Journal of Expertise
2020. Vol. 3(1)

ISSN 2573-2773

The Vanderbilt Holistic Processing Tests for Novel Objects: Validation in Novice Participants

Kao-Wei Chua[1] and Isabel Gauthier[2]
[1]Department of Psychology, New York University
[2]Department of Psychology, Vanderbilt University

Correspondence: Kao-Wei Chua, kaochua@gmail.com

Abstract

In recent work, the Vanderbilt Holistic Processing Tests for novel objects (VHPT-NOs), were used to show that holistic processing for artificial objects increased as a function of parametric variation of experience. Here, novices are tested on the VHPT-Nos to address two questions. First, does the test detect any level of holistic processing for novel objects in novices? Second, how is part matching performance on this test related to object recognition ability, as measured by the Novel Object Memory Test (NOMT)? In a high-powered study, we provide substantial evidence of no holistic processing on the VHPT-NO in novices, including for arguably facelike symmetrical Greebles. Evidence of no correlations between measures of holistic processing suggests that these indices can be considered free of influences from domain-general selective attention. In contrast, overall performance in part matching in the VHPT-NO shows shared variance across categories, which we postulate is related to object recognition. A second study provides direct evidence that part matching measures to a large extent the same ability as whole object learning on the NOMT. Our results suggest that any holistic processing measured in the VHPT-NOs will not be contaminated by domain-general effects and can be considered entirely due to experience with a category. The VHPT-NO will therefore be useful in further examination of how different aspects of experience contribute to the development of holistic processing. Materials for the VHPT-NO are available at https://figshare.com/articles/VHPT_NO_zip/8345597.

Keywords

Object recognition, holistic processing, vision, individual differences

Introduction

Holistic processing has received many definitions, but one that has been central to the study of face recognition and other kinds of perceptual expertise is rooted in difficulty selectively attending to part of an object (Richler, Palmeri & Gauthier, 2012). Failures of selective attention have been used in most studies concerned with individual differences of holistic processing (e.g., Konar, Bennett, & Sekuler, 2010; Richler, Cheung, & Gauthier, 2011; DeGutis, Wilmer, & Cohan, 2013). Holistic processing can be measured with the composite task (Young, Hellawell, & Hay, 1987) wherein subjects are cued to attend to part of a face or object and told to ignore the rest of the image. Subjects then see two sequentially displayed composite objects and are asked to judge whether the attended part is the same in the first and second composites. The goal of the current work is to provide additional evidence for the validity of the recently developed Vanderbilt Holistic Processing Tests for Novel Objects (VHPT-NO), which are designed to measure both object recognition ability and holistic processing with artificial objects.

The composite task has been used to demonstrate that holistic processing is larger for faces than most categories of non-face objects (Farah, Wilson, Drain, & Tanaka, 1998; Richler et al., 2011) and to show it is obtained for experts in domains such as cars (Bukach, Phillips, & Gauthier, 2010), chessboard configurations (Boggan, Bartlett, & Krawczyk 2012), and novel objects after individuation training (Gauthier, Williams, Tarr, & Tanaka, 1998; Wong, Palmeri, & Gauthier, 2009). Although holistic processing can be considered a hallmark of visual expertise, including the vast experience we have for faces, it is difficult to study the role of experience with faces because even those with the least experience individuating faces still have a great deal of it compared to their individuation experience for other categories. Therefore, one approach to understand how experience influences how we process objects, including holistic processing, is to use novel objects so that we can fully control experience.

Until recently, there were no tests designed to measure holistic processing with novel objects. In a recent study (Chua & Gauthier, 2019), the VHPT-NO were used so that holistic processing for artificial non-face objects could be measured as a function of parametric variation of experience. In this study, subjects first received between 1 and a little over 8 hours of training learning to individuate objects from three different artificial object categories (all subjects received experience with all three categories, but subjects varied in terms of which categories they received the most experience with). Testing with the VHPT-NO revealed that holistic processing for artificial objects was related to the amount of experience subjects had received with other exemplars from the same artificial category. A caveat is that many aspects of experience were varied together (number of objects seen, number of names learned, duration of training) such that we know very little about what aspect of experience is critical. The availability of standard tests such as the VHPT-NOs can facilitate the study of this question, as the effect of different training manipulations can be evaluated on a common outcome across different studies. One limitation at the moment is that no data are available on the VHPT-NOs for subjects with no experience at all with these artificial objects. The present work aims to provide further validation of the VHPT-NO in groups of fully novice individuals who had never previously encountered these objects.

