Hadingham. (1994). The mummies of Xinjiang. *Discover*. https://elibrary.ru/item.asp?id=2100927

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences*, *33*(2-3), 61–83; discussion 83–135.

Homa, D. (1973). Prototype abstraction and classification of new instances as a function of number of instances defining the prototype. In *Journal of Experimental Psychology* (Vol. 101, Issue 1, pp. 116–122). https://doi.org/10.1037/h0035772

Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of Personality*, *69*(2), 323–361.

Ji, L. J., Peng, K., & Nisbett, R. E. (2000). Culture, control, and perception of relationships in the environment. *Journal of Personality and Social Psychology*, *78*(5), 943–955.

Ji, L.-J., Zhang, Z., & Nisbett, R. E. (2004). Is it culture or is it language? Examination of language effects in cross-cultural research on categorization. *Journal of Personality and Social Psychology*, *87*(1), 57–65.

Kemler Nelson, D. G. (1984). The effect of intention on what concepts are acquired. *Journal of Verbal Learning and Verbal Behavior*, *23*(6), 734–759.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Development and Psychopathology*, *13*(1), 111–124.

Knight, N., & Nisbett, R. (2007). Culture, Class and Cognition: Evidence from Italy. *Journal of Cognition and Culture*, *7*(3-4), 283–291.

Levitin, D. J. (2002). *Foundations of Cognitive Psychology: Core Readings*. MIT Press.

Lewandowsky, S. (2011). Working memory capacity and categorization: individual differences and modeling. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(3), 720–738.

Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology. General*, *130*(1), 3–28.

Lloyd, G. E. R. (1990). *Demystifying Mentalities*. Cambridge University Press.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*(1), 1–27.

Masuda, T., & Kitayama, S. (2004). Perceiver-induced constraint and attitude attribution in Japan and the US: A case for the cultural dependence of the correspondence bias. *Journal of Experimental Social Psychology*, *40*(3), 409–416.

Masuda, T., & Nisbett, R. E. (2001). Attending holistically versus analytically: comparing the context sensitivity of Japanese and Americans. *Journal of Personality and Social Psychology*, *81*(5), 922–934.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, *35*, 113–138.

Mercado, E., 3rd, Chow, K., Church, B. A., & Lopata, C. (2020). Perceptual category learning in autism spectrum disorder: Truth and consequences. *Neuroscience and Biobehavioral Reviews*, *118*, 689–703.

Miller, J. G. (1984). Culture and development of everyday social explanation. *Journal of Personality and Social Psychology*, *46*(5), 961–978.

Minda, J. P., Desroches, A. S., & Church, B. A. (2008). Learning rule-described and non-rule-described categories: a comparison of children and adults. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *34*(6), 1518–1533.

Minda, J. P., & Miles, S. J. (2010). The influence of verbal and nonverbal processing on category learning. In B. H. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 52, pp. 117–162). Academic Press.

Morgan, E. L., & Johansen, M. K. (2020). Comparing methods of category learning: Classification versus feature inference. *Memory & Cognition*, *48*(5), 710–730.

Morris, M. W., & Peng, K. (1994). Culture and cause: American and Chinese attributions for social and physical events. *Journal of Personality and Social Psychology*, *67*(6), 949–971.

Murphy, G. L., Bosch, D. A., & Kim, S. (2017). Do Americans have a preference for rule-based classification? *Cognitive Science*, *41*(8), 2026–2052.

Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of category learning: Fact or fantasy? *Psychology of Learning and Motivation*, *54*, 167.

Nisbett, R. E., & Miyamoto, Y. (2005). The influence of culture: holistic versus analytic perception. *Trends in Cognitive Sciences*, *9*(10), 467–473.

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological Review*, *108*(2), 291–310.

Norenzayan, A. (1999). *Rule-based and experience -based thinking: The cognitive consequences of intellectual traditions* (R. E. Nisbett (ed.)) [University of Michigan]. https://www.lib.uwo.ca/cgi-bin/ezpauthn.cgi?url=http://search.proquest.com/dissertations-theses/rule-based-experience-thinking-cognitive/docview/304517623/se-2

Norenzayan, A., & Heine, S. J. (2005). Psychological Universals: What Are They and How Can We Know? *Psychological Bulletin*, *131*(5), 763–784.

