Holistic Processing: A Matter of Experience

By

Kao-Wei Chua

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY

In
Psychology
August 31, 2017
Nashville, Tennessee

Approved:
Isabel Gauthier, Ph.D.
Tom Palmeri, Ph.D.
Dan Levin, Ph.D.
James Tanaka, Ph.D.
ACKNOWLEDGEMENTS

I would like to thank everyone in the OPL who helped with collecting data on this project. Susan Benear was instrumental in helping to organize and juggle a hectic schedule that involved 12 in lab sessions with over 60 subjects. Anthony Rios, Parth Patel, and Mareike Eydt-Beebe all helped make what could have been a difficult data collection process one that was as organized and efficient as anyone could ask for. Members of the OPL like Rankin Williams, Magen Speegle, Jackie Floyd, and Mackenzie Sunday have provided incredible support over my graduate school tenure.

I would like to thank my committee members, Dan Levin, Tom Palmeri, and Jim Tanaka. Dan and Tom have been on all of my committees throughout my graduate career and their input on all my projects have been invaluable in shaping the way I think about data and presentation. Jim was always a positive presence at the Perceptual Expertise Network meetings and I went to his lab for a summer to help with his Face Camp. This introduced me to different kinds of outreach opportunities that scientists can undertake in the community.

Isabel Gauthier is an exemplary scientist and her focus on teaching and mentorship was one of the foremost things that struck me upon meeting her. During my visit to Vanderbilt as a prospective graduate student, I noticed that her office door was quite literally always open and she always had time for students. Rankin was working on her dissertation and had many questions, and Isabel made time for her questions amidst a busy schedule interviewing prospective students. It was then that I knew this was the right mentor and environment for me. Isabel carried through that dedication to mentorship with me and kept her office door open for me to barge through and ask (sometimes asinine) questions whenever I needed help. She has also been extraordinarily patient with some of my personal quirks. Sometimes I can’t see those shadows on figures! Sometimes I mistake 1s and 0s!

Through our engaging lab meeting discussions, the way I have read and thought about science have been considerably sharpened. Witnessing her working style, I saw a model that was aspirational. Throughout my graduate career, I have often found myself asking “how did Isabel even do that,” especially when it comes to her preternatural ability to get manuscripts back in days. I have not found cogent answers to that question beyond some physically intractable theories involving time compression. As a scientist and person, Isabel has been a strongly positive example for me over these six years. Whatever position I find myself in in life, I hope to carry her dedication to education and mentorship forward.

Finally, I thank my parents, Chu and Balvin Chua, who are literally and figuratively the reasons why I am here and doing this. They are both PhDs in Biology, and it seemed preordained that I would follow in their footsteps in some form. They would take me to their cardiovascular lab at a young age and all the strange, bubbling machinery was fascinating. Eventually, my father had me reading his manuscripts and paid me a quarter for each typo I caught, and it was not long until I gained an understanding of how scientific argumentation was structured. He also had me press the submit button on manuscripts that he felt too nervous about to submit himself. It was no big deal for an 8 year old! Working with my mother, Chu Chang Chua, I learned diligence and the dedication to get complicated procedures right. Seeing her beam with happiness after perfectly executing a protocol gave me an appreciation for science as a process, that there is value in completing each step of a project even if it “fails.” When I helped her write grants, we thought about evidence and put them together like puzzle pieces to gain a greater understanding of complicated systems. These moments have most strongly informed my decision to pursue science. She is no longer with us, but she is with me still.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapters</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapters</td>
<td></td>
</tr>
<tr>
<td>I.  Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Holistic Face and Object Processing</td>
<td>6</td>
</tr>
<tr>
<td>Previous Training Studies</td>
<td>15</td>
</tr>
<tr>
<td>Experience and Learned Attention</td>
<td>20</td>
</tr>
<tr>
<td>Aims of the Current Study</td>
<td>25</td>
</tr>
<tr>
<td>II. The Creation of Tests of Holistic Processing for Novel Objects</td>
<td>27</td>
</tr>
<tr>
<td>Pilot Data (VHPT-Sheinbugs)</td>
<td>31</td>
</tr>
<tr>
<td>Further Iterations</td>
<td>33</td>
</tr>
<tr>
<td>VHPT Data for Further Iterations</td>
<td>34</td>
</tr>
<tr>
<td>Conclusions</td>
<td>38</td>
</tr>
<tr>
<td>III. Comparing Holistic Processing for Novel Objects and Faces</td>
<td>40</td>
</tr>
<tr>
<td>Discussion</td>
<td>47</td>
</tr>
<tr>
<td>IV. Pre-Screening, Individuation Training, and Post-Tests</td>
<td>49</td>
</tr>
<tr>
<td>Novel Object Matching Tests (NOMT)</td>
<td>49</td>
</tr>
<tr>
<td>Individuation Training Study</td>
<td>51</td>
</tr>
<tr>
<td>Post-training Tests</td>
<td>54</td>
</tr>
<tr>
<td>V.  Training Study Results</td>
<td>55</td>
</tr>
</tbody>
</table>
Novel Object Matching Test (Ziggerins and Greebles) ...............................55
Individuation Training Study ........................................................................57
Post-Training Tests of Holistic Processing ..................................................62
The Effects of Visual Ability and Holistic Processing and Performance..70
VI. Discussion ................................................................................................76
REFERENCES...............................................................................................82
### LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initial Data and Reliability for VHPT-Sheinbugs</td>
<td>32</td>
</tr>
<tr>
<td>2. Gender and Age Summary for VHPT Iterations</td>
<td>35</td>
</tr>
<tr>
<td>3. Mean Accuracy by Condition for VHPT Iterations</td>
<td>36</td>
</tr>
<tr>
<td>4. Reliability for VHPT for Aligned Condition</td>
<td>37</td>
</tr>
<tr>
<td>5. Reliability for VHPT after Regression</td>
<td>37</td>
</tr>
<tr>
<td>6. Correlation between performance on VHPT Conditions</td>
<td>38</td>
</tr>
<tr>
<td>7. Reliability of VHPT by Condition</td>
<td>42</td>
</tr>
<tr>
<td>8. Reliability and mean accuracy for VHPT categories</td>
<td>42</td>
</tr>
<tr>
<td>9. Correlations of holistic processing between categories</td>
<td>43</td>
</tr>
<tr>
<td>10. Correlation between each condition for AMT subjects</td>
<td>44</td>
</tr>
<tr>
<td>11. Correlations within and between conditions for AMT subjects</td>
<td>45</td>
</tr>
<tr>
<td>12. Correlation between overall accuracy in the VHPT tasks across category</td>
<td>46</td>
</tr>
<tr>
<td>13. Percentage of training for each category across the five training groups</td>
<td>52</td>
</tr>
<tr>
<td>14. Number of learned exemplars associated with each level of training</td>
<td>53</td>
</tr>
<tr>
<td>15. Number of test trials for each day of training</td>
<td>54</td>
</tr>
<tr>
<td>16. Reliability of VHPT for the trained sample</td>
<td>62</td>
</tr>
<tr>
<td>17. How visual ability (NOMT) and Experience Level predict holistic processing residuals for each category</td>
<td>71</td>
</tr>
<tr>
<td>18. How Visual Ability (NOMT) and Experience Level predict VHPT Accuracy for each category</td>
<td>71</td>
</tr>
</tbody>
</table>
19. Multiple regressions for VHPT Accuracy and Holistic Processing........72

LIST OF FIGURES

Figures 

1. Relationship between experience and holistic processing ......................3

2. Example Stimuli for the Greeble Training study..................................9

3. Examples of Greebles, Sheinbugs, and Ziggerins.................................28

4. Examples of aligned congruent, aligned incongruent, and RISE trials..........................................................31

5. Magnitude of the congruency effect for each category..........................43

6. Example of study and test trials for the NOMT. ................................50

7. Correlation between NOMT – Greeble and NOMT Ziggerin.................57

8. Accuracy and reaction time for the training across Day........................59

9. Improvement in accuracy and reaction time for each group and experience level.................................................................................................61

10. Magnitude of holistic processing for each category .........................63-64

11. Magnitude of holistic processing by experience level ..........................65

12. Correlation for holistic processing between categories in the trained sample. ........................................................................................................66-67

13. Correlation of holistic processing with the RISE baseline regressed out ... ........................................................................................................69-70

14. Correlations between experience level and NOMT scores with VHPT accuracy and the residuals of holistic processing for each novel object category....73-75
CHAPTER I

INTRODUCTION

To recognize faces, we must be able to process similar facial features (the eyes, nose, and mouth) that appear in the same general configuration. To deal with this difficult problem, people process faces holistically, or as an entire whole rather than by parts or singular features (Young, Hellawell, & Hay, 1987; Hole, 1994). Holistic processing is thought to play a part in identification/discrimination since it involves the integration of features and their relative spatial relations (Calder et al., 2000; Maurer, Le Grand, & Mondloch, 2002). Additionally, holistic processing may support quick feature processing that yields greater visual short-term memory capacity for faces than other objects (Curby, Glazek, & Gauthier, 2009). Thus, holistic processing is important for theories of face recognition and it is also associated with advantages such as more efficient encoding.

Initially, holistic processing was argued to be specific to faces (Farah et al., 1998). However, as people gain experience with non-face categories, holistic processing for that category increases (Busey & Vanderkolk, 2005; Bukach, Phillips, & Gauthier, 2010; Boggan, Bartlett, & Krawczyk, 2011). Holistic processing has been considered a hallmark of visual expertise since it comes as a result of training and experience (Gauthier et al., 1998; Gauthier & Tarr, 2002) and is generally absent for novices (Wong, Palmeri, & Gauthier, 2009), although there are certain situations that can yield holistic processing in novices such as contextual grouping (Curby et al., 2016) and in cases of Gestalt continuation for certain shapes (Zhao, Bülthoff, & Bülthoff, 2016). The
magnitude of holistic processing for lab-trained objects is positively correlated with increases in activity in the fusiform area (Gauthier & Tarr, 1998; Wong et al., 2009), which is thought to be sensitive to objects of expertise (Gauthier, Skudlarski, Gore, & Anderson, 2000; Bilalić, Langner, Ulrich, & Grodd, 2011; McGugin, Gatenby, Gore, & Gauthier, 2012). Thus, there is a link between experience and the presence of holistic processing strategies.

One assumption from this evidence is that holistic processing is associated with expertise, which is often defined as a high level of performance on recognition and other tasks. However, one recent finding showed no correlation between a measure of holistic face processing (Richler, Floyd, & Gauthier, 2014) and recognition performance as measured by the Cambridge Face Memory Test (CFMT, Duchaine & Nakayama, 2006). This study had a large sample size (n = 236) and used a measure of holistic face processing that was more reliable than the composite task, making it well suited for individual differences research. However, even with the statistical power to detect a relationship between holistic face processing and recognition performance, no correlation was found.

What is to be made of holistic processing if it does not relate to performance? This finding seems paradoxical, but there are several factors that could have an impact on performance, such as domain general visual ability and the amount of experience one has in a category. The effect of practice and experience generally follows a power law (Newell and Rosenbloom, 1981; Logan, 1992), reaching an asymptote after a certain amount of experience. Faces are a category in which most people have a great deal of experience. The general expectation is that performance is related to experience level, but
when experience is generally very high, differences in experience may not have a
significant impact on certain mechanisms like holistic processing. Put another way, even
if people have differences in experience for faces, holistic processing may depend on
experience only at lower levels (see Figure 1). It is possible that holistic processing is
influenced primarily by experience level, while variability in performance on face
recognition tasks is driven more so by visual ability. However, to date, the relationship
between holistic processing and experience level has not been well established.

![Figure 1. A schematic for the potential influence of experience on holistic processing and performance.](image)

Previous studies examining holistic processing have measured differences
between novices and experts in a category (Bukach, Phillips, & Gauthier, 2010; Wong et
al., 2012) or assessed group level improvements in holistic processing following training
(Gauthier et al., 1998; Wong, Palmeri, & Gauthier, 2009). One issue with predicting the
effects of experience on holistic processing is that it is difficult to quantify how much
experience a person has with a category. For example, studies of experts often use self-
reported levels of expertise in a category (Bukach, Phillips, & Gauthier, 2010). However, self-reports of expertise may not be an accurate predictor of performance on perceptual tasks (Barton, Hanif, & Ashraf, 2009; McGugin et al., 2012). Other studies of experts have compared holistic processing for coarse categories of expertise such as novices, recreational players, and experts in the case of chess (Boggan, Bartlett, & Krawczyk, 2011). The variability in experience within each of these classifications could be vast and difficult to quantify. Additionally, there may be intrinsic differences between those who become experts and novices. For example, experts may be a self-selected sample, consisting of people with inherent interest or aptitude in a certain domain. Factors such as motivation may also play a part in how experts perform compared to novices. Due to these differences and the difficulty of quantifying experience, it has been difficult to establish a firm relationship between experience level and holistic processing at the individual level from these studies.