The first question we ask is whether any significant holistic processing can be detected by the VHPT-NO in novices. Chua & Gauthier (2019) found evidence of holistic processing as a function of experience in the VHPT-NO, but they did not test complete novices. Two other recent studies used a more standard version of the composite task, one not designed for individual differences measurement, in a training paradigm with novel objects. Richler et al. (2019) trained 246 novices for one hour and 45 min on each of four categories of novel objects. Two of the three categories we use in the present work were among the trained categories (symmetrical Greebles and Sheinbugs, see Figure 1). Subjects were then tested on a few post tests, including a composite test, with new examples of the trained categories as well as objects from a completely untrained category (the third category we use here, vertical Ziggerins). The study found significant holistic processing for objects from the trained categories, but not for the untrained Ziggerins. Chua, Richler, & Gauthier (2015) trained 80 subjects with asymmetrical Greebles for 3 hours and also measured holistic processing with the standard composite task in these subjects, as well as a group of 40 novices. Holistic processing was obtained for new examples of the trained category, but not in novice subjects. These three training studies make a convincing case for training-induced holistic processing with novel non-face objects, an effect that the most recent work shows systematically grows with experience. However, it is less clear whether any holistic processing of these novel objects can be obtained in complete novices, in the more sensitive VHPT-NO and given sufficient power.

This is important for future use and interpretation of the VHPT-NOs, because there has been a concern that measuring holistic processing as a congruency effect in a selective attention task reflects domain-general cognitive control processes like those that lead to Stroop and Flanker effects (Rossion, 2013). Recent work using the Vanderbilt Holistic Processing Tests for faces (VHPT-F), the test that inspired the VHPT-NO, showed that there was no common variance between congruency effects on a set of different Stroop and Flanker tasks and that index of holistic processing for faces (Gauthier, Chua, & Richler, 2018). It is conceivable, however, that this would be different for non-face objects, especially given demonstrations of holistic effects with novel line patterns in novices (Zhao, Bulthoff, & Bulthoff, 2016; Curby, Huang & Moerel, 2019). A related question is whether non-face objects that are more facelike in their appearance, such as the symmetrical Greebles, are more likely to be processed holistically by novices. Indeed, one fMRI study found a neural inversion effect for Greebles before any training and suggested that people may process Greebles as faces, even before training (Brants, Wagemans & Op de Beeck, 2011). In summary, in Study 1, we determine whether the VHPT-NOs can detect any holistic processing in novices, and specifically whether we can find evidence of more holistic processing for symmetrical Greebles compared to the non-facelike Sheinbugs and Ziggerins.

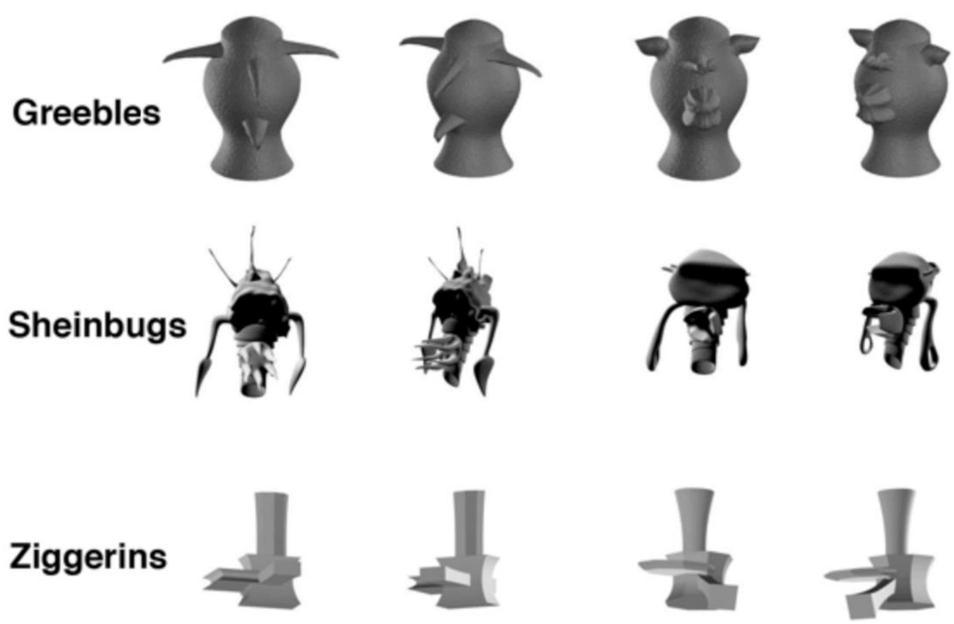

Figure 1. Examples of Greebles, Sheinbugs, and Ziggerins. The objects could appear in one of two views during the study (a front view and a side view).