Norenzayan, A., & Nisbett, R. E. (2000). Culture and Causal Cognition. In *Current Directions in Psychological Science* (Vol. 9, Issue 4, pp. 132–135). https://doi.org/10.1111/1467-8721.00077

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, *26*(5), 653–684.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*(1), 53–79.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. In *Journal of Experimental Psychology* (Vol. 77, Issues 3, Pt.1, pp. 353–363). https://doi.org/10.1037/h0025953

Rabi, R., & Minda, J. P. (2016). Category learning in older adulthood: A study of the Shepard, Hovland, and Jenkins (1961) tasks. *Psychology and Aging*, *31*(2), 185–197.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou (Ed.), *Similarity and analogical reasoning , (pp* (Vol. 592, pp. 21–59). Cambridge University Press, xiv.

Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, *4*(3), 328–350.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. In *Cognitive Psychology* (Vol. 7, Issue 4, pp. 573–605). https://doi.org/10.1016/0010-0285(75)90024-9

Ross. (n.d.). The intuitive psychologist and his shortcomings: Distortions in the attribution process. *Advances in Experimental Social Psychology*. https://www.sciencedirect.com/science/article/pii/S0065260108603573

Saalbach, H., & Imai, M. (2007). Scope of linguistic influence: does a classifier system alter object concepts? *Journal of Experimental Psychology. General*, *136*(3), 485–501.

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–1323.

Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, *75*(13), 1–42.

Smith, E. E., & Grossman, M. (2008). Multiple systems of category learning. *Neuroscience and Biobehavioral Reviews*, *32*(2), 249–264.

Smith, E. E., & Medin, D. L. (1981). *Categories and Concepts*. https://doi.org/10.4159/harvard.9780674866270

Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., Spiering, B., Beran, M. J., Church, B. A., Ashby, F. G., & Grace, R. C. (2012). Implicit and explicit categorization: a tale of four species. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2355–2369.

Smith, J. D., David Smith, J., & Minda, J. P. (1999). "Prototypes in the mist: The early epochs of category learning": Correction to Smith and Minda (1998). In *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Vol. 25, Issue 1, pp. 69–69). https://doi.org/10.1037/h0090333

Smith, J. D., Minda, J. P., & Washburn, D. A. (2004). Category learning in rhesus monkeys: a study of the Shepard, Hovland, and Jenkins (1961) tasks. *Journal of Experimental Psychology. General*, *133*(3), 398–414.

Soto, C. J., & John, O. P. (2017a). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, *68*, 69–81.

Soto, C. J., & John, O. P. (2017b). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143.

Spencer-Rodgers, J., & Peng, K. (Eds.). (2018). *The psychological and cultural foundations of east Asian cognition: Contradiction, change, and holism*. Oxford University Press.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*(3), 506–520.

Varnum, M., Grossmann, I., Katunar, D., Nisbett, R., & Kitayama, S. (2008). Holism in a European Cultural Context: Differences in Cognitive Style between Central and East Europeans and Westerners. *Journal of Cognition and Culture*, *8*(3-4), 321–333.

Ward, T. B., & Scott, J. (1987). Analytic and holistic modes of learning family-resemblance concepts. *Memory & Cognition*, *15*(1), 42–54.