Another approach to studying holistic processing has been to train novices in the lab and measure increases in holistic processing (Gauthier & Tarr, 1998; Wong, Palmeri, & Gauthier, 2009; Chua, Richler, & Gauthier, 2014, 2015; Richler et al., in preparation). This approach has some advantages over studying real world experts since it allows for the examination of changes as novices gain experience over the course of training. Typically, these objects have never been seen or experienced before, so experience level is controlled. Since novices do not process most objects holistically, any increases in holistic processing can be attributed to training and experience. Also, randomly assigning subjects into control/or training groups reduces possible effects of motivational differences between samples. However, one limitation of such designs is that subjects
have generally been given the same amount of training. Thus, subjects go through the
same training regimen and therefore do not vary in experience level, such that individual
differences in visual ability likely account for most of the variability in performance.
Only one study to our knowledge has parametrically manipulated experience and
measured holistic processing (Gauthier & Tarr, 2002). This study suggested that there
was no holistic processing for novel objects until subjects reached a training criterion (the
basic level shift; equal speed identifying objects at the individual and family levels).
However, this was done with five subjects, so there was not a large enough sample to
draw inferences about holistic processing at the individual level. Thus, the overall lack of
parametric manipulations of experience makes it difficult to determine whether individual
differences in holistic processing could be due to differences in experience level.

Another difficulty in establishing a relationship between experience and holistic
processing is one of measurement. The composite task, a test of holistic processing, is
sensitive to holistic effects for faces with a small number of subjects (for a meta-analysis,
see Richler & Gauthier, 2014). However, in its typical instantiation, it also has low
internal reliability (DeGutis et al., 2013; Ross, Richler, & Gauthier, 2014), which
constrains the magnitude of correlations using this measure. This issue makes it difficult
to draw inferences from the composite task at the individual level.

Experience level with a visual category could determine holistic processing, but
thus far, it has been difficult to pinpoint the contributions of different levels of
experience, especially at the individual level. There are also factors such as visual ability
that must be disentangled from experience level in terms of their influence on holistic
processing. The goal of this dissertation is to examine this issue by creating new reliable
measures of holistic processing for several categories of novel objects. These tests should allow for better measurements of holistic object processing and therefore will allow more sensitive tests of correlations of holistic processing across categories and of whether holistic processing is predicted by experience or ability. Next, we will parametrically manipulate experience with three categories of novel objects and compare the magnitude of holistic processing given different levels of experience. If experience level systematically affects the amount of holistic processing for a category and can predict holistic processing between different categories of objects, then we will have established, experimentally, that experience drives holistic processing. We can also compare category-specific experience and domain-general visual ability as predictors of holistic processing.

**Holistic Face and Object Processing**

There have been several measures and definitions proposed for holistic processing, such as better recognition for wholes rather than parts (Tanaka and Farah, 1993), sensitivity to configuration or inversion (Yin, 1969; Thompson, 1980) and featural vs. configural interference (Amishav and Kimchi, 2010). Here, we focus on the composite task, which measures failures to selectively attend to parts of a face (Young, Hellawell, & Hay, 1987). This measure is able to capture differences between faces and objects (Farah et al., 1998), is correlated with face selective markers in the brain (Gauthier & Tarr, 2002; Wong et al., 2009), and is sensitive to group level increases in holistic processing due to training (Gauthier & Tarr, 2002; Wong, Palmeri & Gauthier, 2009). Additionally, the task has good construct validity (Richler, Palmeri, & Gauthier,
2012), and improved versions of the composite task have been used to measure individual differences in holistic face processing (Richler, Floyd, & Gauthier, 2014), making it the measure best suited for this study.

The composite task is a sequential matching task wherein subjects are asked to attend to one part of an object and ignore the other part. If the information in the to-be-ignored half interferes with the ability to match the attended halves, then there was a failure to selectively attend to part of the object. Faces are processed more holistically than other objects (Farah et al., 1998; Gauthier & Tarr, 2002). Because of this finding, some have argued that holistic processing is reflective of a module specialized for the processing of faces (McKone et al., 2007). However, facelike holistic processing has also been observed in experts in nonface domains. For example, car experts process cars holistically but car novices do not (Gauthier et al., 2003). Experts in other visual domains also process objects in that domain holistically, such as chessboard configurations (Boggan, Bartlett, & Krawczyk, 2011), musical notation (Wong & Gauthier, 2010), fingerprints (Busey & Vanderkolk, 2005), and words (Wong & Bukach, 2011). These studies suggest that holistic processing can be a consequence of experience across many different visual domains.

One question is how more experience leads to increases in holistic processing. Previous training regimens have demonstrated that individuation, or learning the names for specific exemplars of objects, increases holistic processing whereas a categorization regimen matched on length and difficulty did not (Wong, Palmeri, & Gauthier, 2009). Thus, individuation is one path to holistic processing, but it is unclear what the mechanism behind this could be. According to the template hypothesis, faces are encoded
to fit whole face templates, such that individual parts of the face are not explicitly and independently represented (Tanaka & Farah, 1993; Farah et al., 1998). However, there is evidence suggesting that face parts can be represented independently. For instance, in the composite task, when faces were manipulated such that the test face (the second in the sequential presentation) was misaligned, holistic processing was still found (Richler, Tanaka, Brown, & Gauthier, 2008). Additionally, holistic processing was found even when the faces parts presented at study were adjacent to one another and thus were not presented in the context of whole faces. These findings indicate that non-holistic encoding of faces still resulted in holistic processing later in the trial, suggesting that holistic processing does not necessarily depend on an explicit representation of an entire face.

Other studies suggested that rather than being reflective of representational properties, holistic processing could be conceptualized as an attentional strategy (Richler, Wong, & Gauthier, 2011; Curby, Goldstein, & Blacker, 2013), wherein all parts of an object are obligatorily attended. Further studies examined the role of learned attention in the acquisition of facelike holistic processing strategies (Chua, Richler, & Gauthier, 2014; 2015). Subjects were trained on two novel object categories (Ploks and Glips). Novel objects were used since novices would have little to no holistic processing for these objects and there would be more room for increases in holistic processing due to training. Each subject learned names from both categories, but the most diagnostic or useful information for individuation was on complementary halves from each category (e.g. Glip top halves, Plok bottom halves, see Figure 2).
Figure 2. Examples of stimuli for the Glip Top/Plok Bottom group. For this group, the tops of Glips and the bottom of Ploks were diagnostic (varied the most) during individuation training (left). During the composite task, stimuli were created from diagnostic and nondiagnostic Greeble parts from each category. Composite task stimuli were either aligned, misaligned, or the task-irrelevant part was phase-scrambled.

For the composite task, Greebles that were made up of diagnostic and nondiagnostic parts were shown. A control group that had no experience with Greebles showed no difference in sensitivity for congruent or incongruent trials, showing no holistic processing at all. In contrast, for trained subjects who saw nondiagnostic parts, there was a main effect of congruency, suggesting some failures of selective attention even for parts that were not attended to during training. However, when presented with diagnostic Greeble parts, subjects showed a pattern indicative of facelike holistic processing. That is, when shown parts of Greebles that were previously important to
attend to, subjects could not help but to attend to those parts even when asked to ignore them. Critically, this study suggests that experience leads to holistic processing by establishing a history of learned attention to object parts most crucial for individuation. One possible source of variation in holistic processing is the strength of the representations for the trained objects or the strength of attentional biases acquired during training. The effect of practice or experience on performance tends to follow a power law (Newell & Rosenbloom, 1981; Logan, 1988, 1992). Therefore, more experience individuating objects may lead to stronger or more automatized representations and responses for objects in that category as people learn to attend to objects parts that are the most useful for a given task.

In the previously outlined studies, giving subjects a specific kind of experience with object parts led to facelike holistic processing. Following from this, one may expect that those with the more of this kind of experience will show a greater degree of holistic processing. However, this claim has not been directly tested. As mentioned previously, the internal reliability of the composite task is poor, which makes it difficult to firmly establish strong correlations between holistic processing and other measures. A proxy has been to relate performance with a category with holistic processing, because experience tends to increase performance. However, variability in performance can be driven by other factors (e.g., differences in ability, motivation) and so performance is not a direct measure of experience.

Because of measurement issues with the standard version of the composite task with faces, a modified version was created to assess holistic processing more reliably (Vanderbilt Holistic Processing Test – Faces, VHPT-F). The VHPT-F includes trials of
varying difficulty to capture a larger range of ability (Richler, Floyd, & Gauthier, 2014). Performance on the VHPT-F correlates well with the composite task to the extent allowed by their respective reliabilities, showing that they measured the same construct (Wang, Ross, & Gauthier, 2016). Surprisingly, despite being more reliable, this measure of holistic face processing does not correlate with performance on the Cambridge Face Memory Test (Duchaine & Nakayama, 2006), a commonly used measure of face recognition ability that is itself quite reliable (Richler, Floyd, & Gauthier, 2015). Thus, it seems that there is little relationship between holistic face processing and performance on face recognition tasks. Moreover, those who demonstrate the best face recognition ability can correctly recognize faces even with degraded or incomplete face information, suggesting that a holistic face representation may not be necessary for successful face recognition (Royer et al., 2015).

These findings can seem puzzling given holistic processing and its status as a hallmark of expertise. What is to be made of holistic processing and its relationship to expertise if it does not relate to recognition ability in the case of faces? It is possible that experience is oversaturated for faces. As people navigate the world, they gain years of experience individuating, categorizing and recognizing multitudes of faces, and thus, the amount of experience people have with faces is very high and may not vary much between people. Thus, variability on face performance may be more due to differences in visual ability as opposed to experience. As suggested previously, more experience with individuating objects in a category may lead to stronger or more automatized representations and responses for items in that category (Chua, Richler, & Gauthier, 2014; 2015). It is possible that as people gain more experience with a category, their
performance improves until it reaches a certain plateau. Individuals may have similar levels of experience with individuating faces. Because of this high level of experience (compared to most objects), the effect of experience may level off, reducing the variability of holistic face processing in the population. However, people may vary much more in their experience with a given category of objects compared to faces. Thus, differences in the extent of holistic processing for objects could stem from different levels of experience with those objects.

Several training studies have demonstrated that experience leads to a holistic processing strategy for trained objects. These training studies have subjects undergo the same training regimen (Gauthier et al., 1998; Wong, Palmeri, & Gauthier, 2009) or train equally between multiple object categories (Richler et al., submitted). One implicit inference from these studies is that a given amount of experience leads to a certain amount of holistic processing for the trained object category (indeed, on average the trained categories result in more holistic processing then untrained categories). However, the strength of the relationship between the amount of experience and holistic processing has not been tested because subjects are typically given similar amounts of experience in training studies.

In one training regimen, subjects were trained on four novel object categories, and there was an increase in holistic processing for these categories relative to an untrained category (Richler et al., submitted). However, there was no correlation across categories, despite a large sample size (n = 285) and with each holistic processing task showing some non-negligible reliability (~.4, Guttman’s λ2). Again, this finding shows that the benefits of experience to holistic processing are domain specific.
This leads to the question of whether there is a domain general aspect to holistic processing. That is, is there some common factor that could account for the magnitude of holistic processing for different categories of objects? If such a factor exists, we should be able to find correlations in holistic processing across different domains.

One open question is what factors result in differences in holistic processing for individuals. It is possible that this variation may derive from different sources, and we are interested in determining which of the sources can explain most of the variance in holistic processing within and between different visual domains. One possibility is that the amount of experience one has with a category could drive holistic processing levels. Although holistic processing has been long thought of as a hallmark of expertise, this has not been established empirically. To date, there have not been manipulations of experience to examine whether different amounts of experience results in different levels of holistic processing. Also, there are few reliable measures of holistic processing for non-face objects, making comparisons of holistic object processing between individuals difficult.

The first step in this dissertation will be to create more reliable tests of holistic processing for objects. Such improvements can be accomplished by mirroring some of the changes made to composite task by the Vanderbilt Holistic Processing Test for Faces (Richler, Floyd, & Gauthier, 2014). Better measurements of reliability would allow for improved sensitivity in measuring the effects of experience on holistic processing and increases in statistical power for within-subjects designs that use these tasks. These improved measures will allow for the examination of questions such as whether variability in holistic processing for objects correlates between different categories,
which would suggest some kind of domain general factor that drives holistic processing. Additionally, more reliable measures would allow for comparisons of holistic processing across categories across different levels of experience.

Next, we will implement a study design that parametrically manipulates the amount of training that subjects receive for three categories of visual objects, followed by assessments of holistic processing. Because experience level is the main factor of interest, measures of domain general visual ability will be used as an initial screening tool to create a restricted range of ability. It is still possible that those who have the most ability will have the most holistic processing after some training, in which case there should be a relationship between visual ability and holistic processing in each category regardless of experience level.