If we found some evidence of holistic processing in novices, it would suggest the possibility that performance on the tests could be contaminated by domain-general influences, for instance individual differences in cognitive control (Rossion, 2013). In the same/difference composite task for faces, congruency effects are robust under a wide variety of task constraints, suggesting the task does measure robust holistic processing. However, some task settings, such as the combination of long exposure durations with feedback (Meinhardt-Injac, Persike, & Meinhardt, 2011) or not indicating at study which part will be relevant on a given trial can influence the magnitude of the congruency effects, especially in children and older adults who may have less efficient attentional control (Meinhardt-Injac, Boutet, Persike, Meinhardt, & Imhof, 2017). Such domain-general effects should by definition be common across categories and therefore lead to correlations of holistic effects across different object categories. In designing the VHPT-NOs, we chose task constraints that should limit such domain general contributions, such as providing no feedback and pre-cueing the relevant part on each trial. Previewing our results in Study 1, we were unlikely to find evidence of such correlations because we did not find holistic processing in novices in the VHPT-NOs.

A second question is whether the average performance on the VHPT-NO can be used as an index of general object recognition ability. The average performance on the VHPT-NO reflects how well subjects perform on matching parts of objects that the test instructs subjects to attend, regardless of the congruency of the part they are told not to attend. In Chua & Gauthier (2019), this measure was found to correlate well with performance on the Novel Object Memory Tests (NOMTs, Richler, Wilmer, & Gauthier, 2017), suggesting that part matching offers a valid measure of the ability measured in the NOMTs. However, Chua & Gauthier (2019) observed this correlation in a group of individuals tested on the VHPT-NO after receiving training (in fact, a variable amount of training across subjects) on the novel categories. In Study 1, we test the prediction that the average performance on part matching for different categories should be correlated. In Study 2, we determine whether part matching scores in the VHPT-NO relate to performance on the NOMT, providing further evidence that a single test VHPT-NO can produce useful measurement of both holistic processing and object recognition ability.

Study 1

Subjects

In Chua & Gauthier (2019), the most holistic processing was obtained for Sheinbugs after 8 hours of training (dz = 2.71). Given this very large effect size, the sample size required to detect such an effect with 95% power at an alpha of .01 is = 5. In the same study, holistic processing was significant after 2 hours of training (with Greebles and Ziggerins) with an average effect size of dz = 1.13. The sample size required to detect such an effect with 95% power at an alpha of .05 is $N = 13$. In that study, the group closest to novices were subjects trained with Greebles and Ziggerins for only 1 hour. This group showed no significant holistic processing, with and effect size of dz = .66. The sample size required to detect such an effect with 95% power at an alpha of .05 is $N = 32$. Acknowledging that this effect size was estimated on the basis of a subset of subjects (10 of the 50 trained subjects), and that a small sample can overestimate effect sizes, we calculate that with 50 subjects, we have more than 95% power to detect an effect that is 80% of that effect size (dz =.53).

Fifty subjects were recruited online from Amazon Mechanical Turk (AMT). The use of online testing was deemed appropriate because the Vanderbilt Holistic Processing Test for faces (VHPT-F, Richler, Floyd, & Gauthier, 2014), from which the VHPT-NO is inspired, was validated in several experiments on AMT. Seventeen more subjects were run but their data were not used because they either did not complete the entire test battery or had data for one or more tests 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), and these subjects were contacted to complete the other two VHPT-NOs 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.

Materials and Procedure

Forty-two objects from each category were used for each test. For each object we used a frontal view and a 3/4 view. Target segments were highlighted with a red box. 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 target part, either in a congruent or an incongruent context. Subjects were instructed to choose 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.

The target segment varied in size (top half, bottom half, top 1/3, bottom 1/3, top 2/3, bottom 2/3, isolated part, see Figure 2). The target segment was blocked[1], and there were 32 trials per block (7*32 = 224 trials total). In addition to these aligned trials, we also included baseline trials where the distractor parts were phase-scrambled images (see Richler, et al., 2014). Phase scrambling was performed with the RISE algorithm (Sadr & Sinha, 2004), which randomizes image components while retaining the low-level attributes of the image such as luminance, contrast, and spatial frequency. This RISE baseline performed similarly to a misaligned baseline with faces in the VHPT-F (Richler, et al., 2014; 2015), and in a training study with novel objects, showing no congruency effect (Chua, Richler, & Gauthier, 2015). One advantage of the RISE baseline over the misaligned baseline is that overall performance on RISE trials in difficulty with performance on the congruent-aligned trials.

