CrossMark

How holistic processing of faces relates to cognitive control and intelligence

Isabel Gauthier [1] · Kao-Wei Chua [2] · Jennifer J. Richler [1]

Abstract

The Vanderbilt Holistic Processing Test for faces (VHPT-F) is the first standard test designed to measure individual differences in holistic processing. The test measures failures of selective attention to face parts through congruency effects, an operational definition of holistic processing. However, this conception of holistic processing has been challenged by the suggestion that it may tap into the same selective attention or cognitive control mechanisms that yield congruency effects in Stroop and Flanker paradigms. Here, we report data from 130 subjects on the VHPT-F, several versions of Stroop and Flanker tasks, as well as fluid IQ. Results suggested a small degree of shared variance in Stroop and Flanker congruency effects, which did not relate to congruency effects on the VHPT-F. Variability on the VHPT-F was also not correlated with Fluid IQ. In sum, we find no evidence that holistic face processing as measured by congruency in the VHPT-F is accounted for by domain-general control mechanisms.

Keywords Face recognition · Individual differences · Stroop · Flanker

Introduction

The idea that faces are processed more holistically than most other objects has been central to the study of face recognition for several decades (Maurer, Grand, & Mondloch, 2002; Richler, Floyd, & Gauthier, 2014). Farah, Wilson, Drain, and Tanaka (1998) proposed that while face recognition depends to some extent on the same part-based processes involved in object recognition, it also recruits a holistic mechanism that represents faces as "wholes." Many tasks have been designed to target this mechanism but there are a surprisingly large number of possible meanings for holistic processing, including sensitivity to configural information, sampling of global information and the "whole being greater than the sum of its parts"(see Richler, Palmeri, & Gauthier, 2012 for review). A different meaning of holistic processing, distinct from those above, is that of a failure of selective attention to parts. The idea of holistic processing as a failure of selective attention originates in the Farah et al. studies (Farah et al., 1998) in which subjects were asked to make matching judgments about face parts while

ignoring other parts and made more errors when task-irrelevant parts were incongruent with the correct decision. Holistic processing in these tasks is operationalized by greater discriminability in congruent versus incongruent trials. Using this definition, there are larger holistic effects for faces than non-face objects (Farah et al., 1998; Richler, Mack, Palmeri, & Gauthier, 2011), although similar effects can be obtained with non-face objects in expert observers (Bukach, Phillips, & Gauthier, 2010; Boggan, Bartlett, & Krawczyk, 2012, Chua, Richler, & Gauthier, 2015; Wong, Palmeri, & Gauthier, 2009). Holistic processing as defined by failures of selective attention also correlates with activity in the inferior temporal cortex in subjects who vary in expertise, providing evidence of predictive validity for this measure (Gauthier & Tarr, 2002; Wong, Palmeri, Rogers, Gore, & Gauthier, 2009).

The Vanderbilt Holistic Processing Test for faces (VHPT-F, Richler et al., 2014) is the first standard test designed to measure individual differences in holistic processing, defined as a failure of selective attention. It shows convergent validity with the complete design of the composite test (Wang, Ross, Gauthier, & Richler, 2016). The VHPT-F typically achieves higher internal consistency (>.6) relative to various implementations of the composite task (e.g., DeGutis, Wilmer, Mercado, & Cohan, 2013; Ross, Richler, & Gauthier, 2015) by including trials that vary the size of the part that is selectively attended, thereby being more sensitive to a larger range of holistic processing in individuals. In prior work, the VHPT-F has shown a test-retest reliability after 6 months of .52, about as high as its internal

✉ Isabel Gauthier
isabel.gauthier@vanderbilt.edu

[1] Department of Psychology, Vanderbilt University, PMB 407817, 2301 Vanderbilt Place, Nashville, TN 37240-7817, USA

[2] Department of Psychology, New York University, New York, NY, USA

consistency, suggesting that it measures a stable construct (Richler et al., 2014). Holistic processing in the VHPT-F is not related to performance on the Cambridge Face Memory Test or CFMT (Richler, Floyd, & Gauthier, 2015). Richler et al. (2015) conducted studies suggesting that the CFMT may only relate to holistic measures that repeat a small number of faces many times over trials, suggesting the correlation is an artifact due to exemplar learning in both tasks.