The main prediction is that the more experience subjects have with a category of objects, the more holistic processing they will show for that category. Additionally, this design allows for correlations of holistic processing between multiple categories of objects matched for level of experience. If the same amount of experience exerts a similar influence on holistic processing for both categories then it follows that levels of holistic processing for these two categories should be correlated in the present design. Finding that experience influences the magnitude of holistic processing should aid in making predictions about holistic processing and the advantages it is associated with, such as faster and more accurate recognition performance.
Previous Training Studies

The following describes previous studies of holistic processing and discuss potential mechanisms for how more experience yields more failures of selective attention. This will highlight the kind of training that has been done thus far, which will inform the design of the current study.

Single Category Training

Previous studies have trained subjects in the lab to determine the conditions under which a category of novel objects becomes processed holistically. In an initial study with novel objects called Greebles, subjects learned the names of individual Greebles over the course of a ten day training study (Gauthier et al., 1998). Following training, subjects performed a composite task and showed a reliable composite effect for Greebles, demonstrating for the first time a composite effect for a nonface category. Subjects went through Greeble individuation training and by the end of the study (7-9 hours), they reliably showed a facelike pattern of holistic processing for Greebles. Another study using an individuation training regimen (Gauthier & Tarr, 2002) compared holistic processing for Greebles and brain activation for Greebles. There was a positive correlation between the magnitude of holistic processing and activation in the right fusiform area for Greebles. This was one of the first studies to show that the magnitude of holistic effects could be predictive of experience-based changes in the brain. Additionally, these studies were the first to demonstrate that experience could lead to holistic processing strategies for non-face objects.
Comparison of two different training regimens

Wong, Palmeri, & Gauthier (2009) compared training regimens to determine the type of experience that would lead to holistic processing. Two groups of subjects trained on the same set of novel objects called Ziggerins and performed two different training regimens. One group did a basic level categorization task while the other did an individuation naming task as in previous training studies (Gauthier et al., 1998; Gauthier & Tarr, 2002). The two training tasks were roughly equated for difficulty and for length of training (ten 1 hour sessions). After training, the categorization and individuation groups performed a composite task to measure holistic processing. Only the group who had individuated Ziggerins showed facelike holistic processing. This study compared the effect of different training regimens (equated for time and difficulty) on holistic processing, finding that it takes a specific kind of training that leads to holistic processing behavior similar to that of faces. In a companion study, subjects who had gone through the categorization and individuation training regimens performed a one back task in an fMRI scanner to measure selectivity for Ziggerins before and after training. There was an increase in activity for Ziggerins in the right fusiform area that was positively correlated with the magnitude of holistic processing for Ziggerins, again showing that performance on the holistic processing task could be related to training effects in the brain at the individual level (Wong et al., 2009).

Comparison of training for multiple novel object categories

A recent training study examined whether there is a common underlying visual ability that could explain the variance in performance for different categories of novel
objects across multiple tasks (Richler et al., submitted). Subjects were trained on an exposure task that involved matching a target identity with another target in a space invaders game, a protocol used previously to increase perceptual expertise without the use of naming (Bukach, Vickery, Kinka, & Gauthier, 2012). Subjects performed this exposure task for four categories of novel objects, with a fifth novel category that was not trained.

In the composite task, subjects demonstrated a congruency advantage in reaction time for the four trained categories relative to the untrained category. Thus, the exposure training did have an effect on holistic processing. However, there was no correlation in the levels of holistic processing between the four trained categories, suggesting that the effects of the training were domain specific. There was still some variability in holistic processing between the categories. However, the amount of experience and exposure for each category was fairly well matched, so this study cannot demonstrate that the variability in holistic processing was due to differences in experience.

In this study, holistic processing for each category showed some reliability in terms of internal consistency. However, it is possible that reliability was relatively low because experience did not vary much. When the range of one predictor (in this case, experience) is restricted, the variance of the measurement is affected, which often results in underestimations of reliability (Fife, Mendoza, & Terry, 2012). Therefore, using a wider range of experience may be necessary to more accurately assess the reliability of holistic effects.
Comparison of novices to experts

Holistic processing has also been found for experts in other visual domains such as chessboard configurations (Boggan, Bartlett, & Krawczyk, 2012). For chess experts, congruency effects were equivalent for faces and chessboards. In contrast, those who played chess recreationally showed a small congruency effect for chessboards, and novices who did not show a congruency effect at all. This suggests a systematic increase of holistic processing as experience is acquired in a given domain. A similar pattern of results was found with English words (Wong et al., 2011). A group of native English readers and a group of Chinese readers who learned English as a second language (ESL) were tested on a composite task with English words. A reliable congruency effect was observed in both groups, but the congruency effect was larger for native English readers. The congruency effect for native speakers was also larger for common words than for nonsense words, suggesting that the congruency effect was modulated by how frequently a combination of letters was encountered. The ESL sample had roughly fifteen years of English learning experience, so even with considerable experience reading English, holistic processing for English words was smaller compared to those with a lifetime of experience reading English.

These studies suggest that the magnitude of holistic processing is related to the amount of experience people have with a visual category. However, as with faces, experience for words tends to be very high and may not vary much between individuals within the same group (native language vs. second language). Additionally, it is difficult to quantify how much experience a person has in a category of familiar objects or with reading words. Therefore, while these studies show differences at the group level related
to experience level, they may not translate into more fine-grained predictions at the individual level.

*Manipulating Experience for Novel Faces*

Studying the effects of experience on face performance has been difficult in the past, because people often have a great deal of experience with faces and differences in levels of face experience are difficult to quantify. One study examined the effects of different levels of experience with an artificial race of faces, Lunaris (McGugin, Ryan, Tamber-Rosenau, & Gauthier, 2017). Experience was manipulated such that people who perform poorly on a face recognition task (CFMT) would receive very little experience with the artificial race (Lunaris) and those who perform well with faces would receive much more training. Thus, in this design, experience and ability are confounded to create a large range in performance with these faces, similar to the variation in experience that can be found in the general population for familiar objects such as cars. Subjects performed a discrimination task for the Lunari before and after training, and then were run in an fMRI contrast with Caucasian and Lunari faces with non-face objects. There was more selectivity for Caucasian than Lunari faces, and neural activity for Caucasian and Lunari faces was highly correlated. Critically, neural activation in the right FFA2 correlated with performance with Lunari faces.

Holistic processing was not measured here, so it is still an open question how the acquisition of holistic processing interacts with experience and visual ability, but this study demonstrates that making selecting on ability in addition to manipulating experience can create a range of recognition performance for faces and that activation in
FFA can reflect experience driven changes in performance. These findings also raise questions about whether performance as measured by tasks such as the CFMT are driven not purely by ability but by some combination of ability and experience with faces. Recognition performance for faces and holistic processing do not correlate (Richler, Floyd, & Gauthier, 2014), so it could be that experience and ability differentially affect performance and holistic processing. This possibility could be a confound for studies using faces and familiar objects, so a more controlled manipulation of experience by using novel objects may be necessary.

**Experience and Learned Attention**

Holistic processing has been conceptualized as an automatized strategy that encourages attention to those parts most useful for individuation (Chua, Richler, & Gauthier, 2014; 2015). Given the importance of learned attentional biases on the acquisition of a holistic processing strategy, it is possible that the magnitude of holistic processing is related to the strength of representations for trained categories of objects. That is, more experience could result in higher levels of holistic processing. There is evidence that learning and familiarity with a stimulus increases selectivity of neurons in the primate inferior temporal cortex (IT), an area related to object recognition (Tanaka, 1996). Neurons in this area increased in activity after associative learning (Miyashita, 1988; Erickson and Desimone, 1999) and became more narrowly tuned to trained or behaviorally relevant stimuli (Kobatake, Wang, & Tanaka, 1998; Freedman et al., 2006; Woloszyn and Sheinberg, 2012). Critically, IT neurons can also be tuned to diagnostic features. For example, after categorization training, there is increased spiking activity for
populations of IT neurons in response to features that are diagnostic for categorization (Sigala & Logothetis, 2002), suggesting that neurons can be tuned to behaviorally relevant features. Furthermore, when monkeys were trained to discriminate baton stimuli with discrete top and bottom parts with distinct features, there was a population of neurons that responded more to the whole stimulus compared to the top and bottom parts combined (Baker, Behrmann, & Olson, 2002). The learned integration of stimulus wholes could be a mechanism associated with a holistic processing strategy.

With enough experience or training, objects can acquire salience, or automatized attention. In a study of acquired salience in cells in the anterior inferotemporal cortex (aIT), monkeys were trained to saccade to a target stimulus when it was accompanied by another stimulus and were rewarded for successful saccades to the targets (Jagadeesh, Chelazzi, Mishkin, & Desimone, 2001). The target stimuli gained behavioral relevance, and neuronal response to distractor stimuli decreased over the course of training. Additionally, neurons most selective for the target stimulus prior to training showed a large increase in response after the monkey learned that the target was behaviorally relevant. These neuronal responses may facilitate the suppression of task irrelevant information and increase selectivity for stimuli that acquire learned salience. When individuating objects, attention may be drawn to parts of objects that are diagnostic for correct identification. Since these parts become behaviorally relevant, a bias may develop to attend to these parts. If all parts of an object are diagnostic, this may lead to automatized attention to the entire object, making selective attention to an object part difficult (Chua, Richler, & Gauthier, 2014; 2015). After individuation learning, neurons in IT may become more responsive to diagnostic object features (Sigala & Logothetis,
and neurons in anterior inferotemporal cortex (aIT) may become more selective for behaviorally relevant parts of stimuli (Jagadeesh, Chelazzi Mishkin, & Desimone, 2001), and the integration of behaviorally relevant object features may lead to increased neuronal activation to whole objects (Baker, Behrmann, & Olson, 2002). Thus, one possible source for variation in holistic processing between individuals comes from differences in the strength of representation stemming from neuronal responses to behaviorally relevant objects or features and the binding of these features into integrated wholes.

A history of behavioral relevance can lead to biases to attend to certain features of objects. For example, while categorizing objects, attention is drawn to features most diagnostic for successful categorization (Kruschke, 1996). When the structure of a category is dependent on the presence of certain features, more attentional weight is placed on those features. In categorization tasks using eyetracking, first fixations and a high proportion of fixation time are allocated to the diagnostic features that best predict category membership (Rehder et al., 2005; Blair et al., 2009). Additionally, statistical learning can lead to biases to attend to regularities in the environment (Fiser & Aslin, 2002; Turke-Brown, Junge, & Scholl, 2005; Zhao, Al-Aidroos, & Turke-Brown, 2013). For example, probability cuing has often been used to induce biases in an area of space (Jiang et al., 2012; 2013). Probability cuing is employed in visual search tasks wherein one area of a visual scene is statistically more likely to contain a target (target “rich” areas), and subjects are faster to find targets in target rich areas. A similar probability cuing paradigm was used in the context of complex objects, specifically two different categories of Greebles (Chua & Gauthier, 2016). This paradigm involved visual search
wherein a target appeared in one half of a Greeble and a distractor appeared in the other half. There was a probability asymmetry such that the target appeared in one half (e.g. top) 89% of the time and the other half (e.g. bottom) 11% of the time. Thus, there was a half of the Greeble that was more likely to contain a target (the “target-rich half”) and a Greeble half where the target was not very likely to appear (the “target-sparse” half). This asymmetry led to faster search times when the target appeared in a target rich half as compared to the sparse half, which is evidence of an attentional bias to areas most likely to contain a target. Further, this bias was also category specific and transferred to new members of each category. Thus, attention can be biased to parts of objects as evidenced by faster target detection in high probability areas and this bias can be generalized across members of a visual category.

Selection history can also play a role in the formation of attentional sets or biases. For example, in item-specific control, features of a stimulus or the stimulus context act as cues that dictate attentional allocation based on past selection history (Jacoby et al., 2003; Bugg & Crump, 2012). Over the course of learning, there is a correspondence between a stimulus and a response, and attentional biases can form due to these stimulus-response mappings. In the context of the Plok/Glip training study outlined previously (Chua, Richler, & Gauthier, 2015), learning to individuate Greebles where one half of the object was diagnostic led to a strategy that prioritized attention to those diagnostic parts due to their behavioral relevance. In future instances when these diagnostic parts are encountered, they act as a cue that attention should be allocated to these parts. When an entire object is made up of diagnostic parts, all parts of the object become automatically
prioritized in attention, and it becomes difficult to selectively attend to a part of the
object.