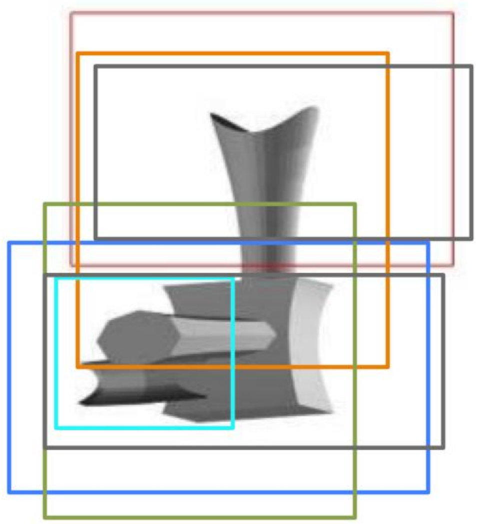

Figure 2. Illustration of the various sizes of to-be-attended region (horizontal shifts introduced to improve visibility).

In this test, holistic processing is inferred when subjects cannot ignore information presented in the task-irrelevant part. On congruent trials, the target and irrelevant parts belonging to the correct answer are the same as in the study item. For incongruent trials, the target part is paired with a distractor part from a different object (for examples of congruent and incongruent trials, see Figure 3). For RISE trials, the distractor part at study and test were phase-scrambled images and congruency is still be defined by the identity of these phase scrambled distractor parts. Holistic processing is measured by subtracting performance from the incongruent condition from performance on the congruent condition (the same is done for the RISE trials, and RISE congruency difference scores can in turn be regressed out of the congruency effect for aligned trials).

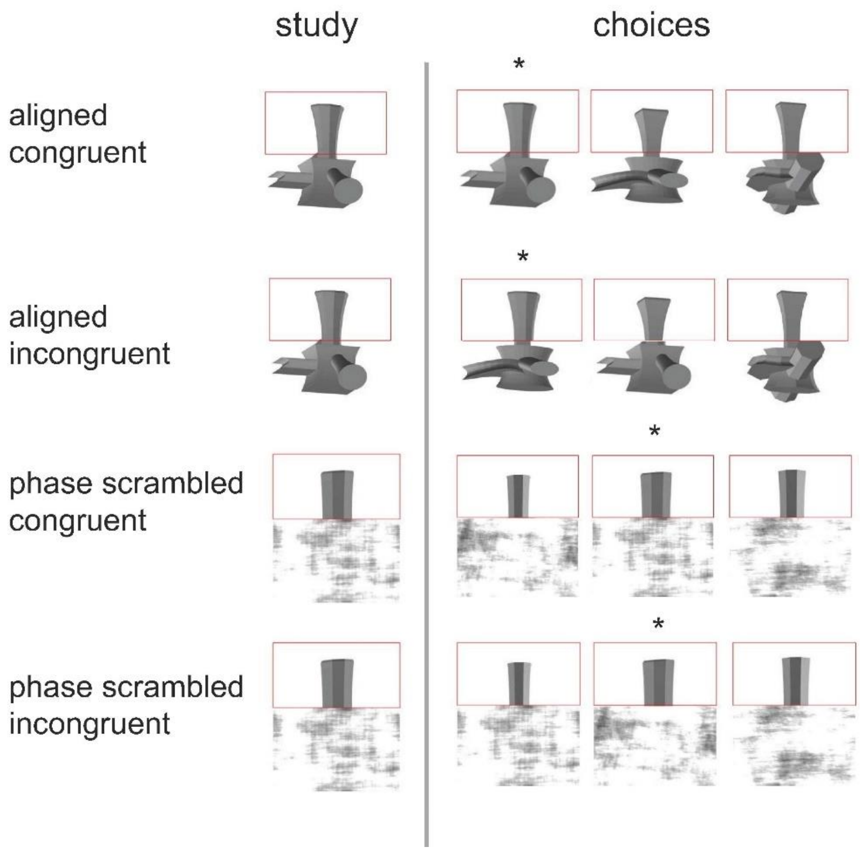

Figure 3. Examples of aligned congruent, aligned incongruent, and phase scrambled trials. The target part is highlighted in red. The target part varied in size, as shown in Figure 2. The correct answer in these examples is marked with an asterisk.

Results

As in previous research (Richler et al., 2012; Ross et al., 2015, DeGutis et al., 2013), we calculated the reliability in each condition using Guttman lambda2, the reliability of the holistic processing difference scores on normal and RISE trials (Rogosa, Brandt, & Zimoowski, 1982), and the reliability of the congruency effect in the normal condition, regressing out the congruency scores from the RISE condition (Malgady & Colon-Malgady, 1991). Reliability and mean accuracy for each condition and category are reported in Table 1. The congruent and incongruent conditions for novel objects were highly correlated (ranging from .59 to .69) whereas for faces, this correlation was .01 (this is the main reason that the index of holistic processing is more reliable for faces than for objects).