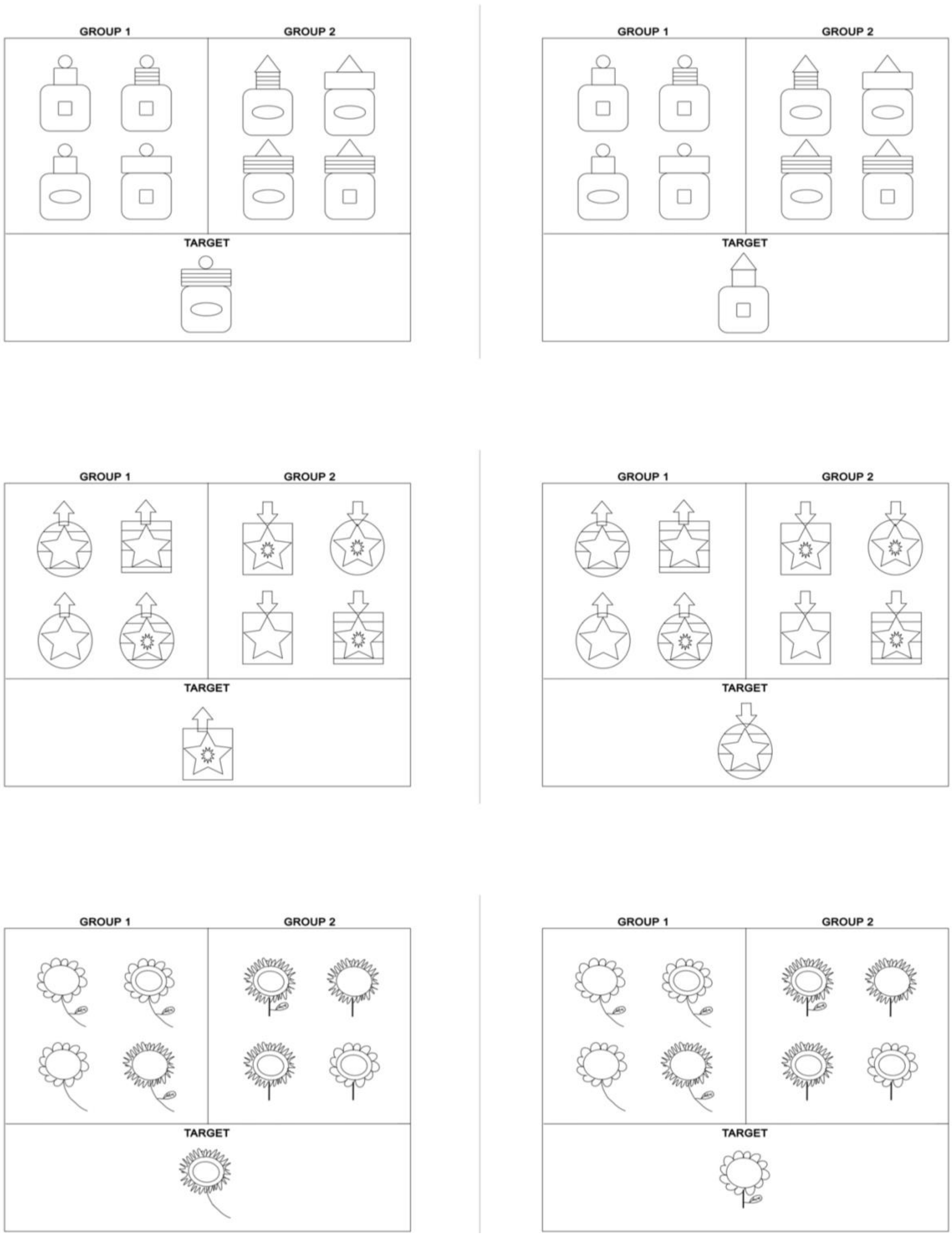
Wickens, J. (1990). Striatal dopamine in motor activation and reward-mediated learning: steps towards a unifying model. *Journal of Neural Transmission. General Section*, *80*(1), 9–31.

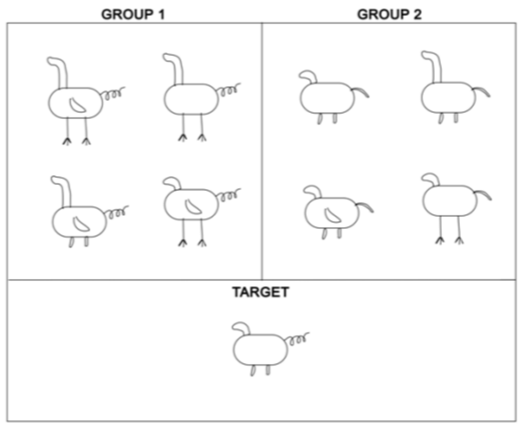
Wiggins, J. S., & Trapnell, P. D. (1997). Chapter 28 - Personality Structure: The Return of the Big Five. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of Personality Psychology* (pp. 737–765). Academic Press.

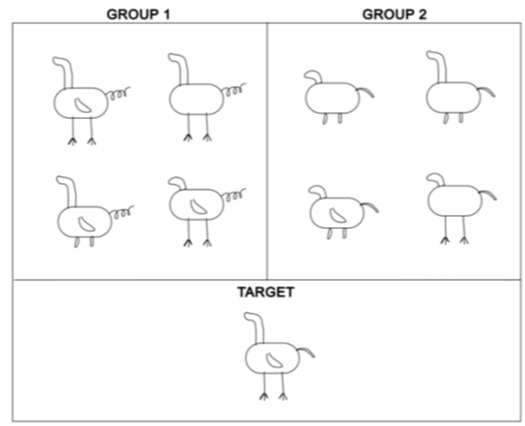
Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, *34*(2), 387–398.