Is holistic processing of faces related to cognitive control?

The fact that there are stable individual differences in holistic processing that do not correlate with face recognition ability in the normal population leads us to ask what this stable variability reflects. Some (Rossion, 2013) have argued that holistic processing measured as a congruency effect, as in the VHPT-F, reflects cognitive control processes, i.e., the ability to select sensory targets and motor responses in the face of conflict. There are some reasons to doubt that may be the case. First, congruency effects in this task are not observed for non-face objects in novices (Meinhardt-Injac, Persike, & Meinhardt, 2014; Richler et al., 2011), unlike failures of selective attention observed in other tasks (e.g., Flanker effect; Eriksen & Eriksen, 1974). Second, response interference was found not to contribute substantially to congruency effects for faces (Richler, Cheung, & Wong, 2009). In addition, congruency effects for faces differ qualitatively from Stroop effects, where the easier of two dimensions is less susceptible to interference (MacLeod, 1991; Melara & Mounts, 1993). In the standard composite task that uses faces divided horizontally across the middle, subjects are better at judgments on top halves but experience more interference from incongruent face bottom halves. However, these arguments are based on group-averaged effects and experimental effects and individual differences even in the same task can reflect different influences (Hedge, Powell, & Sumner, 2017). Here, we wanted to determine the relationship between individual differences in cognitive control effects in Stroop and Flanker tasks and holistic processing of faces.

Correlations between the congruency effect in holistic processing tasks and general selective attention/cognitive control measures have not been explored. Measures of cognitive control present many of the same problems as measures related to face recognition when it comes to individual differences. That is, difference scores are used (performance in one condition relative to a baseline) for theoretical reasons, effects often lack sufficient reliability and the same cognitive control task using different kinds of stimuli do not always correlate (e.g., Kindt, Bierman, & Brosschot, 1996). Despite these limitations, some studies find common variance across versions of the task, for instance across three different Stroop tasks (color-word, number quantity, word-position; Salthouse & Meinz, 1995). Here, we include three Stroop and three Flanker tasks to relate to the VHPT-F.

Is holistic processing of faces related to intelligence?

The ability to recognize faces or objects has been found to be relatively distinct from general intelligence (Richler, Wilmer, & Gauthier, 2017; Shakeshaft & Plomin, 2015; Van Gulick, McGugin, & Gauthier, 2016; Wilmer et al., 2010). However, holistic processing could be different. Fluid intelligence, working memory capacity, and cognitive control are overlapping constructs according to some authors (Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999), although this can depend on the specifics of the task. For instance when incongruent trials are relatively infrequent (as opposed to very frequent), Stroop interference correlates with intelligence (Kane & Engle, 2003). Here, we used 50% incongruent trials in our cognitive control tasks, to keep the proportion the same as on the VHPT-F, so on that basis we might not expect interference in our selective attention tasks to be related to intelligence.

However, another recent study links intelligence with perceptual suppression in motion tasks, with high intelligence strongly associated with rapid perception of small moving targets and with worse perception of motion for larger motion displays (Melnick, Harrison, Park, Bennetto, & Tadin, 2013; Tadin, 2015). While the specific mechanisms invoked to explain these results in motion perception are unlikely to be relevant to face processing, the authors suggested that such individual differences could reflect more general principles related to how several systems in the brain generally must ignore distracting information to achieve task goals. Given that a relation between intelligence and processes that appear to be both lower level (motion perception) and arguably higher level (working memory capacity) than face perception, it seems reasonable to explore a contribution of intelligence to holistic processing of faces.