Table 1. Reliability and mean accuracy of each condition and for congruency indices for VHPT-NO categories and for faces in the VHPT-F. We do not provide an average for the VHPT-F as we do not use this score in correlational analyses, due to the very low correlation between congruent and incongruent trials in this task.

	Sheinbugs		Greebles		Ziggerins		Faces	
	Reliability	Mean (*SD*)	Reliability	Mean (*SD*)	Reliability	Mean (*SD*)	Reliability	Mean (*SD*)
Aligned C	0.80	.73(.08)	0.75	.78(.10)	0.75	.77(.07)	0.64	.73(.08)
Aligned I	0.79	.70(.08)	0.77	.79(.10)	0.75	.76(.07)	0.61	.55(.06)
RISE C	0.76	.74(.10)	0.77	.80(.11)	0.72	.78(.07)	-	-
RISE I	0.77	.74(.09)	0.74	.81(.10)	0.70	.77(.08)	-	-
Aligned C-I	0.42	.02(.07)	0.38	.01(.09)	0.37	.01(.06)	0.63	.18(.11)
RISE C-I	0.47	.00(.09)	0.00	.01(.06)	0.27	.01(.07)	-	-
Aligned C-$I_{(RISE)}$	0.40	-	0.39	-	0.28	-	-	-
Average Scores	0.76	.73(.07)	0.86	.79(.09)	0.73	.77(.06)	-	-

None of the novel object categories showed significant holistic processing (Aligned C-I). We used Dienes's (2014) method to evaluate the support for the null hypothesis for these effects. Because the theoretical prediction is unidirectional (the theory only predicts a positive congruency effect), we specified a H1 of a half-normal distribution with a maximum plausible value of .18 (the congruency effect observed for faces), and an *SD* of .09 (mean/2, as advised by Dienes, 2014). Sheinbugs is the category showing the largest congruency effect (.02) and the Bayes Factor (B) in favor of a difference is .27, representing substantial evidence in favor of the null. B for Greebles was .03 and that for Ziggerins was .04. The large (dz = 1.63) and significant congruency effect for faces was also much larger than that for the novel objects (paired *t*-tests against each of the other categories, *t*s > 8.4, *p*s < .0001). The reliability of the indices of holistic processing (Aligned C-I and Aligned C-$I_{(RISE)}$) varied between .28 and .42. The reliability of holistic processing for faces was .63 (note that this is achieved despite the fact that the reliability for each condition is lower than for novel objects, indicating that the low reliability for holistic effects with novel objects is most plausibly explained by the lack of holistic processing). The holistic processing scores (we used the Aligned C-I scores which were most reliable) were not significantly correlated across categories (see Table 2), providing no evidence of a domain-general influence as measured by these tasks. We evaluate the support for the null hypothesis for these correlations. Because the theoretical prediction is unidirectional (the theory predicts only positive correlations), for correlations among VHPT-NOs we specified a H1 of a half-normal distribution with a maximum plausible value of .4 (the maximum correlation given a reliability of .4 for each measure, and an *SD* of .2 (mean/2). Fisher Transform was applied to correlations (although at these low values it has little effect). The Bayes Factor (B) in favor of a correlation was .07 (Greeble/Ziggerin), .17 (Greeble/Sheinbug) and .04 (Ziggerin/Sheinbug). These values, all lower than .3, offer substantial support for the null hypothesis. This is of course expected given the absence of holistic effects on average, but we nonetheless wanted to test this, given recent work arguing for the importance of testing for individual differences even when average effects are null (Miller & Schwarz, 2018).

Table 2. Correlations of holistic processing between categories. Cutoff for $p = 0.05$, $r = 0.28$. Disattenuated correlations are not included because both correlations and reliabilities are low.

	Greebles HP	Ziggerins HP	Sheinbugs HP
Ziggerins HP	-.09		
Sheinbugs HP	.03	-.19	
Faces HP	-.01	-.04	-.16

As a second goal, we asked whether the part matching judgments in the VHPT-NOs may capture domain-general object recognition abilities. To this end, we averaged the performance in the congruent and incongruent conditions (in the aligned and the RISE conditions) for each task and considered correlations across categories. We did not average the two VHPT-F conditions because for faces there is evidence of different mechanisms at play in the congruent vs. incongruent conditions, as evidenced by a low correlation between those conditions, we. These correlations are presented in Table 3. Average performance on part matching in the 3 VHPT-NOs was significantly correlated, sharing between 29% and 50% of the variance according to the disattenuated correlations. Interestingly, face part matching during the incongruent trials of the VHPT-F were significantly correlated with part matching for Greebles (and these correlations were almost significant for Ziggerins, $p = .05$ and for Sheinbugs, $p = .07$). In fact, when part-matching for the three categories is averaged (after Fischer transform), part matching in the VHPT-NO is significantly correlated with performance on incongruent trials for faces ($r = .35$, $p = .005$) but not on incongruent trials ($r = .06$, $p = .34$). However, the difference between these two correlations did not reach significance ($Z = 1.45$, $p = .07$).

