Learned Attention to Parts Results in Holistic Processing

Kao-Wei Chua¹, Jennifer J. Richler¹, & Isabel Gauthier¹

¹Department of Psychology, Vanderbilt University

In press, JEP:G

Note: J.R. and I.G. played no role in the review process and did not have access to the manuscript file at any point

Corresponding Author:
Kao-Wei Chua
Department of Psychology, Vanderbilt University
Email: kao-wei.chua@vanderbilt.edu
Phone: 706-224-0905, Fax: 615-322-4706

REGULAR MAIL (via U.S. Postal Service)
Vanderbilt University
Abstract

Attention helps us focus on what is most relevant to our goals, and prior work shows that aspects of attention can be learned. Learned inattention to parts can abolish holistic processing of faces, but it is unknown whether learned attention to parts is sufficient to cause a change from part-based to holistic processing with objects. Here, we trained subjects to individuate non-face objects (Greebles) from two categories, Ploks and Glips. Diagnostic information was in complementary halves for the two categories. Holistic processing was then tested with Plok-Glip composites that combined the kind of part that was diagnostic or non-diagnostic during training. Exposure to Greeble parts resulted in general failures of selective attention for non-diagnostic composites, but face-like holistic processing was only observed for diagnostic composites. These results demonstrate a novel link between learned attentional control and the acquisition of holistic processing.

Keywords: holistic processing; learning; attention, perceptual expertise
Introduction

As we interact with the world, attention allows us to react to salient or surprising events (Theeuwes, 1994; 1998) and find information relevant to current goals (Folk et al., 1992; Bacon & Egeth, 1994). Attention can facilitate learning (Shiu & Pashler, 1992, Ahissar & Hochstein, 1993), but some learning effects are best characterized as changes in how we attend, or learned attention. Attention can be guided by statistical learning (Zhao, Al-Aidroos, and Turke-Brown, 2013), or reward and past selection history (Awh, Belopolsky, & Theeuwes, 2012). One kind of learned attention, item-specific control, occurs when mappings between a stimulus and an attentional set are learned (Jacoby et al., 2003; Bugg & Crump, 2012). While most studies on learned attention use simple stimuli, complex objects such as faces can also trigger attentional sets. For instance, in a study of cognitive control, subjects learned associations between face sex and specific proportions of congruent responses (Cañadas et al., 2013). Learned attentional settings can also be applied to categories of objects, with the control settings relevant at training transferring over to novel members of those categories (Bugg et al., 2011). Similarly, learned attention could account for phenomena related to perceptual learning (Nosofsky, 1986; Goldstone, 1994). For example, over the course of a categorization study, eye movements can reveal how subjects shift from attending to all stimulus dimensions equally to only attending to dimensions most diagnostic for categorization (Blair, 2009). Thus, attentional weights to stimulus dimensions can be shifted for object categorization.
Similarly, learned attention to dimensions of complex objects such as faces may account for expert visual object processing phenomena such as holistic processing, the tendency to process objects as unified wholes rather than as parts (Young et al., 1987). In a recent experiment (Chua et al., 2014), we examined the role of learned attention in holistic face processing. Subjects learned to individuate faces from two novel categories, Lunaris and Taiyos. Diagnostic information for identifying each face was placed in complementary halves of the two races. For example, only the top halves of Taiyos and the bottom halves of the Lunaris provided diagnostic information for individuation.

After training, subjects saw composites made of diagnostic and non-diagnostic face parts in the composite paradigm, a common measure of holistic processing (Farah, 1998; Richler & Gauthier, in press). In this task, subjects are asked whether the target half (e.g., top) of two sequentially presented composite faces (made of top and bottom halves from different faces) was the same or different while ignoring the other part (e.g., bottom). Holistic processing is inferred when subjects cannot ignore information in the task-irrelevant half, specifically for aligned face halves. In Chua et al. (2014), holistic processing was only found for face parts that were diagnostic at training, suggesting that learned attention to diagnostic face parts is associated with holistic face processing. In terms of learned attention, the features that define diagnostic face parts could act as cues that trigger an attentional set encouraging attention to those parts or features (Crump et al., 2008). When shown composite faces wherein both parts are diagnostic, subjects cannot help but to attend to parts that were previously diagnostic even when instructed not to, indicating a holistic processing strategy. Holistic processing is often described as having a perceptual basis (Rossion, 2013), and the role of face representations where
parts are not differentiated is often invoked (Tanaka & Farah, 1993). The Lunari-Taiyo study suggested that a history of attending to face parts can result in holistic processing.