Different amounts of experience with a category of objects could impact the
strength of object representations, specifically in terms of the strength of learned attention
to diagnostic parts. In the Instance Theory of Attention and Memory (Logan, 1988;
2002), attention to a stimulus will commit it to memory, and the strength of retrieval will
depend on the conditions of attention at encoding. Each encounter with a stimulus will
leave memory traces that accumulate with experience. With enough instances, there is a
transition from a novice-like algorithmic approach to a task to a more automatized
approach because the solutions to the task are being repeatedly retrieved from memory.
According to this theory, learned attentional biases can be conceptualized as the
accumulation of sufficient instances that lead to a certain object or stimulus feature being
selected or retrieved often enough to yield an automatized response, usually resulting in
faster responses. In a similar way, as experience with an object category increases, the
stimulus dimensions that are most relevant for the task become automatically prioritized
in attention (Kruschke, 1992; Logan, 2002). When one encounters an object that consists
of parts that have been previously attended, this leads to an automatized holistic
processing strategy. Therefore, different levels of experience could impact the magnitude
of holistic processing, with more experience leading to stronger associations between
objects and learned attentional settings involving automatized attention to entire objects.
Aims of the Current Study

The main purpose of this dissertation is to parametrically manipulate the amount of experience individuals receive with different categories of novel objects, and we intend to determine whether differences in experience will be associated with differences in the magnitude of holistic processing. Thus far, most studies of holistic processing have looked at group level training effects with comparisons between lab trained subjects and novices. Holistic processing tends to be domain specific (Bukach, Phillips, & Gauthier, 2010; Richler et al., submitted) and this could be because one’s amount of experience with these categories is also domain specific. If experience level drives holistic processing and the effects of experience are domain specific, there should be positive correlations in holistic processing for two categories that have received similar amounts of experience.

One consideration for this study is that visual ability could have an impact on the acquisition of holistic processing. Manipulating experience can result in a wide variation in recognition performance and neural selectivity in the FFA (McGugin et al, 2017). However, it is unclear whether variations in visual ability could best explain holistic processing. Learning tests using novel objects have been proposed to provide a useful index of a domain general visual ability because performance on novel categories can predict performance on tasks with both familiar and other novel objects (Richler, Wilmer, & Gauthier, 2017). Thus, it is possible that people with more visual ability acquire holistic processing strategies more quickly. Because the primary interest of this study is in the effects of experience, the current study design constrains visual ability into a middle range such that there are few people who are extremely high or low in visual
ability. Visual ability may have an impact on learning such that those high on visual ability may reach maximum performance too quickly. Conversely, even with sufficient experience, those with low visual ability may reach a plateau in holistic processing very quickly. It is still possible that visual ability could have an effect on holistic processing in this design but these constraints are meant to highlight differences based on experience.

The contribution of this work is twofold. First, the composite task has been useful for measuring differences between groups of subjects but its low reliability limits its ability to accurately assess individual differences in holistic processing (DeGutis et al., 2013; Ross et al., 2014). This research plan will include the creation of reliable measures of holistic processing for novel objects that can be used for this and future studies, mirroring that of previous efforts with faces (Richler, Floyd, & Gauthier, 2014). These tests can be used to make comparisons between holistic object processing and other factors such as visual ability.

Second, a missing component in previous studies is that they have mostly consisted of comparisons of groups of subjects who receive the same training regimen, so experience level does not vary much. This study design will manipulate experience parametrically to determine whether different levels of experience yield different levels of holistic processing. In conjunction with more reliable holistic processing tasks for novel objects, we should be able to better measure the contributions of experience on holistic processing strategies. We predict that the magnitude of holistic processing in a category of objects will correspond with level of experience in that category.
CHAPTER II

THE CREATION OF TESTS OF HOLISTIC PROCESSING FOR NOVEL OBJECTS

*Vanderbilt Holistic Processing Test – Novel Objects [VHPT-N]*

To properly address the question of whether holistic processing is related to experience levels, reliable tests need to be created for measuring holistic processing for novel objects. The Vanderbilt Holistic Processing Test for Novel Objects (VHPT-NO) was developed for this purpose. This test was structured after the fashion of the Vanderbilt Holistic Processing Test – Faces (VHPT-F), which is a more reliable individual differences measure of holistic processing compared to the standard composite task (Richler, Floyd, & Gauthier, 2014). There is one separate test for each five different novel object categories. Here, we will use three of these categories: Sheinbugs, Greebles, and Ziggerins (see Figure 3). The tests for each of these categories follow the same format.
Figure 3. Examples of Greebles, Sheinbugs, and Ziggerins. The objects could appear in one of two views during the study (a front view and a side view).

In the VHPT, subjects are instructed to attend to a part and holistic processing is defined by a congruency effect, but the trials also varied in the saliency of the ignored part to capture a wider range of holistic processing ability. Instead of a same/different matching task, the VHPT involves choosing between three alternatives, reducing chance performance from 50% to 33%. Additionally, there are trials that vary in the size of the attended part. The size of the attended segments affects the extent of holistic processing, with smaller attended segments resulting in larger congruency effects and with larger attended segments resulting in smaller congruency effects. This manipulation helps capture more variance in ability since those who process objects the most holistically should have positive congruency effects even when the attended segment is very large (and thus the task irrelevant segment is smaller). Conversely, those who exhibit the least
holistic processing should have relatively small congruency effects even when the attended segment is small.

**Procedure**

Forty-two exemplars from each category were used for each test. Each exemplar had an image that was the frontal view and another showing the same object from the side. Each object consisted of a target segment that was highlighted with a red box, and the rest of the object is a task irrelevant distractor. For each trial, a study composite object was shown for 2 seconds, followed by the test display. The test display consisted of three composite objects, one of which contained the same identity as the study object. Participants were instructed to choose, on each trial, which of the three test objects matched the target segment. The test and study objects could be shown in one of two views, a view from the front and a view of the object at a rotated angle. Study and target objects were seen from the same view within a single trial. Because subjects individuated exemplars from both views during training, holistic processing should generalize to both views.

The target segment varied in size (top half, bottom half, top 1/3, bottom 1/3, top 2/3, bottom 2/3, isolated part). The target segment was blocked, and there were 32 trials per block (7*32 = 224 trials total). In addition to standard aligned trials, there was a baseline consisting of phase scrambled images. In these trials, the distractor part was decomposed using an algorithm (Sadr & Sinha, 2004) that randomizes image components while retaining the low level attributes of the image such as luminance, contrast, and spatial frequency. This phase scrambled images have resulted in minimal interference in
the VHPT-F, and they are not confounded by differences in distance induced by spatial misalignment, making it a useful control condition (Richler, Floyd, & Gauthier, 2014; 2015). These RISE trials can be used as a control to ensure that holistic processing is specific to whole objects from the trained categories (we would not expect interference from these scrambled images, and to the extent that interference is present, we will regress it out).

In this test, holistic processing is indexed by failures of selective attention. Holistic processing is inferred when subjects cannot ignore information in the task-irrelevant part. On congruent trials, the target and irrelevant parts belonging to the correct answer were the same as the study item. For incongruent trials, the target part was paired with a distractor that belongs to a different object at test (for examples of congruent vs incongruent trials, see Figure 4). For these trials, the distractor part at study and test were phase scrambled images. Congruency here can still be defined by whether the phase scrambled distractor image at study vs test is the same or different. Better performance on congruent than on incongruent trials indicates an inability to selectively attend. That is, the task irrelevant object part influenced performance, despite instructions to ignore it.
Figure 4 – Examples of aligned congruent, aligned incongruent, and phase scrambled trials. The target part is highlighted in red. The target part varied in size (1/2, 1/3 or 2/3 shown here). The correct answer in these examples is marked with an asterisk.

Pilot Data (VHPT-Sheinbugs)

Twenty-five subjects participated in an initial version of the task with Sheinbug stimuli on Amazon Mechanical Turk. Reliability of the measurements was calculated using Guttman’s λ2 (Guttman, 1945) for each condition (see Table 1). Additionally, the reliability of the congruency effect for each alignment condition (aligned, phase-
scrambled) was calculated as a difference score (congruent – incongruent) with a formula that takes the variance between measures into account (Rogosa, Brandt, & Zimowski, 1982).

<table>
<thead>
<tr>
<th></th>
<th>Guttman $\lambda_2$</th>
<th>Subtraction Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aligned</td>
<td>Incongruent</td>
</tr>
<tr>
<td>Iteration 1</td>
<td>.64</td>
<td>.51</td>
</tr>
<tr>
<td>Iteration 2</td>
<td>.77</td>
<td>.84</td>
</tr>
</tbody>
</table>

Table 1. Reliability by condition (Guttman’s $\lambda_2$) and reliability of the congruency effect for each alignment condition for the VHPT-Sheinbugs.

The reliability of the congruency effect for aligned trials was quite low (0.01). Reliability for the congruency effect for the RISE baseline was higher (.214) but it is still low for the purposes of measuring individual differences. To improve reliability on this test, an item analysis was performed on each trial, correlating individual subject performance on each trial with the mean performance for all subjects within each condition. This correlation indexes how informative or predictive each trial is for overall performance in each condition. This analysis helped to identify trials that were uninformative or not difficult enough and these trials were altered (by changing the target part or adding more confusable foils) for the next iteration.

Forty-six subjects were run on the next version of this test. Mean accuracy for aligned congruent trials was 0.91 and for aligned incongruent trials, the mean was 0.88. The difference between congruent and incongruent trials was .03, yielding a small but significant congruency effect, $t(45) = 2.07, p < 0.05$, Cohen’s $d = 0.31$ (effect size formula = $t / n^2$), suggesting some failures of selective attention even for these novel
objects. For RISE trials, the mean for congruent trials was 0.89 and the mean for incongruent trials was 0.88, yielding a congruency effect of 0.01. This difference was not reliably different from zero, $t(45) = 1.04$, $p = 0.30$, Cohen’s $d = 0.16$, so there was no holistic processing for the scrambled baseline. This suggests that the RISE baseline worked as intended, yielding very few failures of selective attention. The congruency advantage for aligned and RISE trials was different, $t(45) = 3.73$, $p < 0.001$, Cohen’s $d_z = 0.57$, with greater congruency for aligned trials. Interestingly, even for novices, there was some holistic processing for novel objects. This finding is in contrast to previous studies, where subjects exhibit very little holistic processing for novel objects (Gauthier & Tarr, 2002; Wong, Palmeri, & Gauthier, 2009). It should be noted that the reliability of this congruency effect was still fairly low (.23), and the effect size of this congruency effect is relatively small. In a meta-analysis of 48 studies of holistic processing for faces, the mean effect size ($\eta^2_p$) found for the congruency effect was 0.32. Comparatively, the effect size obtained for the congruency effect for Sheinbugs in this sample ($\eta^2_p = 0.10$) was much smaller than that what is typically found for faces, so there is room for increases in holistic processing for novel objects after training.

After one iteration, reliability was markedly improved for all conditions and for the congruency effects found in both aligned and RISE trials. In the next section, we describe additional iterations designed to further increase reliability for this test.

**Further Iterations**

Similar tests to the VHPT-Sheinbugs were created for Greebles and Ziggerins. To further improve the internal reliability of the VHPTs, item analyses were carried out to
identify trials that were at ceiling performance or were not much information. Low information trials show little correlation between individual subject performance and mean performance over the sample within each condition, and these items do not have much predictive power for performance on the task. Additionally, we added manipulations to attempt to reduce accuracy on the task overall. It is important to avoid ceiling effects because holistic processing here is defined as a subtraction between two conditions, congruent minus incongruent, and so if performance is very high for incongruent trials then that for the congruent trials may hit ceiling. We predicted that holistic effects would increase with training, and this could be due to worse performance on incongruent trials, improved performance on congruent trials, or both.

Starting with the 3rd iteration of the tests, the encoding time was changed from 2 seconds to 1 second. Additionally, a pattern mask was inserted in between the study and test displays. This mask was made up of parts from the each respective category. These manipulations were intended to reduce accuracy to avoid ceiling performance. Other than those manipulations, the basic structure of the tests was the same as outlined above.

**VHPT Data for Further Iterations**

*Sample*

All subjects were recruited using Amazon Mechanical Turk. A separate sample was recruited for each iteration and each category (see Table 2). The subjects were paid $0.70 for participation in the study. Subjects who finished the studies in less than five minutes (indicating that they were skipping through trials) or were at chance performance
(33%) were removed. Twenty-two subjects total were removed from the analyses for this reason.

<table>
<thead>
<tr>
<th></th>
<th>Mean Age</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sheinbug</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iteration 1 (n = 25)</td>
<td>42.2 (12.7)</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Iteration 2 (n = 50)</td>
<td>42.2 (11.1)</td>
<td>21</td>
<td>29</td>
</tr>
<tr>
<td>Iteration 3 (n = 50)</td>
<td>38.2 (12.8)</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td><strong>Ziggerin</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iteration 1 (n = 25)</td>
<td>44.2 (12.5)</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Iteration 2 (n = 50)</td>
<td>40.7 (13.8)</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Iteration 3 (n = 50)</td>
<td>37.4 (13.2)</td>
<td>16</td>
<td>34</td>
</tr>
<tr>
<td><strong>Greeble</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iteration 1 (n = 25)</td>
<td>41.0 (13.0)</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>Iteration 2 (n = 50)</td>
<td>41.2 (13.1)</td>
<td>21</td>
<td>29</td>
</tr>
<tr>
<td>Iteration 3 (n = 50)</td>
<td>38.6 (12.8)</td>
<td>26</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 2. Mean age and gender information for each sample. The standard deviation of sample ages is in parentheses.