Methods

Subjects

We collected data from 134 individuals from the community in and around the Vanderbilt Campus (91 female, mean age = 21.34 years, SD = 3.81). We expected effect sizes to be small between congruency effects in cognitive control tasks, based on other work (e.g., Salthouse & Meinz, 1995, rs varying between .14 and .52 for pairs of Stroop tasks). We chose to calculate power based on an effect size of r=.25, and determined that we would need 120 subjects to detect such an effect with 80% power at an alpha of .05. We collected more data in prevision of possible outlier exclusions.

We only excluded outliers if they were an extreme outlier on one of the congruency measures (VHPT-F, Stroop and Flanker tasks), defined as a value above *high hinge* + 3.0 (*high hinge – low hinge*) or below *low hinge* - 3.0 (*high hinge – low*

hinge), with *low* and *high hinges* being the 25th and 75th percentile. Data for four individuals were flagged using this procedure (two in the color/word Stroop task, one in the quantity/number Stroop task and one in the Flanker arrow task). The final sample had 130 subjects (88 female, mean age = 21.19, SD = 3.72).

Materials and procedure

Subjects performed the tasks in four blocks (fluid IQ tasks, Stroop tasks, Flanker tasks and VHPT-F), with the order of these blocks randomized across subjects. Within each block, task order and trial order was fixed. In the fluid IQ block, subjects completed the Ravens matrices first, followed by letter sets and numbers. In the Flanker block, subjects completed the letter task first, then the arrow and finally the color task. In the Stroop block, subjects completed the color/word task first, followed by the quantity/number task and finally the size/word task.

Vanderbilt Holistic Face Processing Test, VHPT-F (ver. 2.1)

This test has been described in detail in Richler et al. (2014). The only difference from version 2.0 is that 16 trials were modified from version 2.0 based on data from 525 subjects across several prior studies, because they were not correlated with overall condition scores (e.g., congruent trials that were negatively correlated with overall performance on congruent trials). On each trial, a study composite face was shown for 2 s and a test display followed that showed three composite faces. Subjects were instructed to select the composite face containing the target part with the same identity (but different image) as the target part in the study composite, while ignoring the rest of the face. The target part was outlined in red at study and test. The correct target part was paired with either the same distractor parts (congruent trials) or different distractor parts (incongruent trial) relative to study (see Fig. 1a). There were nine blocks of 20 trials, each with a different target segment (top 2/3, bottom 2/3, top half, bottom half, top 1/3, bottom 1/3, eyes, mouth, nose), for a total of 180 trials.

Stroop tasks

The Stroop tasks (Fig. 1b) were inspired from those used by Salthouse and Meinz (1995). On the color/word Stroop, subjects reported the color of a word and the word was presented in either a congruent or incongruent color (options were blue, green, red, and yellow for words and colors). In the quantity/number Stroop, subjects reported the quantity of a group of numbers while the numbers themselves were either congruent or incongruent with the quantity (options were 1, 2, 3, and 4). On the size/word Stroop, subjects reported the size of the words small, medium, or large printed in 12-, 50-, or 100-

point font. On each Stroop task there were 48 trials, half of which were congruent. On each block, trials began with a 500-ms fixation cross followed by a 250-ms inter-stimuli interval and the stimulus was presented until a response was made.

Flanker tasks

The Flanker tasks (Fig. 1c) were inspired from those used by (Peschke, Hilgetag, & Olk, 2013). On the letter Flanker task, subjects were asked to respond to a central letter (A or D) flanked by either four congruent (e.g., AAAAA) or incongruent (e.g., AADAA) Flankers. On the arrow Flanker task, subjects were asked to respond to a central arrow (< or >) flanked by either four congruent (e.g., >>>>>) or incongruent (e.g., >><>> Flankers. On the color Flanker task, subjects were asked to respond to the central color of 5 colored squares flanked by either four congruent (e.g., same in color) or incongruent (e.g., different in color) squares (the colors were red and blue). On each Flanker task there were 48 trials, half of which were congruent. On each block, trials began with a 500-ms fixation cross followed by a 250-ms inter-stimuli interval and then the stimulus was presented until a response was made.