This conclusion was also supported by the finding that when subjects were shown face composites made of parts that were non-diagnostic during training, no holistic processing was observed. Face parts with a history of not being attended do not trigger obligatory attention, resulting in little to no interference from the task-irrelevant part during the composite task. However, while processing diagnostic and non-diagnostic composites was clearly different, novices who had never seen Taiyo or Lunari faces also processed them holistically. It is possible that the main effect of training was to abolish holistic face-like processing for non-diagnostic composites rather than to increase holistic processing for diagnostic composites.

Reducing holistic processing through learned inattention does not guarantee that learned attention is sufficient to cause face-like holistic processing. Therefore, we take on the bigger challenge of addressing this question directly, by measuring the acquisition of holistic processing with novel objects. We trained participants to individuate Greebles, objects that novices do not process holistically. We used two kinds of Greebles that contained diagnostic information in different parts and then tested holistic processing for Greebles combining parts never presented together before. Since holistic processing in novel objects such as Greebles may be sensitive to factors such as context (Richler, Bukach, Gauthier, 2009), we included a phase-scrambled baseline to better detect face-like holistic processing. Our results establish for the first time a link between learned attentional control settings and the acquisition of holistic processing.
Methods

Subjects

Eighty subjects participated in a within-subjects design, randomly assigned to two counterbalancing conditions: a GlipTop/PlokBottom training condition (18 male, 22 female, mean age = 21.6) or a GlipBottom/PlokTop condition (16 male, 24 female, mean age = 21.9). Group assignment dictated which part was diagnostic for each Greeble category during individuation training. A control group of 40 subjects received no training (16 male, 24 female, mean age = 20.5). Subjects received $15/hour for participation. The study was approved by the Vanderbilt University IRB. Sample size was predetermined based on the interaction between group, congruency, and alignment in Chua et al. (2014), $\eta^2_p = .04$, and on our expectation that using fewer parts in the composite task should improve its reliability (see Ross et al., 2014) to at least $r = .4$. With 80 subjects and an alpha of .05, power for the critical interaction should reach .90.

Stimuli

Stimuli were asymmetrical Greebles (Gauthier & Tarr, 1997; Rossion et al., 2004). Ploks and Glips had distinct body shapes and textures (Figure 1), and Ploks have parts pointing down, whereas Glips have parts pointing up. All Greebles were presented in grayscale and tilted 40 degrees clockwise to facilitate making composites without cutting any of the parts.

During individuation training, different stimulus sets were used depending on counterbalancing condition. For example, in the GlipTop/Plok Bottom condition, three Glip bottom halves were combined with 10 unique top halves each, and three Plok tops
halves were combined with 10 unique Plok bottoms each. This resulted in 30 Glips that varied 10 times more in the top than the bottom and 30 Ploks that varied 10 times more in the bottom than the top. A separate set of 60 Greebles was created for the GlipBottom/Plok Top condition, with the reverse assignment of variability to parts.

For the composite task, two sets of five unique top and bottom halves from the two families were used that were not seen during training. The composite Greebles varied on the top and bottom. Therefore, all subjects were tested on Greebles with the same amount of variation; any differences in composite task performance could only be attributed to training. The top and bottom Greeble halves were randomly combined to form Plok-Glip composites (400 x 400 pixels). A white line 6 pixels thick separated the Greeble halves so that it was unambiguous where the top half ended and the bottom half began. Misaligned composites were made by shifting the top half 35 pixels to the left and the bottom half 35 pixels to the right. Note that an advantage of the Greebles being tilted is that the entire object remains inscribed within a box approximately the same width as the aligned object, eliminating a potential confound with the standard misalignment manipulation in which misaligned trials are wider. Another baseline condition was added wherein task-irrelevant parts were phased-scrambled (83%) using an algorithm (Sadr & Sinha, 2004) that controls for luminance, contrast, and spatial frequency. With faces, performance in a phase-scrambled baseline is indistinguishable from that in a misaligned baseline (Richler, Floyd, & Gauthier, 2014). An example of a phase-scrambled composite task trial is shown in Figure 2.