**Accuracy**

The mean accuracy for each iteration and condition are in Table 3. Congruency effects were measured by subtracting congruent and incongruent trials for both the aligned and RISE conditions. Congruency effects for most of the conditions were not significantly different from 0, with the exception being in the aligned trials for the first iteration of the Sheinbug VHPT. This suggests that there was no holistic processing for these objects for novices for aligned or RISE trials.
Table 3. Mean accuracy for each condition for the first three iterations in each category. For congruency, $* = p < .05$, significantly different than 0. A = Aligned, R = RISE, C = Congruent, I = Incongruent.

<table>
<thead>
<tr>
<th></th>
<th>Sheinbug</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aligned</td>
<td>Aligned</td>
<td>Aligned</td>
<td>RISE</td>
<td>RISE</td>
<td>RISE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Congruent</td>
<td>Incongruent</td>
<td>Congruency</td>
<td>Congruent</td>
<td>Incongruent</td>
<td>Congruency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.91</td>
<td>.88</td>
<td>.03*</td>
<td>.89</td>
<td>.88</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.88</td>
<td>.86</td>
<td>.02</td>
<td>.91</td>
<td>.89</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.79</td>
<td>.77</td>
<td>.02</td>
<td>.80</td>
<td>.79</td>
<td>.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Ziggerin</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aligned</td>
<td>Aligned</td>
<td>RISE</td>
<td>RISE</td>
<td>RISE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruent</td>
<td>Incongruent</td>
<td>Congruency</td>
<td>Congruent</td>
<td>Congruency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.86</td>
<td>.83</td>
<td>.03*</td>
<td>.85</td>
<td>.85</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.84</td>
<td>.82</td>
<td>.02</td>
<td>.89</td>
<td>.91</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.79</td>
<td>.76</td>
<td>.03</td>
<td>.80</td>
<td>.78</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Greeble</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aligned</td>
<td>Aligned</td>
<td>RISE</td>
<td>RISE</td>
<td>RISE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congruent</td>
<td>Incongruent</td>
<td>Congruency</td>
<td>Congruent</td>
<td>Congruency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>.98</td>
<td>.95</td>
<td>.03</td>
<td>.94</td>
<td>.96</td>
<td>-0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.96</td>
<td>.95</td>
<td>.01</td>
<td>.94</td>
<td>.92</td>
<td>.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>.81</td>
<td>.81</td>
<td>0</td>
<td>.78</td>
<td>.81</td>
<td>-0.03</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reliability

The reliability for each iteration is shown in Table 4. As in the initial set of data, reliability for each condition was calculated using Guttman’s $\lambda_2$, and the reliability of the Aligned and RISE conditions was calculated for the difference between congruent and incongruent trials. The third iteration included the manipulation of the mask and lowering the encoding time. With each iteration, the reliability for the aligned trials increased, with a final mean reliability around 0.38.
A second index of holistic processing was calculated by regressing congruency in the baseline condition (the RISE trials) from congruency in the condition of interest (the aligned congruency difference). In the past, holistic processing has been calculated by taking the difference in congruency for a main condition and subtracting the congruency difference in some baseline condition such as spatial misalignment or RISE trials. These difference scores can be affected both by variability in the condition of interest as well as variability from the baseline (Cronbach, 1970). Regressing out the variability from the baseline condition can improve the validity of the measure. These results are shown in Table 5. The mean reliability obtained after regressing out the RISE baseline was 0.36.

As a proxy for how the congruent and incongruent conditions differed in terms of performance, the correlation for these conditions was calculated (Table 6). The correlation for congruent and incongruent trials was quite high, suggesting that subjects were approaching these conditions similarly. It is unclear why these conditions are similar, but due to their high correlation, they seem to be tapping into the same ability.
These subjects were all novices and there was little holistic processing, so performance in these conditions was not well differentiated.

Sheinbug
Aligned
Sheinbug
RISE
Ziggerin
Aligned
Ziggerin
RISE
Greeble
Aligned
Greeble
RISE

<table>
<thead>
<tr>
<th></th>
<th>Iteration 1</th>
<th>Iteration 2</th>
<th>Iteration 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheinbug</td>
<td>0.66</td>
<td>0.67</td>
<td>0.76</td>
</tr>
<tr>
<td>Sheinbug</td>
<td>0.75</td>
<td>0.71</td>
<td>0.64</td>
</tr>
<tr>
<td>RISE</td>
<td>0.86</td>
<td>0.68</td>
<td>0.67</td>
</tr>
<tr>
<td>Ziggerin</td>
<td>0.82</td>
<td>0.68</td>
<td>0.79</td>
</tr>
<tr>
<td>RISE</td>
<td>0.73</td>
<td>0.73</td>
<td>0.79</td>
</tr>
<tr>
<td>Greeble</td>
<td>0.75</td>
<td>0.83</td>
<td>0.77</td>
</tr>
<tr>
<td>Aligned</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RISE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Correlations between performance on congruent and incongruent trials. These data are separated for the aligned and RISE conditions.

Conclusions

The tests found little evidence of holistic processing for novel objects, and there was no significant difference between the magnitude of holistic processing for the objects compared to the RISE baseline. This is expected, as none of the subjects had any experience with these categories, and there tends to be little holistic processing for novices in a category (Wong et al., 2008, Chua, Richler, & Gauthier, 2015). As experience with these objects increases, there should be more failures to selectively attend to part of the objects.

After each iteration, the VHPTs improved in reliability for both ways of estimating holistic processing (in the aligned condition alone, or after regressing out congruency effects in the RISE condition). The Guttman λ2 reliability for each condition also increased on average for each iteration, suggesting that the internal consistency of the trials was improving. Overall, the mean reliability for the third iteration of tests was .38 and .36 for subtraction and regression, respectively. This is an increase in reliability over the standard composite task for faces (~.1 - .24, DeGutis et al., 2013; Ross et al, 2014), although it is not as reliable as the VHPT-Faces (~.56, Richler, Floyd, & Gauthier,
It is worth noting that these reliability calculations are specific to each set of measurements. It is possible that reliability of holistic processing cannot reach high levels in a sample of participants who do not process objects holistically because they are novices. Indeed, we find evidence that congruent and incongruent trials are treated similarly (correlations between these conditions range from .64 - .86). The higher this correlation, the lower the reliability is for the difference score. The overall reliability within each condition is fairly good ($\lambda^2 = .66 - .81$ for the last iteration), so what is limiting the reliability of the measure here seems to be the fact that there is no holistic processing strategy. With training, we expect higher holistic processing and a lower correlation between congruent and incongruent performance, hopefully resulting in increases in reliability for the congruency effect.
CHAPTER III

COMPARING HOLISTIC PROCESSING FOR NOVEL OBJECTS AND FACES

One source of variation in holistic processing could be a domain general cognitive control mechanism that governs selective attention, which seems logical since holistic processing is often defined as a failure of selective attention (Young, Hay, & Hellawell, 1987). However, in the literature on cognitive control and selective attention, there is not much evidence that variability in a general cognitive control mechanism predicts performance on selective attention tasks. For instance, Stroop interference, which is similar to the composite task in that it measures selective attention with congruent and incongruent trials, does not consistently correlate between different versions of the task (r = .14-.52, Salthouse & Meinz, 1995; Yehene & Meiran, 2007). Additionally, the magnitude of Stroop interference can differ greatly based on the associative strengths between stimuli and task set (Ward et al., 2001), suggesting that this kind of interference may be related to one’s experience with a particular domain. In a similar way, interference in holistic processing is governed by experience in a visual domain, since holistic processing from expertise seems to be domain specific (Bukach et al., 2010). Thus, the evidence for a domain-general selective attention mechanism is mixed.

One large-scale training study in multiple categories demonstrated little to no correlation in holistic processing across different novel object categories (Richler et al., submitted). This study used the standard composite task, so it is possible that the test’s low reliability constrained the correlations in holistic processing between each category of novel objects. The VHPTs have better reliability than the composite task so they may
be better tools in detecting correlations in performance between categories. If holistic processing is governed primarily by a generalized selective attention mechanism, then holistic processing should be correlated across different categories of objects. To test this possibility, we collected data from 50 subjects in the VHPTs for Greebles, Sheinbugs, Ziggerins, and faces and correlations for holistic processing were calculated between each category.

Subjects

Fifty subjects were recruited online from Amazon Mechanical Turk. Seventeen more subjects were run but were not used because they either did not complete each test or had data for one or more tests that was at or below chance performance. The mean age of the sample was 40 years old (SD = 12.4), with 17 males and 33 females. Subjects were compensated $1.50 for completing one half of the tests (split between the Sheinbugs/Faces or Greebles/Ziggerins in randomized order), and these subjects were contacted to complete the other two VHPTs for another $1.50. To be eligible for participation, subjects were required to have U.S. IP addresses. Each subject was consented according to Vanderbilt University’s Institutional Review Board standards.

Results

First, reliability for each category was calculated (Table 7 and 8). The reliabilities for holistic processing in both the aligned condition and holistic processing after regressing out the RISE baseline were similar to that of iteration 3.
Table 7. Reliability of each condition (Guttman’s λ2).

<table>
<thead>
<tr>
<th></th>
<th>Aligned Congruent</th>
<th>Aligned Incongruent</th>
<th>RISE Congruent</th>
<th>RISE Incongruent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheinbug</td>
<td>0.77</td>
<td>0.78</td>
<td>0.76</td>
<td>0.77</td>
</tr>
<tr>
<td>Ziggerin</td>
<td>0.80</td>
<td>0.82</td>
<td>0.72</td>
<td>0.70</td>
</tr>
<tr>
<td>Greeble</td>
<td>0.75</td>
<td>0.77</td>
<td>0.66</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Table 8. Reliability and mean accuracy for VHPT categories.

<table>
<thead>
<tr>
<th></th>
<th>Aligned Reliability</th>
<th>Regression Reliability</th>
<th>Aligned Congruent Accuracy</th>
<th>Aligned Incongruent Accuracy</th>
<th>RISE Congruent Accuracy</th>
<th>RISE Incongruent Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheinbug</td>
<td>0.40</td>
<td>0.38</td>
<td>0.77</td>
<td>0.75</td>
<td>0.74</td>
<td>0.74</td>
</tr>
<tr>
<td>Ziggerin</td>
<td>0.45</td>
<td>0.39</td>
<td>0.77</td>
<td>0.76</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td>Greeble</td>
<td>0.39</td>
<td>0.32</td>
<td>0.79</td>
<td>0.80</td>
<td>0.80</td>
<td>0.81</td>
</tr>
<tr>
<td>Faces</td>
<td>0.71</td>
<td>0.61</td>
<td>0.74</td>
<td>0.54</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The amount of holistic processing for each test was calculated by subtracting the mean accuracy of incongruent trials from congruent trials. There was no significant holistic processing for Ziggerins (mean = .013, SD = .07, t(49) = 1.46, p = 0.15, Cohen’s d = 0.21) or Greebles (mean = -0.01, SD = .10, t(49) = -0.61, p = 0.54, Cohen’s d = 0.08).

There was marginally significant holistic processing for Sheinbugs (mean = .02, SD = .08, t(49) = 1.91, p = .062, Cohen’s d = 0.27). For faces, there was a significant congruency effect (mean = .199, SD = .07, t(49) = 15.5, p = 0.0001, Cohen’s d = 2.19) that was much larger than that of the novel objects. Despite comparable variability in each category, there was no evidence of holistic processing for the novel objects. These results are shown in Figure 5.
Figure 5. Magnitude of the congruency effect for each category. Error bars are 95% confidence intervals.

Of particular interest in this study was whether there could be some domain general selective attention mechanism that governs holistic processing. To that end, correlations were calculated for the congruency advantage between each category (see Table 9). None of the correlations between categories was significantly different than 0, providing no evidence for a relationship in holistic processing across category as measured by these tasks.