Fluid intelligence tasks

We used three tasks known to load on fluid intelligence and used in several previous studies (see Fig. 1d; Redick et al., 2013; Van Gulick et al., 2016). There were specific time limits for each block, but no time limits for a response on each trial and within each block, and trials were ordered from easiest to most difficult with practice trials preceding every block. Subjects completed as many of 18 trials as possible in 10 min from the Raven's Advanced Progressive Matrices (RAPM; Raven, Raven, & Court, 1998). In Letter Sets (Ekstrom, French, Harman, & Dermen, 1976) subjects saw five sets of letter strings with all but one of the letter strings following a specific rule. Subjects had 7 min to complete as many of the 30 trials as possible. In number series (Thurstone, 1938), each trial showed an array of 5–12 numbers forming some type of pattern. Subjects chose which of five number options would follow the presented array (e.g., if the array was 1 2 3 4 5, the correct response was 6). They had 5 min to complete as many of the 15 trials as possible.

Results

We report descriptive statistics for all mean congruency effects in Table 1, and for all individual conditions and for the congruency effects in Table 2. All the congruency effects were significant in accuracy and in response times. Despite correct response times in each condition often showing moderate to good reliability (.52–.97), none of the congruency effects in mean response times were reliable. The reliability of
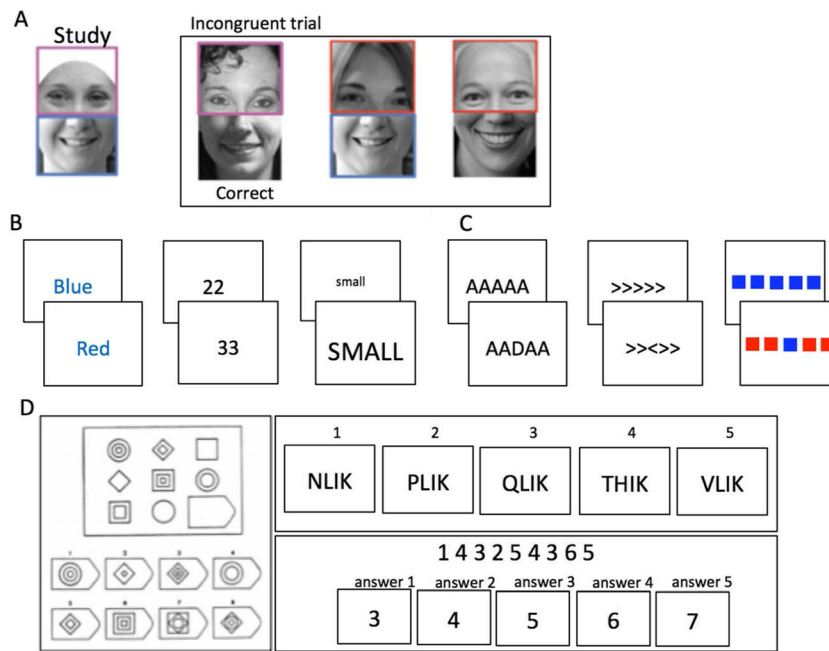
Fig. 1 Sample trials for the different tasks. (**A**) An incongruent trial from the VHPT-F. Color outlines indicate the identity of the face parts. On the test, only the top part on this trial would be outlined in red to indicate the part that is relevant on this trial. (**B**) Examples of congruent and incongruent trials for the three Stroop tasks and the three Flanker tasks. (**C**) Example trials from the three intelligence tasks, Raven's Matrices, Letter Sets, and Number Series

difference scores drops to 0 as the correlation between the two conditions begins to approach the average of their separate reliabilities. In all our Stroop and Flanker tasks, this was the case. To use one illustrative example, for the Stroop color/word condition, the correlation between congruent and incongruent trials was r = .78, which is almost as high as the average (Fisher-transformed) reliability of the two conditions (.81). This suggests that the two conditions are perfectly correlated to the limit of the measurement error and there is no reliable difference construct that can be extracted. We found the same result when using mean response time as opposed to mean correct response times. The problem of unreliable individual differences in cognitive tasks, especially for difference scores,

has been thoroughly discussed elsewhere (Hedge et al., 2017; Ross et al., 2015).