*Experimental Procedure*
Subjects completed 3 sessions of individuation training (approx. 60 minutes each) over the course of a week. After the third session, subjects completed the composite task.

*Individuation Training*

Subjects learned unique names for a total of 16 Greebles, 8 for each category. During training trials, subjects were shown Greebles with one syllable names (e.g. Awg, Dak, Ror). Names were randomly assigned to objects for each subject. In all trials, subjects pressed the first letter of the Greeble’s name. Training trials were followed by test trials where Greebles were presented without names. If a Greeble appeared that had no learned name association, subjects pressed “n” to indicate “no name.” There were 22 “no name” Greebles used during training. Out of the thirty Greebles for a given category, only eight were assigned names for participants to learn. Incorrect responses were followed by feedback showing the correct name (including “no name”).

There were three phases on each day of training. The details of each phase and the number of blocks per phase on each training day are shown in Table 1. All trained Greebles were introduced by the end of Day 1 and were repeated on subsequent training days.

*Composite Task*

All Greebles in the composite task were Plok-Glip or Glip-Plok composites made of parts not seen during training (Figure 3). Five top and five bottom parts from each category, as well as their phase-scrambled versions, were used across all trials. For each subject, depending on the counterbalancing training condition, half of the composites
were made of parts similar to parts that were *diagnostic* during training, and the other half made of parts similar to parts that were *nondiagnostic* during training. Each trial started with a fixation cross (200 ms), followed by a study Greeble (200 ms), a blank screen (500 ms), and a test Greeble (200 ms). Subjects were instructed to judge if the cued half of the study and test composites were the same or different, and to ignore the other, irrelevant half. On congruent trials, the cued and irrelevant halves were associated with the same response (e.g., both parts same, or both parts different); on incongruent trials, the irrelevant and cued halves were associated with different responses (e.g., one part same, the other part different). A congruency effect (better performance on congruent vs. incongruent trials) indicates an inability to selectively attend: the irrelevant object half influenced performance, despite instructions to ignore it. On misaligned trials, only the test Greeble was misaligned to prevent pseudo-holistic effects that are not sensitive to configuration (see Richler, Bukach, & Gauthier, 2009; Richler, Wong, & Gauthier, 2011). The signature of holistic processing is a congruency x alignment interaction, with larger congruency effects on aligned than misaligned trials (Richler & Gauthier, in press). Likewise, congruency can still be defined on phase-scrambled trials (Figure 2), depending on whether the task-irrelevant phase-scrambled half is the same or different between study and test. Therefore, holistic processing using the new phase-scrambled baseline would be defined as a congruency x condition interaction, with larger congruency effects on aligned than phase-scrambled trials.

There were 15 trials for each combination of composite type (diagnostic/nondiagnostic), congruency (congruent/incongruent), alignment (aligned/misaligned/phase-scrambled), cued part (top/bottom) and correct response
(same/different), for a total of 720 trials. Cued part was blocked, with order counterbalanced across subjects. All other factors were randomized.

Results

Individuation Training

From Day 1 to Day 3, subjects become more accurate (Day 1: M = .92; Day 3: M = .96; $F(2,79) = 110.6, p < .0001, \eta_p^2 = .58$) and faster (Day 1: M = 946.8 ms; Day 3: M = 709.0 ms; $F(2,79) = 246.2, p < .0001, \eta_p^2 = .75$) on the training task.