<table>
<thead>
<tr>
<th></th>
<th>Sheinbug</th>
<th>Ziggerin</th>
<th>Greeble</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ziggerin</td>
<td>-0.19 [-.10, .44]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greeble</td>
<td>-0.03 [-.25, .31]</td>
<td>0.08 [-.20, .35]</td>
<td></td>
</tr>
<tr>
<td>Faces</td>
<td>-0.10 [-.18, .37]</td>
<td>-0.09 [-.36, .19]</td>
<td>0.13 [-.15, .39]</td>
</tr>
</tbody>
</table>

Table 9. Correlations of holistic processing between categories. Cutoff for $p = 0.05, r = 0.27; p = 0.01, r = 0.38$. In parentheses are 95% confidence intervals.
In addition to the correlations of holistic processing between categories, zero order Pearson’s correlations were computed for the accuracy between conditions for all categories (Table 10).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shein AC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shein AI</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shein RC</td>
<td>0.66</td>
<td>0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shein RI</td>
<td>0.60</td>
<td>0.62</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zig AC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zig AI</td>
<td>0.23</td>
<td>0.20</td>
<td>0.45</td>
<td>0.31</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zig RC</td>
<td>0.48</td>
<td>0.38</td>
<td>0.43</td>
<td>0.40</td>
<td>0.49</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zig RI</td>
<td>0.34</td>
<td>0.33</td>
<td>0.40</td>
<td>0.46</td>
<td>0.68</td>
<td>0.52</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greeble AC</td>
<td>0.37</td>
<td>0.29</td>
<td>0.35</td>
<td>0.40</td>
<td>0.46</td>
<td>0.44</td>
<td>0.44</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greeble AI</td>
<td>0.36</td>
<td>0.26</td>
<td>0.34</td>
<td>0.32</td>
<td>0.31</td>
<td>0.35</td>
<td>0.40</td>
<td>0.33</td>
<td>0.60</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greeble RC</td>
<td>0.35</td>
<td>0.33</td>
<td>0.36</td>
<td>0.38</td>
<td>0.46</td>
<td>0.47</td>
<td>0.53</td>
<td>0.54</td>
<td>0.76</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greeble RI</td>
<td>0.26</td>
<td>0.31</td>
<td>0.27</td>
<td>0.33</td>
<td>0.38</td>
<td>0.32</td>
<td>0.37</td>
<td>0.36</td>
<td>0.72</td>
<td>0.77</td>
<td>0.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face C</td>
<td>0.07</td>
<td>0.16</td>
<td>-0.02</td>
<td>-0.01</td>
<td>0.01</td>
<td>0.04</td>
<td>0.15</td>
<td>0.20</td>
<td>0.02</td>
<td>0.02</td>
<td>0.06</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Face I</td>
<td>0.29</td>
<td>0.18</td>
<td>0.21</td>
<td>0.19</td>
<td>0.25</td>
<td>0.25</td>
<td>0.03</td>
<td>0.35</td>
<td>0.32</td>
<td>0.32</td>
<td>0.30</td>
<td>0.19</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 10. Pearson correlations between conditions. A = Aligned, R = RISE, C = Congruent, I = Incongruent. p < 0.05, r > 0.27; p < 0.01, r > 0.38.

Even though there is no correlation in holistic processing between categories, accuracy for the VHPT was generally correlated across conditions and categories. In novel objects, the correlations for congruent and incongruent trials are also generally similar, and this trend applies to both the aligned and RISE conditions. This is consistent with the lack of holistic processing for the novel objects, which suggests that the congruent and incongruent trials are not well differentiated.

Comparisons for the different categories and conditions were made to determine whether there could be some correspondence between performance on non-face objects in comparison to performance with faces. To make these comparisons, the correlation
coefficients for each condition were converted to Fisher Z values. Then, these correlations were averaged together and converted back to single correlation coefficients for each condition of interest. The correlation coefficients were then compared and tested for significance.

First, the average correlation within each category was calculated (Table 11). For the novel objects, performance in one condition was highly correlated with the other conditions within that category. For example, within the Sheinbug category, performance for aligned congruent was correlated for aligned incongruent, RISE congruent, and RISE incongruent. In contrast, there was no correlation for the face congruent and incongruent conditions, which is consistent with the fact that there was a strong congruency effect for faces.

<table>
<thead>
<tr>
<th>Sheinbug</th>
<th>Ziggerin</th>
<th>Greeble</th>
<th>Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.61 [.4, .76]</td>
<td>0.58 [.36, .74]</td>
<td>0.75 [.60, .85]</td>
<td>0.01 [-.27, .29]</td>
</tr>
</tbody>
</table>

Table 11. Correlations for VHPT performance within each category. In parentheses are 95% confidence intervals.

Next, correlations were calculated between the different novel object categories ($r = 0.46, 95\% \text{ CI} [0.21, 0.65]$), which were contrasted against the correlation between the face conditions ($r = 0.01, 95\% \text{ CI} [-0.27, 0.29]$). The difference between the correlations was significantly different, $z = 2.36, p = 0.02$, suggesting that performance within the non-face categories was higher than between the face conditions. Performance with faces was not predictive of performance in the other categories ($r = 0.14, 95\% \text{ CI} [-0.14, 0.40]$). For faces, performance could be affected by both ability and experience level, while performance with novel objects may dependent primarily on a domain general visual ability.
It is also possible that overall accuracy on the VHPTs is related for each category when all conditions are combined (Table 12). If such a relation were found, that would suggest some common factor that can explain performance in these matching tasks across different categories.

<table>
<thead>
<tr>
<th></th>
<th>Sheinbug Total</th>
<th>Ziggerin Total</th>
<th>Greeble Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheinbug Total</td>
<td>0.51* [.27, .69]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ziggerin Total</td>
<td>0.44* [.19, .64]</td>
<td>0.56* [.33, .73]</td>
<td></td>
</tr>
<tr>
<td>Greeble Total</td>
<td>0.20 [-.08, .45]</td>
<td>0.26 [-.02, .50]</td>
<td>0.17 [-.13, .43]</td>
</tr>
</tbody>
</table>

Table 12. Correlation between overall accuracy in the VHPT tasks across category. In parentheses are 95% confidence intervals. *, p < 0.01.

Using the methods outlined above, the correlations of performance were averaged and compared within the novel objects and between faces and novel objects (r = .50 within novel objects and r = 0.21 for objects and faces, p = 0.05). Thus, overall accuracy for the tasks was more similar within the non-face categories compared to performance between non-face and face categories. This may be indicative that the mechanisms that govern performance for faces could be influenced by both experience and ability, whereas performance on novel objects is more similar because they depend primarily on domain general visual ability.

Looking within face conditions, there is a consistent trend in the way that the face congruent and incongruent conditions correlate with the non-face conditions. Generally, the correlation for the face incongruent condition with each of the novel object conditions is higher than the face congruent condition and the novel object conditions, with positive correlations for face incongruent trials and correlations close to 0 for the face congruent trials. A non-parametric sign test was conducted comparing the face congruent and incongruent with the non-face conditions, and the correlations were larger for the face
incongruent correlations 10/11 times \((z = 2.89, p = 0.002)\). The average correlation for the face incongruent trials and the non-face conditions \((r = .25)\) was higher than the correlation for face congruent trials with the non-face conditions \((r = 0.04)\), but this difference is not significant, \(z = 1.04, p = 0.30\). Although the difference in correlation is small and not significant, because we have many conditions we were able to detect a significant pattern whereby performance with incongruent face trials, but not congruent face trials is predicted by object matching. This was not a planned comparison so we did not have a prediction for this relationship, but this effect, although small, seems to be systematic.

**Discussion**

In this study, the same subjects performed holistic processing tests for faces and three novel object categories. For novel objects, there was no holistic processing as measured by these tests, which is to be expected because subjects had no experience with these categories of objects. Although on average there was no holistic processing, there was still sufficient variation in holistic processing for each category that could be compared. If there is some domain general selective attention mechanism, then it should be reflected by positive correlations in holistic processing across each category. However, when comparing holistic processing across categories, there was no correlation between any of the categories, including faces. Thus, it seems that, at least in novices, variability in selective attention for different categories of objects is domain specific and does not seem to correspond well between categories. This applies in situations when there is no experience (the novel objects) or between categories with a gulf in experience.
(novel categories and faces). This makes a domain general selective attention mechanism unlikely, at least in terms of failures of selective attention that are associated with holistic object processing in the composite task. However, one finding from examining the correlation between conditions is that performance on incongruent trials can be predicted by some domain general mechanism, since it correlates more with non-face performance. That is, performance for incongruent face trials is similar to performance with objects. Therefore, if there are aspects of the VHPTs that are sensitive to increases in experience, they should be reflected in performance on congruent rather than incongruent trials.

In the following training study, we present the alternative possibility that holistic processing will be governed by variability in experience. We will introduce variations in experience to determine whether experience level can predict holistic processing across different categories of objects.
CHAPTER IV

PRE-SCREENING, INDIVIDUATION TRAINING, AND POST-TESTS

The training study consists of three parts. Subjects were initially screened on a measure of visual ability for novel objects, and subjects in the middle range of ability were selected to minimize floor and ceiling effects of ability on learning. This selection was applied to reduce the variation in visual ability between our training groups, since the primary interest of the study is examining the influence of experience. After selecting a sample with a middle range of ability, subjects were randomly assigned to groups that received different amounts of training with three categories of novel objects. The training consisted of individuation naming, which has resulted in holistic processing for the trained category in prior work (Gauthier et al, 1998; Gauthier & Tarr, 2002; Wong, Palmeri, & Gauthier, 2009; Chua, Richler, & Gauthier, 2014; 2015). Finally, following the individuation training, subjects took a post-training test of holistic processing (VHPT) for each of the trained categories.

Novel Object Matching Tests (NOMT)

To measure visual ability, subjects performed a prescreening matching test for novel objects, the Novel Object Matching Tests (NOMT, Richler, Wilmer & Gauthier, 2017). Subjects performed two versions of the NOMTs with Ziggerins and Greebles that were distinct from the ones used for the individuation training. For this task, subjects initially studied a set of six test objects. Each object was viewed once in three different
views. Subjects were shown a display consisting of one of the viewed objects with two distractor items and were asked to select the object they had previously viewed.

This process was repeated for six different objects, resulting in 18 trials during the introductory phase. Afterwards, all six learned objects in a category were shown together for twenty seconds. On each trial of the test phase, subjects saw one of the studied objects along with two distractors, and again the task was to select the previously studied object. After 18 trials, subjects were refreshed on the six memorized objects for another twenty seconds. Following this memorization stage, there were another 36 trials, for a total of 54 test trials (Figure 6). Combining the trials from the introductory and test phases, there were 72 trials total.

Figure 6. Example of study and test trials for the NOMT.
Because the test measures recognition ability for novel objects, they can be used as measures of visual ability that will not be conflated with previous experience. The NOMTs have good internal reliability ($\alpha = 0.8$-$0.89$). These tests also correlate across three visually different novel object categories ($r = .43-.53$), suggesting that there is a domain general component (Richler, Wilmer, & Gauthier, 2017). Due to this positive correlation, performance for the NOMT for one category can act as a baseline measure of visual object recognition ability. To reduce the influence of a domain general visual ability, subjects who performed within two standard deviations from the mean were selected to undergo the training.

**Individuation Training Study**

Individuation training has been shown to increase holistic processing in categories of novel objects (Wong et al., 2009; Chua, Richler, & Gauthier, 2014; 2015). In this training regimen, there are five groups of subjects that receive different levels of experience for three categories of novel objects. The total amount of training for each subject is the same (10 training sessions that last roughly one hour), regardless of training group. However, the groups vary in the amount of training they receive for each category.

The training consisted of individuating objects from three novel categories. Two of the novel categories were equated in terms of training time, with a third category receiving the remaining amount of time in an 10 hour allotment (Table 13). If experience level determines holistic processing, over the whole subject sample, holistic processing for the two matched categories should be positively correlated, and negatively correlated
with the remainder category (wherein experience is inversely related to the matched category).

<table>
<thead>
<tr>
<th>Group</th>
<th>Ziggerins (matched)</th>
<th>Greeble (matched)</th>
<th>Sheinbugs (remainder)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9%</td>
<td>9%</td>
<td>82%</td>
</tr>
<tr>
<td>2</td>
<td>14%</td>
<td>14%</td>
<td>72%</td>
</tr>
<tr>
<td>3</td>
<td>20%</td>
<td>20%</td>
<td>60%</td>
</tr>
<tr>
<td>4</td>
<td>30%</td>
<td>30%</td>
<td>40%</td>
</tr>
<tr>
<td>5</td>
<td>45%</td>
<td>45%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table 13. Percentage of training for each category across the five training groups.

These levels of experience were chosen to encompass a wide range of training time to generate variability stemming from differences in experience. In addition to the time trained, each level differed in the number of exemplars and foils encountered, with higher amounts of time associated with more exemplars (Table 14). Thus, the differences in experience are not only in how much time subjects are trained in a category but with the diversity of exemplars learned and encountered. Experience level as defined here is reflective of a confounded combination of training time and diversity of the learning set. The number of exemplars was varied between training level to approximate differences in the variety of exemplars that real life experts encounter vs. the number that novices would encounter. Experts in a category generally see a wide variety of exemplars compared to novices and this manipulation was included to reflect this difference.
Table 14. Number of learned exemplars and foils associated with each level of training (expressed by proportion of trials in a category).