Accuracy congruency effects were generally more reliable, with the exception of

the Flanker color task, also because the correlation between the congruent and incongruent conditions (r=.67) was as high as the most reliable of the two individual conditions. Despite highly significant mean congruency effect across subjects and moderately reliable individual differences in the congruency effects, one concern with accuracy is that performance was very high in each of the individual Stroop and Flanker conditions. Such a ceiling on accuracy could cause a range restriction and limit correlations with other tasks. To address both the limitations in reliability of the response times and the ceiling effects in accuracy, we calculated a combined measure called the Rate Correct Score (RCS, (Vandierendonck, 2017; Woltz & Was, 2006). The RCS is the number of correct responses divided by the sum of correct and incorrect response times (RTs) and is easily interpretable as the number of correct responses per second. In a comparison of several methods to combined accuracy and response times (Vandierendonck, 2017), RCS was found preferable to other measures including inverse efficiency. RCS recovers the information from accuracy and response times and can account for a larger proportion of the variance than the separate measures. Table 1 shows that the effect size for the mean congruency effects were larger for RCS than either accuracy and correct RTs, and all the congruency effects were highly significant. RCS scores achieved a level of reliability higher than response times and

Table 1 Effect size of the mean congruency effects in each task, for accuracy, correct response times, and rate correct score (Vandierendonck, 2017)

	Congruency effect size, Cohen dz (p-value)		
	Accuracy	Correct RTs	RCS
VHPT Faces	1.19 (<.0001)	-	-
Flanker arrow	0.64 (<.0001)	0.56 (<.0001)	1.58 (<.0001)
Flanker letters	0.57 (<.0001)	0.46 (<.0001)	0.79 (<.0001)
Flanker color	0.25 (.006)	0.48 (<.0001)	1.05 (<.0001)
Stroop color/word	0.23 (.01)	0.94 (<.0001)	1.13(<.0001)
Stroop quant/number	.90 (<.0001)	1.37 (<.0001)	1.72 (<.0001)
Stroop size/word	.78 (.0001)	1.15 (<.0001)	1.88 (<.0001)

Table 2 Mean, standard deviation, and internal consistency for accuracy and mean response times for correct responses for all tasks. For accuracy, internal consistency is Cronbach alpha for Fluid IQ measures and Guttman lambda2 for congruency measures. For response times, internal consistency is Spearman-Brown corrected split-half reliability (split over even and odd trials). If reliability was negative, 0 is reported