Table 3. Correlations between mean performance and categories (with correlations disattenuated for measurement error in parentheses). Cutoff for $p = 0.05$, $r = 0.28$.

	Greebles	Ziggerins	Sheinbugs	Faces cong.
Ziggerins	.56 (.71)			
Sheinbugs	.44 (.54)	.51 (.68)		
Faces cong.	-.03 (-.04)	.12 (.18)	.06 (.08)	
Faces inc.	.32 (.43)	.28 (.41)	.26 (.38)	.01 (.01)

In summary in Study 1, we found that the reliability of accuracy in each condition of the 3 VHPT-NOs was acceptable (> .7) whereas the reliability of congruency effects were lower (~.4). Average congruency effects suggested that novel objects were not processed holistically, and correlations produced very little evidence of shared variance accounting for congruency effects across categories. We also found evidence that part matching judgments in the VHPT-NOs tap into an ability common across three distinct categories of novel objects. While it was not predicted, and thus should be viewed as an exploratory result, performance on the VHPT-NO only correlated with face part judgments in the incongruent condition. Performance on congruent and incongruent trials for faces are not correlated at all (.01), and performance on incongruent trials with faces is related to subjects' performance matching non-face object parts.

Study 2

To test if part judgment and whole object memory depend on a common ability, we recruited subjects to perform one VHPT-NO and one NOMT, using different object categories. In this situation, none of the common variance should be due to the category or to the specific task and comparing performance on these tests should provide a test of the common ability that contributes to these tasks.

Subjects

Fifty-three subjects with IP addresses in the United States (29 males, mean age 36 with a SD of 11.5) were recruited from Amazon Mechanical Turk (we aimed for 50 and tested a few more in case of outliers). They were paid $5 to complete the study. A power analysis shows that with 50 subjects, the smallest correlation we can detect with 80% power is $r = .34$ (one-tail). Given that correlations among average part matching scores showed correlations above .4, and that the NOMTs have produced scores with reliability above .8 in prior work (Richler et al., 2017), we judged this to be sufficient power.

Procedure

All subjects first completed the VHPT-NO for Ziggerins as in Study 1, followed by a NOMT task with a family of asymmetric Greebles (from Richler et al., 2017). On this task, subjects learn each of 6 targets by studying each one from 3 different viewpoints, followed by 3 3-AFC test trials per target. They then study all 6 targets from one viewpoint in one display for 20 seconds, and are tested for their recognition of these 6 target objects in a series of 54 3-AFC trials in which targets and distractors could appear in different views but always the same view within a trial. Chance performance on this test is 24 correct, a perfect score is 72.

Results

Reliability and means are presented in Table 4. Reliability for the average performance on the VHPT-NO and the NOMT were both very high, and the correlation between the two tests was also high ($r = .76$, 95%, C.I. .62-.85, $p < .0001$, $r_{corrected} = .83$, Figure 5). This suggests that almost 70% of the non-error variance between these two tasks is shared, despite the tasks and the obhect categories being different, offering strong evidence that judgments in the VHPT-NO index the same object recognition ability measures in the NOMT (Richler et al., 2017).

Table 4. Reliability, means, and standard deviations for each condition of the VHPT-NO Ziggerin, and the average performance on that test as well as the NOMT with Greebles.

	Reliability	Mean	*SD*
Aligned C	0.88	0.70	0.16
Aligned I	0.87	0.64	0.16
RISE C	0.88	0.69	0.16
RISE I	0.91	0.67	0.16
Aligned C-I	0.24	0.07	0.09
RISE C-I	0.26	0.01	0.09
Aligned C-I$_{(RISE)}$	0.20	-	-
Average VHPT-NO	0.97	0.67	0.16
NOMT	0.86	0.62	0.13

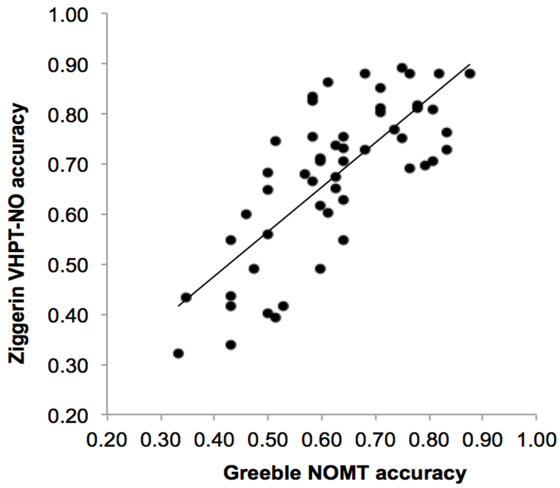

Figure 5. Correlation between average performance on the VHPT-NO for Ziggerins and the NOMT for Asymmetric Greebles.