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Named Exemplars</th>
<th>Foils</th>
<th>Total Objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>9, 10</td>
<td>4</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>20</td>
<td>8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>30</td>
<td>10</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>40, 45</td>
<td>12</td>
<td>16</td>
<td>28</td>
</tr>
<tr>
<td>60</td>
<td>14</td>
<td>18</td>
<td>32</td>
</tr>
<tr>
<td>72</td>
<td>18</td>
<td>22</td>
<td>40</td>
</tr>
<tr>
<td>82</td>
<td>20</td>
<td>26</td>
<td>46</td>
</tr>
</tbody>
</table>

Following training, subjects performed a holistic processing test for each of the trained categories of novel objects using a more reliable measure (the Vanderbilt Holistic Processing Test for Novel Objects, VHPT-NO). Correlations between the levels of holistic processing were compared between the trained categories across different levels of training. This design is advantageous since between the three trained categories, the total amount of training time will be equated, reducing possible external factors such as differences in motivation levels between subjects. For example, it is possible that if the amount of training was uneven for different groups, the least motivated subjects may terminate the study early in a disproportionate fashion. Additionally, this design allows for the examination of correlations within the same subjects for different categories that are matched in training time. Similar comparisons can be made between holistic processing in the matched categories with the third trained category.

*Individuation Procedure*

For the individuation training, there was a learning and test phase. During the learning phase, subjects saw novel objects with their name labels. For test trials, an object
was shown without the name label and the task was to type in the first letter of the object’s name. Each object could be shown in one of two views, so subjects learned to encode the objects’ identity from multiple views. Following the training session was a testing phase with feedback where the previously seen objects appeared without labels. The task was the same as before, but participants were given feedback if they typed in an incorrect answer. There were also unnamed foils included in the stimulus set, and subjects typed in “n” for “no name” to designate that the object was not a learned exemplar. Each day of training consisted of 700 test trials total, split between the three categories (Table 15). To minimize possible effects of attrition due to the disproportionate difficulty of the training between categories, subjects learned half of the exemplars (as well as half of the foils) for the first five days of training. Starting on day 6, the entire set was introduced.

<table>
<thead>
<tr>
<th>Group</th>
<th>Greeble</th>
<th>Ziggerin</th>
<th>Sheinbugs</th>
<th>Total Trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>64</td>
<td>64</td>
<td>572</td>
<td>700</td>
</tr>
<tr>
<td>2</td>
<td>96</td>
<td>96</td>
<td>508</td>
<td>700</td>
</tr>
<tr>
<td>3</td>
<td>140</td>
<td>140</td>
<td>420</td>
<td>700</td>
</tr>
<tr>
<td>4</td>
<td>208</td>
<td>208</td>
<td>284</td>
<td>700</td>
</tr>
<tr>
<td>5</td>
<td>312</td>
<td>312</td>
<td>76</td>
<td>700</td>
</tr>
</tbody>
</table>

Table 15. Number of test trials for each day of training.

**Post-Training Tests**

On the final day, subjects performed the VHPT for Greebles, Ziggerins, and Sheinbugs. The procedure of the tests is the same as outlined in Chapter 2. Each test took roughly 25 minutes, and the order of the tests was randomized for each subject.
CHAPTER V

TRAINING STUDY RESULTS

Sample

Out of 77 total subjects tested in the NOMT screening tests, 50 were qualified and completed the individuation study and the post-tests (mean age = 22.4 years, 17 male, 33 female). Seven were ineligible for the training because they scored two standard deviations above the group means. Two subjects were eligible for the individuation training study but terminated the study early (both did three days of training). Another 9 subjects completed the training but the VHPT post-test data was lost due to software problems, so a new set of 9 subjects was recruited to replace them.

For the individuation training, the subjects were randomly assigned to five different groups, with 10 subjects in each group. Subjects were paid $15 an hour. The entirety of the study consisted of the pre-training screening (~30 minutes), 10 individuation training sessions (~1 hour each), and the post training tests (~1 hour).

Novel Object Matching Test (Ziggerins and Greebles)

For each test, raw scores were calculated for each subject (out of a total of 72). Based on a previous sample of subjects tested on the same tasks (Richler, Wilmer, & Gauthier, 2017), upper and lower boundary cutoffs were set that were two standard deviations above the mean (the 5th and 95th percentile). In that sample (n = 672), the mean score for the NOMT-Greeble was 50.6 (70.3%), and the mean score for the NOMT-Ziggerins was 60.8 (84.3%). The standard deviation for the sample for these two tests
was 7.3 and 8.1, respectively. The final cutoffs for inclusion in the present study were combined scores above 81 and below 137 (out of 144).

For the sample that went on to do the individuation training (n = 50), the mean score for the NOMT-Greeble was 52.6 (SD = 8.0, 73.1% correct, coefficient of variation = 0.15), and the mean score for the NOMT-Ziggerins was 63.0 (SD = 8.10, 87.5% correct, coefficient of variation = 0.13), with a mean combined score from the two tests of 115.6. As in a previous study (Richler, Wilmer, & Gauthier, 2017), there was a positive correlation between performance on the NOMT-Ziggerin and NOMT-Greeble, \( r = 0.32, r^2 = 0.10, 95\% \text{ CI } [0.05, 0.55], t(48) = 2.36, p = 0.02 \) (see Figure 7). With the total sample of 77 subjects, mean score for the NOMT-Greeble was 53.26 (SD = 8.7, 74.0% correct, coefficient of variation = 0.16) and the mean score for NOMT-Ziggerins was 63.20 (SD = 8.5, 87.8% correct, coefficient of variation = .15). The correlation between the two NOMT categories was higher with the full sample, \( r = 0.51, r^2 = 0.26, 95\% \text{ CI } [0.32, 0.66], t(75) = 5.13, p < 0.0001. \) Without constraints on ability, this correlation was therefore similar to previously reported samples using these tests \( (r = 0.50, r^2 = 0.25; \) Richler, Wilmer, & Gauthier, 2017). The positive relationship in performance between the categories suggests a common factor that influenced recognition performance for each test, making the NOMT tests a good index of visual ability. Subjects were randomly assigned to groups and there was no main effect of group on the combined test scores, \( F(4,45) = 0.87, p = 0.49, \eta_p^2 = 0.000. \) This suggests that the training groups could be considered to have comparable visual ability.
Individuation Training Study

Reaction Time and Accuracy by Day

For the entire sample, the accuracy and reaction time was plotted by Day (see Figure 8). Trials that were faster than 100 milliseconds and slower than 20 seconds were removed from the analysis (0.35% of trials removed). It is important to note that the first and second halves of training are not comparable in terms of difficulty. In the first half of the study, half of the total exemplars were learned, and the entire set was introduced starting in the second half.

There was a linear trend in improvement by day, with faster reaction times and increased accuracy with more training. In the second half of the study, when more learned exemplars were introduced, accuracy and reaction times returned to similar levels to the initial days of training (of course subjects by then performed at the same level...
while the chance level went down with more objects). However, there was an improvement in accuracy and reaction time in the days following the introduction of the new exemplars. When each half of the training is examined, there is a trend of improvement with each additional day for both accuracy and reaction time. By the end of training, when the full set of exemplars have been introduced, accuracy and reaction time were much improved compared to the beginning of the training session.
Figure 8. Accuracy and Reaction Time for all subjects across training day. Error bars are 95% confidence intervals.
Improvement in Accuracy and Reaction Time by Experience Level

It is possible that the amount of exemplars learned and the amount of experience that one receives with a category influenced the degree of improvement on the individuation training. Each group of subjects received a different amount of training for each category. There was no significant category-by-day interaction for either reaction time, $F(9,423) = 1.53, p = 0.13$, $\eta_p^2 = 0.03$, or accuracy, $F(9,423) = 0.43, p = 0.91$, $\eta_p^2 = 0.00$. To examine the improvement over the course of the study, the data was binned by quarter, with each bin containing 25% of the total training data. Each bin encompassed 2.5 days of training. Improvement for the first half of the training (Days 1-5) was calculated by subtracting the mean performance in the second quarter from the first quarter. Improvement for the second half of the training (Days 6-10) was calculated by subtracting the mean performance in the fourth quarter from the third quarter. The first and second half improvement were averaged for an index of overall improvement for both reaction time and accuracy. This data is broken down by group and experience level in Figure 9. There is a general trend of improvement with increasing experience level. A linear trend on experience level significantly predicted improvement in accuracy, $b = 21.0, t(8) = 3.97, p = 0.004$, $R^2 = 0.66$. A similar linear trend was found with reaction time, $b = 0.0003, t(8) = 4.02, p = 0.004$, $R^2 = 0.67$. Thus, experience was positively associated with improvement on the training task. This trend may be attributed to the fact that the higher experience levels were associated with more exemplars to learn and more foil objects to distinguish from. Therefore, initial performance may be lower for groups that are learning more objects, with more room to improve in these cases. In contrast, the
groups that are receiving less experience also learned fewer exemplars, so they may have been near ceiling at the beginning of the training.

**Reaction Time Improvement**

![Reaction Time Improvement Chart](chart)

**Accuracy Improvement**

![Accuracy Improvement Chart](chart)

Figure 9. Improvement in accuracy and reaction time for each group and experience level. Error bars are 95% confidence intervals.
Post-Training Tests of Holistic Processing

Reliability of the Trained Sample

As in the previous sample, we calculated the reliability for both the holistic processing defined as the congruency effect on aligned trials, and regressing out from this index the congruency effect for RISE trials (Table 16). The mean reliabilities for the two methods did not differ much (.47 for subtraction, .44 for regression), so reliability was not largely affected by regressing variance from the RISE baseline. The reliability of the trained sample was slightly higher than that of the subjects recruited by Amazon Mechanical Turk in Chapter 3 (.41 for subtraction, .36 for regression in that sample).

<table>
<thead>
<tr>
<th>Subtraction</th>
<th>Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Greeble</td>
<td>Sheinbug</td>
</tr>
<tr>
<td>0.48</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 16. Reliability for the trained sample for aligned congruent and the aligned congruent with the RISE condition regressed out.

Holistic processing by training group

Holistic processing is defined as the difference in performance for congruent and incongruent trials. Holistic processing for the aligned condition was calculated for each category and arranged by experience level (see Figure 10). The first comparison of interest was whether similar levels of experience would yield similar holistic processing between groups. To that end, the magnitude of holistic processing was compared for the matched categories (Ziggerins and Greebles). Category (Sheinbug, Greeble, Ziggerin) and Experience Level were entered into a between subjects ANOVA. There was no main effect of Category, $F(1,45) = 1.18, p = 0.87, \eta^2 = 0.00$, and no Category x Experience level interaction, $F(4,45) = 0.34, p = 0.85, \eta^2 = 0.01$. Since there was no significant
difference in the magnitude of holistic processing between the matched categories across different experience levels, they were combined for the rest of these analyses to increase statistical power.

![Ziggerin Aligned Congruency](image1)

![Greeble Aligned Congruency](image2)
Figure 10. The magnitude of holistic processing for each category. The holistic processing for the matched categories (Greeble and Ziggerins) was combined.

The next comparison was made to determine whether experience level could predict the magnitude of holistic processing. A linear contrast was used for holistic processing at the five different levels of experience for each category. For Sheinbugs, a linear trend on experience level significantly predicted holistic processing, $b = 0.18, t(48) = 3.02, p = 0.004, R^2 = .16$. For the matched categories, the weights in the linear contrast were reversed since experience in the matched and Sheinbug categories are inversely
related. For the matched categories, there was a linear trend on experience level that significantly predicted holistic processing, $b = 0.01$, $t(98) = 2.89$, $p = 0.005$, $R^2 = 0.08$. Thus, for both the Sheinbugs and the two matched categories, experience level was able to predict the magnitude of holistic processing, with higher experience levels yielding more holistic processing. The effect of experience level on holistic processing was roughly double for the Sheinbug category compared to the matched ones, and this could be due to the larger range in experience for the Sheinbugs (9 – 82%) compared to the matched categories (10 – 45%). Holistic processing across each experience level is shown in Figure 11.

![Holistic Processing by Experience Level](image)

**Figure 11.** Holistic processing by Experience Level. Error bars are 95% confidence intervals.

*Regressing out baseline condition*

In the VHPT-Novel Objects, a condition with phase-scrambled images of the stimuli was used as a baseline. Similar to the congruency effect described above, a linear contrast was made for aligned congruency, regressing out the congruency for the RISE
baseline, at the each of the five different levels of experience for each category. For Sheinbugs, a linear trend on experience level significantly predicted holistic processing, \( b = 0.18, t(48) = 9.02, p = 0.004 \), with an \( R^2 \) of .16. For the matched categories, there was a linear trend on experience level that significantly predicted holistic processing, \( b = 0.011, t(98) = 8.67, p = 0.004 \), with an \( R^2 \) of .08. As with the aligned trials, the residuals for each category could be predicted with a linear trend, with more experience yielding higher values.