		Accuracy				Mean correct response times			
		Mean (SD)	Skewness	Kurtosis	Reliability*	Mean (SD)	Skewness	Kurtosis	Reliability**
VHPT-Faces	Congruent	.70 (.10)	-.93	1.45	0.8				
	Incongruent	.58 (.10)	-.56	-.13	0.76				
	Congruency	.12 (.10)	.04	.61	0.62				
Flanker arrow	Congruent	.99 (.05)	-7.46	61.98	0.95	510 (175)	4.12	20.69	0.82
	Incongruent	.95 (.07)	-2.37	7.70	0.68	613 (273)	5.24	34.91	0.60
	Congruency	.04 (.06)	1.03	1.99	0.53	104 (192)	7.89	78.80	0
Flanker color	Congruent	.95 (.06)	-1.42	2.31	0.28	495 (176)	5.61	41.21	0.77
	Incongruent	.94 (.06)	-1.34	2.01	0.24	534 (216)	5.18	32.18	0.69
	Congruency	.01 (.05)	-.13	.96	0	40 (84)	2.63	15.19	0.11
Flanker letter	Congruent	.98 (.03)	-2.80	9.39	0.24	490 (152)	5.12	37.14	0.77
	Incongruent	.95 (.06)	-1.73	3.97	0.67	510 (115)	4.44	30.52	0.57
	Congruency	.04 (.07)	.49	2.90	0.52	21 (56)	1.86	11.74	0
Stroop color/word	Congruent	.97 (.03)	-1.39	1.41	0.24	710 (160)	2.88	17.74	0.95
	Incongruent	.95 (.10)	-2.05	5.03	0.86	789 (155)	1.72	5.99	0.77
	Congruency	.02 (.09)	.46	1.39	0.72	79 (74)	-0.15	3.21	0
Stroop quant/numb	Congruent	.99 (.02)	-2.22	5.21	0.19	614 (173)	6.1	45.07	0.97
	Incongruent	.93 (.07)	-1.41	2.17	0.49	706 (182)	4.1	25.42	0.57
	Congruency	.06 (.07)	1.27	2.16	0.39	92 (66)	1.69	6.61	0
Stroop size/word	Congruent	.99 (.03)	-2.86	10.16	0.36	602 (164)	4.29	24.9	0.52
	Incongruent	.94 (.07)	-1.44	2.64	0.56	707 (199)	4.91	37.79	0.84
	Congruency	.05 (.06)	1.33	2.08	0.41	105 (83)	1.54	6.40	0
Fluid IQ	Letters	18.12 (4.40)	-0.39	-0.10	0.85				
	Number Sets	10.24 (2.40)	-0.09	-0.50	0.67				
	Ravens	11.45 (2.69)	-0.10	-0.48	0.64				
	COMBINED	39.87 (7.08)	-0.61	0.14	0.85				

often comparable to that of accuracy scores (Table 3). The RCS congruency scores for the Flanker/color congruency task were deemed too unreliable to use in further analyses.

In Table 4 we present zero-order correlations between congruency effects (for faces and cognitive control tasks) and Fluid IQ, using both the accuracy and RCS dependent variables for the cognitive control tasks. We also include their value disattenuated for measurement error (Nunnally, 1970). Disattenuated correlations can overcorrect and be imprecise (i.e., have large confidence intervals; Wetcher-Hendricks, 2006), sometimes exceeding 1 when reliability is low.

They are provided to give a sense for what the true effect size would have been without measurement error. As expected based on prior work, the Flanker and Stroop tasks do not show a high degree of common variance, but out of ten pairwise correlations, seven of them were significant for accuracy, and five used RCS. The three Stroop tasks in particular are all correlated with each other, with both dependent variables. In

contrast, the congruency effect for faces showed no evidence of sharing variance with any of the cognitive control tasks (see Fig. 2).

The VHPT-F congruency effect was also not related to Fluid IQ, which itself showed only one significant correlations with the cognitive control measures: using RCS only, the Flanker arrow task was significantly related to Fluid IQ.

For completeness, we provided the zero-order correlations among all individual conditions in the congruency tasks, using both accuracy (Table 5) and RCS (Table 6). In this case, Spearman rhos are reported because most of the cognitive control single conditions were highly skewed (although the RCS scores were more normally distributed). It is important to note that, unlike congruency scores, individual conditions in congruency tasks do not reflect the constructs of interest that the tasks were designed to measure (e.g., holistic processing or cognitive control). These are presented for exploratory purposes, and while significance is presented at the .05

Table 3 Rate Correct Score (RCS) mean, measures of the shape of RCS distribution and reliability (Spearman-Brown split half). Under normality, the expected values of measures of skewness and kurtosis are 0 and 3, respectively