Composite Task

Data from two trained subjects and two control subjects were removed for below chance performance, leaving 78 subjects trained to individuate objects from both Greeble categories, and 38 untrained control subjects. Trials with reaction times less than 100 ms or greater than 2000 ms were discarded (1.62% of trials).

Sensitivity ($d'$) for the control and trained subjects (separated into diagnostic composites and non-diagnostic composites) for aligned, misaligned, and phase-scrambled trials is presented in Figure 4. We found no evidence of holistic processing of Greebles in control subjects, that is, there was no significant interaction between congruency (congruent/incongruent) and Trial Type (aligned/misaligned/phase-scrambled) ($F(2,74) = .32, p = .73, \eta_p^2 = .009$).

To assess training effects, all analyses were conducted twice, once comparing aligned trials to misaligned trials as the baseline, and once comparing aligned trials to phase-scrambled trials as the baseline. In all cases holistic processing is indexed by a
congruency (congruent/incongruent) x Trial Type (aligned/misaligned or aligned/phase-scrambled) interaction. With each baseline, we first conducted a Condition (control/diagnostic/non-diagnostic) x Congruency (congruent/incongruent) x Trial Type (aligned/misaligned, or aligned/phase-scrambled) ANOVA, treating Condition as a between subjects variable even though diagnostic and non-diagnostic conditions are from the same individuals. Based on Ross et al. (in press), the correlation between holistic processing across conditions was expected to be small, so there should be a negligible cost in power for this strategy\(^1\). To unpack any interaction with Condition, we ran within-subject ANOVAs comparing the diagnostic and non-diagnostic conditions in the trained subjects, and between-subject ANOVAs comparing each trained condition to the control group.

*Misaligned baseline.* The three-way interaction between Condition (control/diagnostic/non-diagnostic), Trial Type (aligned/misaligned), and Congruency was significant, \(F(2,191) = 3.16, p = .04, \eta_p^2 = .03\). With trained subjects only, the Condition x Congruency x Trial Type interaction was only marginally significant, \(F(1,77) = 3.06, p = .084, \eta_p^2 = .04\), with significant holistic processing (Congruency x Trial Type interaction) for diagnostic composites, \(F(1,77) = 6.67, p = .01, \eta_p^2 = .08\), but not for non-diagnostic composites, \(F(1,77) = 0.12, p = .73, \eta_p^2 = .002\). As expected, the interaction between Condition, Congruency, and Trial Type was not significant when the

\(^1\) As expected, there was no significant correlation between holistic processing in the diagnostic and non-diagnostic conditions using the misaligned baseline (\(r = -.10, p = .37\)) and that using the phase-scrambled baseline was significant but modest in size (\(r = .36, p = .001\)). This could represent an advantage of the phase-scrambled baseline for future work.
control group was compared to trained-non-diagnostic, $F(1,114) = 0.59, p = .44, \eta^2_p = .0005$, indicating that exposure to non-diagnostic parts during training did not result in holistic processing. In contrast, the same three way interaction was significant when the control group was compared to trained-diagnostic, $F(1,114) = 4.50, p = .036, \eta^2_p = .038$. This reveals more holistic processing for trained subjects judging diagnostic composites than for the untrained control group. In summary, with the misaligned baseline, the within-subject training effect was marginal, but only diagnostic composites were processed holistically by trained subjects and not untrained control subjects.

**Phase-scrambled baseline.** The three-way interaction between Condition, Trial Type, and Congruency was significant, $F(2,191) = 6.24, p < .002, \eta^2_p = .06$. With trained subjects only, the Diagnosticity x Congruency x Trial Type interaction was also significant, $F(1,77) = 9.85, p < .002, \eta^2_p = .11$, with significant holistic processing (Congruency x Trial Type interaction) for diagnostic composites, $F(1,77) = 8.62, p < .0001, \eta^2_p = .20$, but not for non-diagnostic composites, $F(1,77) = 1.38, p = .24, \eta^2_p = .02$. As expected, the interaction between Condition, Congruency, and Trial Type was not significant when the control group was compared to trained-non-diagnostic, $F(1,114) = 1.31, p = .25, \eta^2_p = .001$. However, this same interaction was significant when the control group was compared to trained-diagnostic, $F(1,114) = 9.97, p = 0.002, \eta^2_p = .08$, with greater holistic processing for diagnostic composites than the control group. In summary, there was more holistic processing in the diagnostic condition than in the non-diagnostic condition for trained subjects, and untrained control subjects did not process these objects holistically.
Discussion