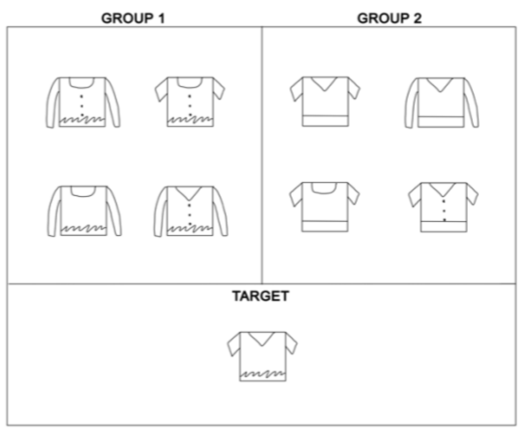
Appendices

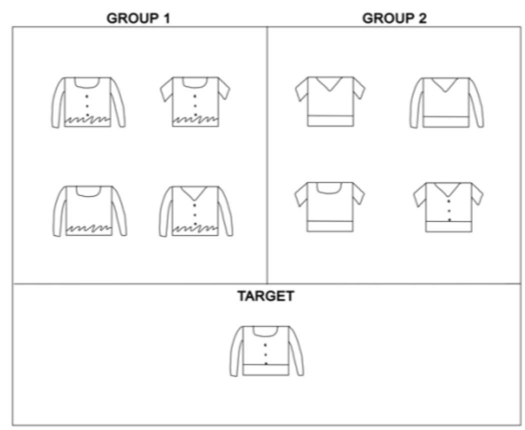
Appendix A: Experiment 2 stimuli.