**Correlation between Holistic Processing for Trained Categories**

The correlations for the congruency advantage between each category were calculated between each category (Figure 12).
Figure 12. Correlation for holistic processing between categories in the trained sample.
The Greeble and Ziggerin categories had equal experience levels across groups, while the Sheinbugs had a level experience that was inversely related to the matched categories. Comparisons were made for holistic processing in each category. First, the categories unmatched on experience were compared. There was no significant correlation for the Sheinbugs and Ziggerins, \( r = -0.01 \), 95% CIs [-0.29, 0.27], \( t(48) = -0.09 \), \( p = 0.93 \), or between Sheinbugs and Greebles, \( r = -0.19 \), 95% CIs [-0.44, 0.09], \( t(48) = -1.31 \), \( p = 0.20 \). Thus, there was no relationship in holistic processing for the unmatched categories.

In contrast, there was a positive correlation for Ziggerins and Greebles, which received an equal amount of experience, \( r = 0.39 \), 95% CIs [0.13, 0.60], \( t(48) = 2.96 \), \( p = 0.005 \). Thus, the amount of experience one has with a category seems to predict holistic processing across different object categories.

Next, holistic processing between the groups was assessed after regressing variance from the RISE baseline (Figure 13). There was no significant correlation for the Sheinbugs and Ziggerin residuals, \( r = -0.04 \), 95% CIs [-0.24, 0.31], \( t(48) = -0.26 \), \( p = 0.80 \), or for Sheinbugs and Greebles, \( r = -0.24 \), 95% CIs [-0.49, 0.04], \( t(48) = -1.68 \), \( p = 0.10 \). As with the aligned condition, there was a positive correlation for the Ziggerins and Greebles, which were matched on experience, \( r = 0.35 \), 95% CIs [0.08, 0.57], \( t(48) = 2.56 \), \( p = 0.01 \). Thus, the amount of experience one has with a category seems to predict holistic processing across different object categories.
Figure 13. Correlation of holistic processing with the RISE baseline regressed out in the trained sample.

The Effects of Visual Ability on Holistic Processing and Performance

This experimental design was set up to explore the effects of experience on holistic processing. To that end, the range of visual ability as measured by the Novel Object Matching Tests was constrained such that there were few extreme performers on the high or low end. Additionally, each group did not differ in terms of overall visual ability. Still, it is possible that the development of holistic processing can be influenced by visual ability at the individual level. To that end, combined NOMT scores were compared to holistic processing for the entire sample. If the acquisition of holistic processing is influenced by visual ability, one might expect more holistic processing for those with more ability regardless of how much experience they have received. That is, within each level of experience, those with the most visual ability may see more gains in holistic processing compared to those with less ability. The correlations for the residuals
of holistic processing with NOMT scores and experience level (defined as the percentage of training for each group) are shown in Table 17. These are partial correlations, in each case controlling for the other variable (NOMT or Experience Level). As shown previously, experience level could predict holistic processing for each category. However, there was no significant relationship between visual ability and holistic processing in any of the trained categories. In this study, experience level accounted for holistic processing more than visual ability.

<table>
<thead>
<tr>
<th></th>
<th>Sheinbug Residuals</th>
<th>Ziggerin Residuals</th>
<th>Greeble Residuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMT Score</td>
<td>0.13 [-.15, .39]</td>
<td>0.04 [-.24, .31]</td>
<td>0.09 [-.19, .36]</td>
</tr>
<tr>
<td>Experience Level</td>
<td>0.39* [.13, .60]</td>
<td>0.29* [0.01, .53]</td>
<td>0.30* [0.02, .53]</td>
</tr>
</tbody>
</table>

Table 17. How visual ability (NOMT) and Experience Level predict holistic processing residuals for each category - Pearson r. *p < 0.05.

Next, to determine whether visual ability has an impact on overall performance on the matching tests (averaging congruent and incongruent conditions rather than subtracting them), partial correlations were calculated for NOMT scores and experience level and overall VHPT accuracy for each category (Table 18).

<table>
<thead>
<tr>
<th></th>
<th>Sheinbug Accuracy</th>
<th>Ziggerin Accuracy</th>
<th>Greeble Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMT Score</td>
<td>0.58** [.36, .74]</td>
<td>0.55** [.32, .72]</td>
<td>0.67** [.48, .80]</td>
</tr>
<tr>
<td>Experience Level</td>
<td>0.12 [-.16, .39]</td>
<td>-0.09 [-.36, .19]</td>
<td>0.05 [-.23, .32]</td>
</tr>
</tbody>
</table>

Table 18. How visual ability (NOMT) and Experience Level predict VHPT Accuracy for each category - Pearson r. *p < 0.05, **p < 0.002.

NOMT scores were able to predict performance on the VHPT tasks. Experience level did not predict accuracy on the VHPTs for any category. Next, we wanted to compare the effect of experience and ability on VHPT with experience level and holistic processing. To that end, NOMT and Experience Level were entered into multiple regressions for overall accuracy and residuals of holistic processing for each category (Table 19).
<table>
<thead>
<tr>
<th></th>
<th>Sheinbug VHPT Accuracy</th>
<th>Sheinbug HP (Residuals)</th>
<th>Ziggerin VHPT Accuracy</th>
<th>Ziggerin HP (Residuals)</th>
<th>Greeble VHPT Accuracy</th>
<th>Greeble HP (Residuals)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R-squared (adjusted)</td>
<td></td>
<td>R-squared (adjusted)</td>
<td></td>
<td>R-squared (adjusted)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>= 31.7%</td>
<td></td>
<td>= 13.0%</td>
<td></td>
<td>= 28.3%</td>
<td>= 4.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(b)</td>
<td>(se)</td>
<td>(t)</td>
<td>(p)</td>
<td>(b)</td>
<td>(se)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.26</td>
<td>0.12</td>
<td>2.22</td>
<td>0.032</td>
<td>-0.190</td>
<td>0.142</td>
</tr>
<tr>
<td>NOMT Score</td>
<td><strong>0.69</strong></td>
<td><strong>0.14</strong></td>
<td><strong>4.81</strong></td>
<td><strong>0.000</strong></td>
<td>0.150</td>
<td>0.176</td>
</tr>
<tr>
<td>Experience Level</td>
<td>0.00</td>
<td>0.00</td>
<td>0.84</td>
<td>0.410</td>
<td><strong>0.000</strong></td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 19. Multiple regressions for overall accuracy in the VHPT and holistic processing with NOMT score and experience level for each category.

For all categories, the score on the NOMT predicted accuracy for the VHPT tasks but they did not predict the magnitude of holistic processing. In contrast, experience level predicted the magnitude of holistic processing but did not predict VHPT accuracy. Thus, there was a dissociation in the relative influences of visual ability and experience level on performance and holistic processing. The partial correlations for NOMT score and experience level with VHPT accuracy and holistic processing for each category are depicted in Figure 14.
Figure 14. Correlations between experience level and NOMT scores with VHPT accuracy and the residuals of holistic processing for each novel object category.
CHAPTER VI

DISCUSSION

The main purpose of this research was to determine the factors that best predict holistic processing. First, we examined the possibility of a domain general selective attention mechanism. If such a mechanism exists, then there should be some correspondence in holistic processing across different categories, even for novices. To that end, subjects were run in the Vanderbilt Holistic Processing Tests for faces and three categories of novel objects. There was no correspondence in the levels of holistic processing between any of these categories, so there was not much evidence for a domain general selective attention mechanism that could explain holistic processing for different categories.

The next possibility was that experience level had an influence on holistic processing. To test this relationship, experience was manipulated in a training regimen and holistic processing was measured at different levels of experience. A linear trend of experience predicted the magnitude of holistic processing, and holistic processing for two different categories matched in experience was positively correlated. Thus, within the same subjects, experience level could predict the magnitude of holistic processing across different categories. This trend was predictive at the group level with different levels of experience and at the individual level across the entire sample of trained subjects. Previous studies had often compared trained and untrained sample of subjects, but the amount of experience for the trained subjects was generally the same. These results show
for the first time a systematic effect of experience level on holistic processing, with more experience resulting in more failures of selective attention.

Importantly, there was a distinction made here between performance on the matching tests and holistic processing. Holistic processing has been discussed as a hallmark of expertise, and expertise had been defined as good performance on tasks like the CFMT. In this study, holistic processing was primarily predicted by experience level, but overall performance on the test was more influenced by visual ability than experience level. This finding might help explain the lack of correlation for holistic processing in faces (Richler, Floyd, & Gauthier, 2014). For familiar objects, there is a range of experience that can influence the magnitude of holistic processing. Many people have a great deal of experience with faces and there might not be as large a range of experience with faces compared to other objects. However, performance with faces and objects might be influenced by a combination of visual ability and experience level, resulting in a larger range of performance. Therefore, tasks such as the CFMT and other measures of performance might be indexing this combination of visual ability and experience.

Although holistic processing is considered a hallmark of expertise, it might more accurately be considered an indirect proxy of experience level. For example, two people who might have different levels of visual abilities but receive the same amount of training in a category might perform differently on matching tests, but they might have similar levels of holistic processing.

In this study, visual ability as measured by the Novel Object Matching Tests exerted little influence on holistic processing. The range of visual ability was restricted here, as we eliminated people who were high or low in visual ability, and having these
restrictions in range can constrain the magnitude of correlations (Fife, Mendoza, & Terry, 2012). There is a possible relationship between visual ability and holistic processing, but such a relationship may require more individuals at the extremes that were excluded in this study. A training regimen that controls experience level and has an appropriate range of visual ability would be better poised to determine the degree to which visual ability can affect holistic processing.

In terms of mechanism, individuation training results in increased attention to those parts of objects that are most important for identification. As people gain more experience with objects, stronger attentional biases develop such that the entire object becomes obligatorily attended. In category-specific cognitive control, there is a learned correspondence between features of objects and certain responses (Jacoby et al., 2013), and here we show that the automatized attentional strategies are strengthened with more experience. In this study, more experience led to more holistic processing, but it is unclear here whether adding more experience would lead to diminishing returns. Combining the two groups that received the most experience (Sheinbugs at 82% and 73%), the Cohen’s d of the congruency effect is 1.21. The sample of subjects who took the VHPT-Faces had a congruency effect with a Cohen’s d of 2.19. There is still a large difference in the effect sizes for faces compared to the groups of subjects in this sample who received the most training. Based on the trend, one might expect more experience to an additional effect on holistic processing, but the comparison here is being made between a few hours of training in a lab setting and with faces, a category that people have a lifetime of experience individuating. Experience can lead to increases in holistic
processing for novel objects, but it is unclear how much experience is necessary for the magnitude of holistic processing to be comparable to that of faces.

One open question is if experience increases automatized responses to objects, then is individuation the only route to holistic processing? This and previous training studies (Gauthier & Tarr, 1998, Wong et al., 2009) demonstrate that individuation is sufficient for holistic processing, but it is possible that other kinds of learned attention can also lead to holistic processing, such as probability cuing, reward, and statistical learning. There have been few parametric manipulations of this kind of attentional training, but one study found that combining factors such as reward and probability cuing can have an additive effect on the strength of attentional biases (Stankevich and Geng, 2014), suggesting that learned attention can have contributions from multiple sources. One possibility for future studies could examine whether other forms of attentional training could yield holistic processing without the need to individuate.

The main claim from this study has some caveats. First, the overall reliability of the VHPT was not very high (mean reliability = ~.46). Although it is higher than standard versions of the composite task, tests with reliabilities this low are not ideal for individual differences research. This lack of reliability constrains the correlations found within this study. For example, even though there was little relationship found for the categories that were inversely related in experience level, it is possible that such a relationship could be detected with better measures. There are some modifications that can be included to improve the reliability, such as the continued iteration, identification and replacement of low information trials, including more trials overall, and creating and using more unique exemplars to reduce repetition. Repetition of objects in these tests could be a measure of
learning rather than holistic processing (Endress & Potter, 2014; Richler, Floyd, & Gauthier, 2015), and the use of repeated stimuli may result in the adaptation of strategies towards those stimuli that are not reflective of selective attention. With these types of tweaks, it is possible that reliability can be further improved and these tests can be stronger measures of holistic processing at the individual level.

A second caveat is that the definition of experience used here is a combination of training time and quantity of exemplars learned. This manipulation was set up to reflect differences in experts and novices, since experts will have seen a wider variety of exemplars compared to novices in addition to having more exposure and experience with those objects. Given this setup, it cannot be determined whether experience level or the number of exemplars one learns exerts more of an influence on holistic processing, but this study shows that a combination of both of these factors does result in differences in holistic processing as more experience (as defined here) is gained. Future studies can isolate time and diversity of exemplars to determine which factor exerts more of an influence on increases in holistic processing.

Another direction for future study would involve examining whether holistic processing has neural correlates. One previous study using Lunari faces purposefully confounded visual ability and experience level, with those with the most ability receiving the most training (McGugin et al., 2017). This resulted in a range of performance similar to that of familiar objects, and there was a correlation between FFA2 and performance on a Lunari recognition task. Holistic processing was not measured in this task, but a similar design could be used to determine whether these differences in experience and visual ability yield different levels of holistic processing. If that were the case, then neural
selectivity for Lunaris (or trained objects) could be compared to indexes of both performance and holistic processing.
REFERENCES