General Discussion

Prior work found that expertise training with novel objects can lead to significant holistic processing in a composite task (e.g., Gauthier & Tarr, 1997; Wong et al., 2009; Chua et al., 2015, Richler et al., 2019). Indeed, using the VHPT-NO, we recently reported that holistic processing for artificial non-face objects increased monotonically with experience. With a great deal of statistical power (95%), the present work further adds evidence that without any experience these objects are not processed holistically, with substantial evidence in support of null effects. This provides reassurance that the tests are not sensitive to domain-general failures of selective attention. Notably, we also found no evidence of holistic processing for symmetrical Greebles, which have been argued by other authors to be processed in a face-like manner regardless of experience (Brants et al., 2011). These results suggest that the congruency effects measured by the VHPT-NOs in work with trained observers are entirely accounted by experience with a category. The congruency effects obtained in the VHPT-NOs by Chua & Gauthier (2019) reached, after 8 hours of training, almost 70% of the congruency effects we detected here for faces in our subjects. In fact, in terms of effect size (dz), the congruency effect in trained Sheinbug experts was larger than for faces (dz = 2 for Sheinbugs vs dz = 1.64 for faces). Of course, people have years, not hours, of experience with faces, which suggests that a plateau on the level of holistic processing, at least as measured by these tasks, is reached relatively rapidly.

As tests of holistic processing sensitive to experience, the VHPT-NOs should be useful in future studies interested in understanding what aspects of experience with objects leads to holistic processing. For instance, in Chua & Gauthier (2019), several aspects of experience were manipulated together, such that people who trained for longer also saw a larger number of unique objects and learned the names of more of these objects. In the real world, these different aspects of experience may often be confounded (e.g., a bird watcher with many more years of experience than another will likely have seen more birds and learned more semantic information about them), but theoretically, it is interesting to ask how these different aspects contribute to the development of holistic processing. Not all aspects are likely to be necessary—for instance, in Richler et al., (2019), subjects processed artificial objects holistically after less than 2 hours of training in a space invader game that required attention to the objects' unique shape, but did not teach subjects any names for the objects. But even if learning names is not necessary, it does not mean that it cannot facilitate holistic processing. Because training studies are expensive, a comprehensive comparison of different training experiences will likely require many studies, and the use of sensitive tests like the VHPT-NO can facilitate such a research program. A caveat to the present research is that we did not measure test-re-test reliability for the VHPT-NOs. This should be done in the future to provide evidence of the stability of the construct it measures. This would ideally be performed with trained observers, as in Chua and Gauthier (2019), so that the stability of holistic processing measurement can be assessed. We note that the VHPT-f, the test with faces that the VHPT-NOs are modeled after, demonstrated a 6 months test-rest reliability of .5, which was very close to its internal consistency (around .6), indicating that it captured a very stable construct.

Aside from measuring holistic processing, we found that part matching on the VHPT-NOs (averaging across congruent and incongruent conditions) provides reliable measurement of a domain-general object recognition ability related to that measured by the recently developed Novel Object Memory Tests (NOMTs, Richler et al., 2017). This is interesting because the VHPT-NO focuses on part judgments and does not require learning of objects over trials, whereas the NOMTs require learning to recognize 6 whole objects over a set of trials. Given the demonstration that the VHPT-NO can detect holistic processing (Chua & Gauthier, 2019), the present results demonstrate that the VHPT-NO can provide, in a single test, measures of both holistic processing and of object recognition ability. In addition, the VHPT-NO and the NOMT, given their shared variance yet distinct task requirements, may be useful together as part of test batteries to measure object recognition as a latent construct. Indeed, the creation of a variety of tests that can serve as different indicators of latent abilities of interest is critical in the development of a research program where individual differences in unobservable constructs can be studied while abstracting away from the more idiosyncratic aspects of specific measurements (Bollen, 2002).

Endnote

1. Before each block, an object with a red box and the words: In this block the target is the (top half/bottom third etc.) of the object. In a few blocks the words were incorrect (e.g., stated "top half" but the red box was a central part). The red box was correct, and was present on all the study trials, and there was no indication that subjects were confused about the target part (by that time they were used to following the red box). We corrected these instructions in the tests made available online.

Acknowledgments

This work was supported by the National Science Foundation (SBE-0542013 and SMA-1640681). K-W.C. was supported by a National Science Foundation graduate fellowship. We thank Susan Benear for help with data collection. Dr. Chua thanks Dr. Chu Chang Chua for continued guidance.