We previously provided evidence that learned inattention to non-diagnostic parts can abolish holistic processing with faces (Chua, Richler, & Gauthier, 2014). Here, we used a similar design with Greebles, novel objects that are not processed holistically by novices, and demonstrate that a history of learned attention to diagnostic parts is sufficient for the acquisition of holistic processing. This reveals a link between learned attentional strategies and visual object processing. Accounts of item-specific attentional strategies suggest that stimulus features and responses are jointly encoded in memory so that attentional filters are later cued by the perception of features with which they were previously associated (Crump et al., 2008). In the current study, novices do not process Greebles holistically at all, and trained subjects only exhibit holistic processing for diagnostic Greeble parts. Thus, learned attention to diagnostic object parts is necessary for the acquisition of holistic processing. We propose that when trained subjects were shown Greebles made of diagnostic parts in the composite task, they retrieved an attentional filter related to their past history of devoting attention to those parts. When both Greeble parts in the composite task are diagnostic, subjects cannot selectively attend to one Greeble part, a hallmark of face-like holistic processing.

In contrast, holistic processing was not found for non-diagnostic Greeble parts. However, training did have an impact since there was a main effect of congruency in the trained group not found in Greeble novices ($F(1,114) = 4.10, p < .05, \eta_p^2 = .044$). Non-diagnostic parts may have become associated with an attentional routine to look for more useful information in a different part of the object. This is consistent with accounts of
learned inattention and blocking during learning (Kruschke & Blair, 2000; Kruschke, 2003), whereby people learn to ignore irrelevant cues. In contrast, the attentional routines that our subjects acquired for diagnostic parts were specific to the aligned configuration, as typically observed for holistic processing of faces (Richler & Gauthier, in press). The configural specificity of these congruency effects may be akin to learned attention that is specific to the encoding context (e.g., Chun & Jiang, 1998). However, the composites our subjects were tested with represent a new context in many ways: new exemplars were shown in a new task, parts were combined with parts from a different category with which they were never paired during training, and over the course of the task the composite Greebles varied equally in both parts, unlike during training. Despite these changes, the task-irrelevant diagnostic parts were difficult to ignore when they were aligned with a task-relevant diagnostic part. This points to spatial configuration as particularly critical for holistic processing, but not because of a special representational status for aligned objects. Rather, configuration may influence the allocation of attention during composite task judgments, perhaps through object-based attention for parts that are grouped perceptually (Vecera & Farah, 1994; Baker, Olson & Behrmann, 2004).

Other work suggests a role of grouping in the composite task, based on evidence that misaligned colored backgrounds behind aligned face parts reduce holistic processing (Curby et al., 2013).

Our effects are similar to those obtained in a number of other paradigms where attention is biased toward certain information with a history of being attended. Because our effects are measured as reduced selective attention in a congruency task, they bear a similarity to a literature on item-specific control during Stroop tasks (Canadas et al.,
2012; Bugg et al., 2011). Because our effects lead to spatial bias attached to specific visual context, they are similar to contextual cueing in which subjects implicitly learn spatial invariants in visual scenes (Chun & Jiang, 1998). However, in contrast to the contextual cueing and item-specific control paradigms, our subjects carried attentional settings from a naming task to the composite task. This renders explanations based on conflict (which did not exist during naming) or learned response associations (because naming responses do not predict congruency) improbable. In the category learning literature, attention can shift to a dimension that was useful in a previous categorization task, and like our effects, this can occur with transfer objects and to a new task (Goldstone, 1994), and sometimes last for days (Folstein, Newton, Van Gulick, & Palmeri, 2012). As in the item-specific control literature, the effects can be stimulus-specific (Aha & Goldstone, 1992; Van Gulick & Gauthier, 2014) and may become category specific after multiple exemplars are encoded (Nosofsky, 1986). Category-specific (or context specific) control has also been described within Stroop paradigms (e.g., Bugg, Jacoby & Toth, 2008). Finally, our results are also similar to some effects obtained in the task-switching literature, when task-stimulus associations have long-term effects, primarily on task switch trials when there is conflict, as in the composite task (Waszak et al., 2003). Which of these effects is most similar to holistic processing remains to be determined. Critically, these attentional accounts do not require the encoding of unitary representations, a common account for holistic processing (Tanaka & Farah, 1993).