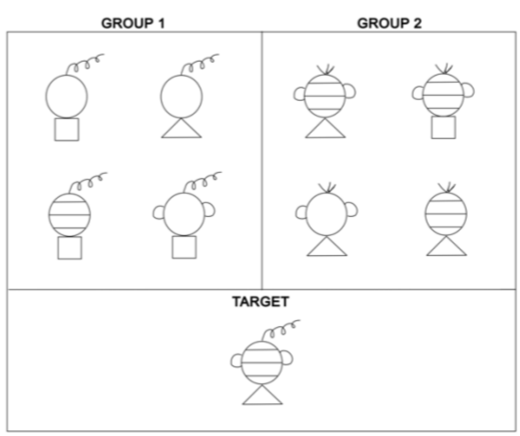


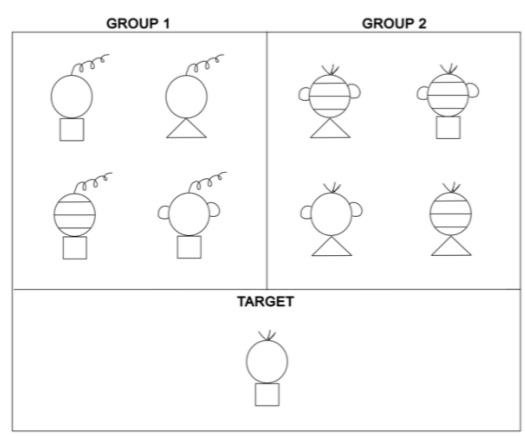

GROUP 1
GROUP 2
TARGET

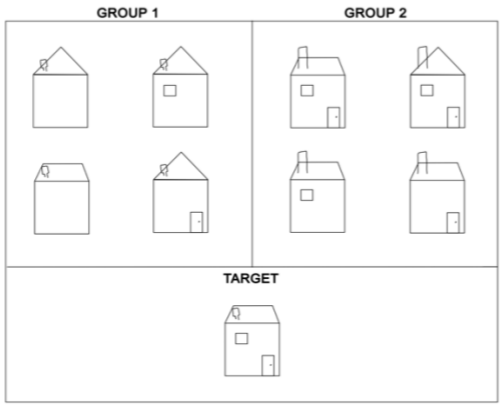

GROUP 1
GROUP 2
TARGET

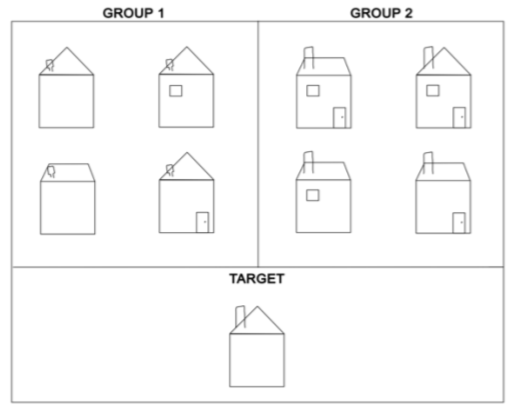

GROUP 1
GROUP 2
TARGET

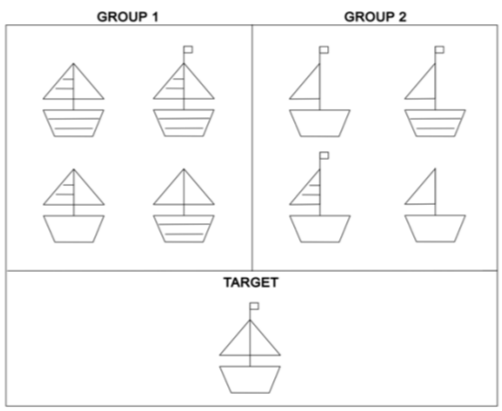

GROUP 1
GROUP 2
TARGET

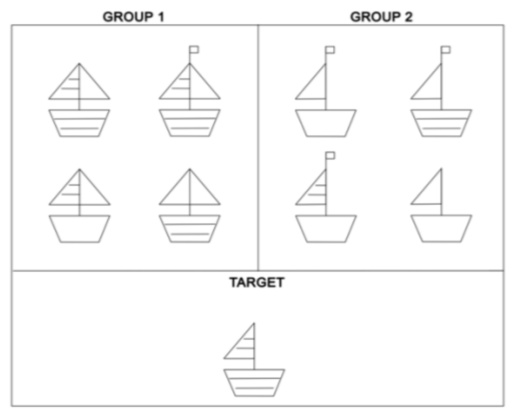

GROUP 1
GROUP 2
TARGET

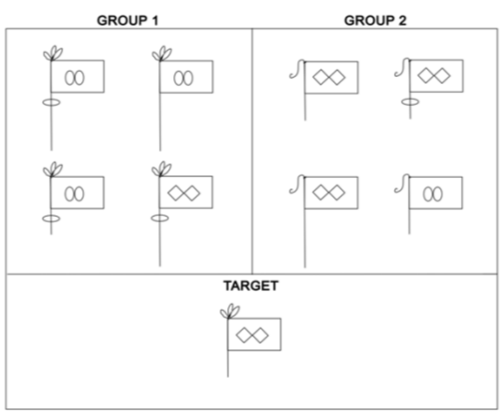

GROUP 1
GROUP 2
TARGET

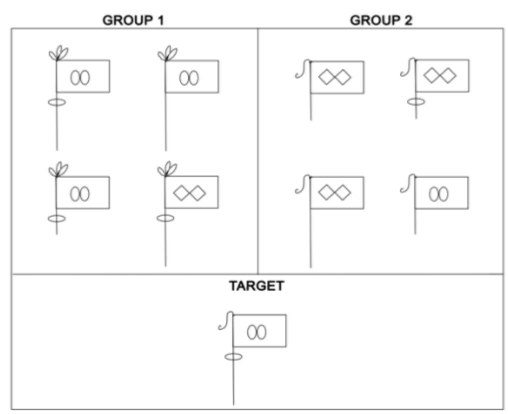

GROUP 1
GROUP 2
TARGET

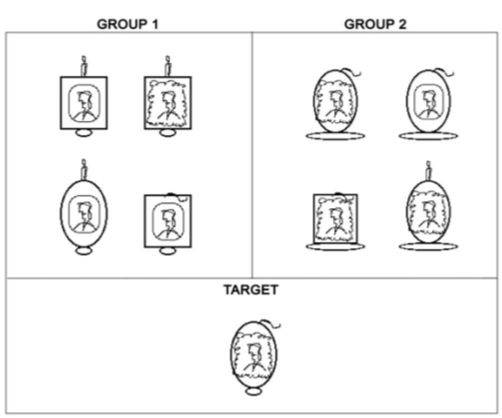

GROUP 1
GROUP 2
TARGET

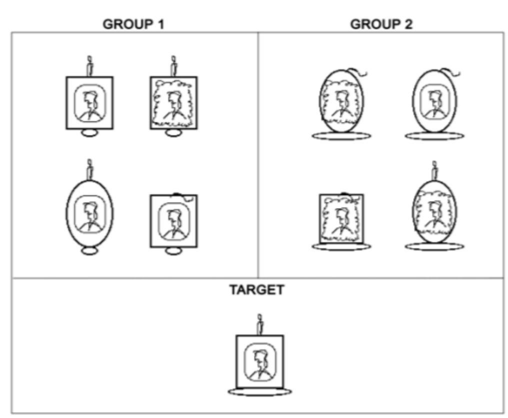

GROUP 1
GROUP 2
TARGET


GROUP 1
GROUP 2
TARGET


GROUP 1
GROUP 2
TARGET


GROUP 1
GROUP 2
TARGET


GROUP 1
GROUP 2
TARGET


GROUP 1
GROUP 2
TARGET