Author's Declarations

The authors declare that there are no personal or financial conflicts of interest regarding the research in this article.

The authors declare that they conducted the research reported in this article in accordance with the Ethical Principles of the Journal of Expertise.

The authors declare that they are not able to make the dataset publicly available but are able to provide it upon request.

References

Boggan, A. L., Bartlett, J. C., & Krawczyk, D. C. (2012). Chess masters show a hallmark of face processing with chess. *Journal of Experimental Psychology: General*, 141(1), 37-42.

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology, 53*, 605-34.

Brants, M., Wagemans, J., & Op de Beeck, H. P. (2011). Activation of fusiform face area by

Greebles is related to face similarity but not expertise. *Journal of Cognitive Neuroscience*, *23*(12), 3949-3958.

Bukach, C. M., Phillips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, & Psychophysics*, *72*(7), 1865-1874.

Chua, K. W., Richler, J. J., & Gauthier, I. (2015). Holistic processing from learned attention to parts. *Journal of Experimental Psychology: General*, *144*(4), 723-729.

Chua, K. W., & Gauthier, I. (2019). Domain-Specific Experience Determines Individual Differences in Holistic Processing. *Journal of Experimental Psychology: General*.

Curby, K. M., Huang, M., & Moerel, D. (2019). Multiple paths to holistic processing: Holistic processing of Gestalt stimuli do not overlap with holistic face processing in the same manner as do objects of expertise. *Attention, Perception, & Psychophysics*, *81*(3), 716-726.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2013). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*, *126*(1), 87-100.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 781.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is" special" about face perception? *Psychological Review*, *105*(3), 482-498.

Gauthier, I., & Tarr, M. J. (1997). Becoming a “Greeble” expert: Exploring mechanisms for face recognition. *Vision Research*, 1673-1682

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training ‘greeble’ experts: A framework for studying expert object recognition processes. *Vision Research*, 38(15-16), 2401-2428.

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2010). Holistic processing is not correlated with face-identification accuracy. *Psychological Science*, 21(1), 38-43.

Malgady, R. G., & Colon-Malgady, G. (1991). Comparing the reliability of difference scores and residuals in analysis of covariance. *Educational and Psychological Measurement*, 51(4), 803-807.

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2011). The context effect in face matching: Effects of feedback. *Vision Research*, 51: 2121-2131.

Meinhardt-Injac, B., Boutet, I., Persike, M. & Meinhardt, G., Imhof, M. (2017). From development to aging: holistic face perception in children, younger and older adults. *Cognition*, 158: 134-146.

Miller, J., & Schwarz, W. (2018). Implications of individual differences in on-average null effect effects. *Journal of Experimental Psychology: General*, 147:377-397.

Richler, J. J., Cheung, O. S., & Gauthier, I. (2011). Holistic processing predicts face recognition. *Psychological Science*, 22(4), 464-471.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology*, *3*.

Richler, J. J., Floyd, R. J., & Gauthier, I. (2014). The Vanderbilt Holistic Face Processing Test: A short and reliable measure of holistic face processing. *Journal of Vision*, *14*(11), 10.

Richler, J. J., Floyd, R. J., & Gauthier, I. (2015). About-face on face recognition ability and holistic processing. *Journal of Vision*, *15*(9), 15.

Richler, J. J., Wilmer, J. B., & Gauthier, I. (2017). General object recognition is specific: Evidence from novel and familiar objects. *Cognition*, *166*, 42-55.

Richler, J.J., Tomarken, A.J., Vickery, T.J., Ryan, K.F., Floyd, J.R., Sheinberg, D., Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to the measurement of change. *Psychological bulletin*, 92(3), 726.

Ross, D. A., Richler, J. J., & Gauthier, I. (2015). Reliability of composite-task measurements of holistic face processing. *Behavior research methods*, *47*(3), 736-743.

Rossion, B. (2013). The composite face illusion: A whole window into our understanding of holistic face perception. *Visual Cognition*, *21*(2), 139-253.

Sadr, J., & Sinha, P. (2004). Object recognition and random image structure evolution. *Cognitive Science*, 28(2), 259-287.

Wong, A. C. N., Palmeri, T. J., & Gauthier, I. (2009). Conditions for facelike expertise with objects: Becoming a Ziggerin expert—but which type?. *Psychological Science*, 20(9), 1108-1117.

Young, A. W., Hellawell, D., & Hay, D. (1987). Configurational informationin face perception. *Perception*, 16, 747–759.

Zhao, M., Bülthoff, H. H., & Bülthoff, I. (2016). A shape-based account for holistic face processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(4), 584-97.

Received: June 2019
Revision received: 19 November 2019
Accepted: 19 November 2019