Together with our previous finding that holistic processing was only evident when both parts of a test face benefitted from a history of attention (Chua et al., 2014), the
present results suggest that failures of selective attention to diagnostic parts are qualitatively different from those to non-diagnostic parts. Face-like holistic effects appear to require that both the task-relevant and task-irrelevant parts have a history of being attended and that the parts be perceptually grouped, allowing this attentional effect to apply to the entire object. More work is required to account for this configurally-specific learned attention in computational models of attention, but our results provide clear answers to the questions that motivated this work: holistic processing can be acquired for non-face objects, and all that is required is a history of attention to parts.

**Acknowledgements**

This work was supported by the NSF (Grant SBE-0542013), VVRC (Grant P30-EY008126) and NEI (Grant R01 EY013441-06A2). Riaun Floyd, Amit Khandhadia, and Nicole Goren aided with data collection. We would like to thank Matt Crump for his comments on the manuscript. We would also like to thank Dr. Chu Chang Chua for continued guidance.
Table 1. Training structure for individuation training for one Greeble category. Both Ploks and Glips were trained in the same way.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Greebles</th>
<th>Training trials/block</th>
<th>Test trials/block</th>
<th>Day 1</th>
<th>Day 2</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4 named</td>
<td>8</td>
<td>21</td>
<td>2 blocks</td>
<td>----</td>
<td>2 blocks</td>
</tr>
<tr>
<td></td>
<td>6 unnamed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>6 named</td>
<td>12</td>
<td>42</td>
<td>2 blocks</td>
<td>----</td>
<td>2 blocks</td>
</tr>
<tr>
<td></td>
<td>14 unnamed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>8 named</td>
<td>24</td>
<td>63</td>
<td>6 blocks</td>
<td>10 blocks</td>
<td>6 blocks</td>
</tr>
<tr>
<td></td>
<td>22 unnamed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>Trials</td>
<td></td>
<td></td>
<td>1008</td>
<td>1260</td>
<td>1008</td>
</tr>
</tbody>
</table>
Figure 1. Example Glips and Ploks for a subject who saw Glips with diagnostic top halves and Ploks with diagnostic bottom halves. The part that was diagnostic for each family was counterbalanced between two groups of subjects. Note that the non-diagnostic half did vary (there were 3 parts for each category, only one is shown), but there was ten times more variation for the diagnostic half.
Figure 2. Example of an incongruent phase-scrambled trial in the composite task, where the top half was cued and the correct answer is “same.” In this condition, the task-irrelevant half is phase-scrambled. The cued halves are the same and the phase-scrambled halves are different, so this is an example of an incongruent phase-scrambled trial.
Figure 3. Examples of stimuli for the Glip Top/Plok Bottom group. For this group, the tops of Glips and the bottom of Ploks were diagnostic during individuation training (top panel). During the composite task, stimuli were created from diagnostic and nondiagnostic Greeble parts (bottom panel). Composite task stimuli were either aligned, misaligned, or the task-irrelevant part was phase-scrambled.
Figure 4. Sensitivity (d’) as a function of trial type (aligned/misaligned/phase-scrambled) and congruency for the control group, and non-diagnostic and diagnostic composites for the trained group. The aligned and misaligned conditions are connected to highlight the typical congruency x alignment interaction that is only observed in the diagnostic condition.
References


face perception? *Psychological review, 105*(3), 482.