Task	Mean (SD)	Skewness	Kurtosis	Reliability
Flanker arrow congruency	.38 (.24)	-0.60	1.11	.58
Flanker color congruency	.15 (.19)	-0.05	-0.51	.08
Flanker letter congruency	.23 (.22)	0.97	2.23	.28
Stroop color/word congruency	.17 (.15)	0.99	2.41	.35
Stroop quantity/number congruency	.31 (.18)	0.43	1.28	.53
Stroop size/word congruency	.32 (.17)	0.30	0.35	.37

(uncorrected level), we invite readers to use this threshold as a rule of thumb and inspect the pattern of effect sizes. In accuracy, there is some indication that ceiling effects may limit correlations: this is suggested both by the very low correlations between Fluid IQ and most other tasks and the relatively low correlations in several of the paired conditions that come from the same task (e.g., Flanker arrow congruent and incongruent trials, r=.08). Of most interest here is how the VHPT-F conditions relate to other tasks. The VHPT-F incongruent condition showed more significant correlations with other conditions (Flanker conditions) than the VHPT-F congruent condition. However, in all cases, the VHPT-F scores were correlated with both the congruent and incongruent Flanker conditions.

The zero-order correlations among conditions using the RCS scores are overall much higher, likely both because they are not restricted in range and also in part reflect speed, known to correlate with domain-general cognitive skills (Sheppard & Vernon, 2008). Accordingly, with RCS, all conditions in the cognitive control tasks were significantly related to Fluid IQ. In addition, as might be expected because of task-specific and stimulus-specific variance, every single correlation that comes

from the paired (congruent and incongruent) conditions in the same task was higher (.7 or higher) than the highest correlation among unpaired conditions (.64). Again, of most interest here is how the VHPT-F conditions relate to other tasks, and it is striking in the context of the strong correlations among conditions with RCS that the VHPT-F incongruent condition was almost never correlated with RCS in other tasks (the exception is the Stroop Size/Word task). This does not appear to stem from the format of the task (three-alternative forced choice in the VHPT-F) or the fact that the VHPT-F scores use accuracy because this is also true of the VHPT-F congruent scores, which show more robust correlations with RCS (seven out of 12 correlations significant).

Discussion

We explored the relation between holistic processing of faces and measures of cognitive control that are also operationalized using congruency effects. Cognitive control refers to how people come up with, monitor, and adjust the strategies

Table 4 Zero-order correlations (and correlations corrected for measurement error) across congruency effects and Fluid IQ, using accuracy (top) or Rate Correct Score (RCS) (bottom) for the cognitive control tasks.

|	1	2	3	4	5	6
1.VHPT face congruency						
2. Flanker arrow congruency acc	.08 (.13)					
3. Flanker letter congruency acc	-.07 (-.12)	**.20 (.38)**				
4. Stroop color/word congruency acc	-.04 (-.06)	-.05 (-.07)	**.25 (.41)**			
5. Stroop quant/num congruency acc	-.05 (-.09)	**.18 (.40)**	.05 (.12)	**.19 (.36)**		
6. Stroop size/word congruency acc	.02 (.04)	.16 (.34)	**.26 (.56)**	**.23 (.42)**	**.21 (.53)**	
7. Fluid IQ	-.04 (-.05)	.08 (.12)	.01 (.01)	-.12 (-.15)	.03 (.05)	-.17 (-.29)
	1	2	3	4	5	6
1.VHPT face congruency						
2. Flanker arrow congruency RCS	-.06 (-.10)					
3. Flanker letter congruency RCS	.03 (.08)	.11 (.27)				
4. Stroop color/word congruency RCS	.03 (.07)	.04 (.08)	**.27 (.86)**			
5. Stroop quant/num congruency RCS	.03 (.05)	.11 (.19)	.09 (.23)	**.38 (.89)**		
6. Stroop size/word congruency RCS	.13 (.27)	-.10 (-.22)	**.29 (.91)**	**.38 (1.1)**	**.35 (.79)**	
7. Fluid IQ	-.04 (-.05)	**.25 (.36)**	.06 (.12)	-.02 (-.03)	.03 (.05)	-.10 (-.17)
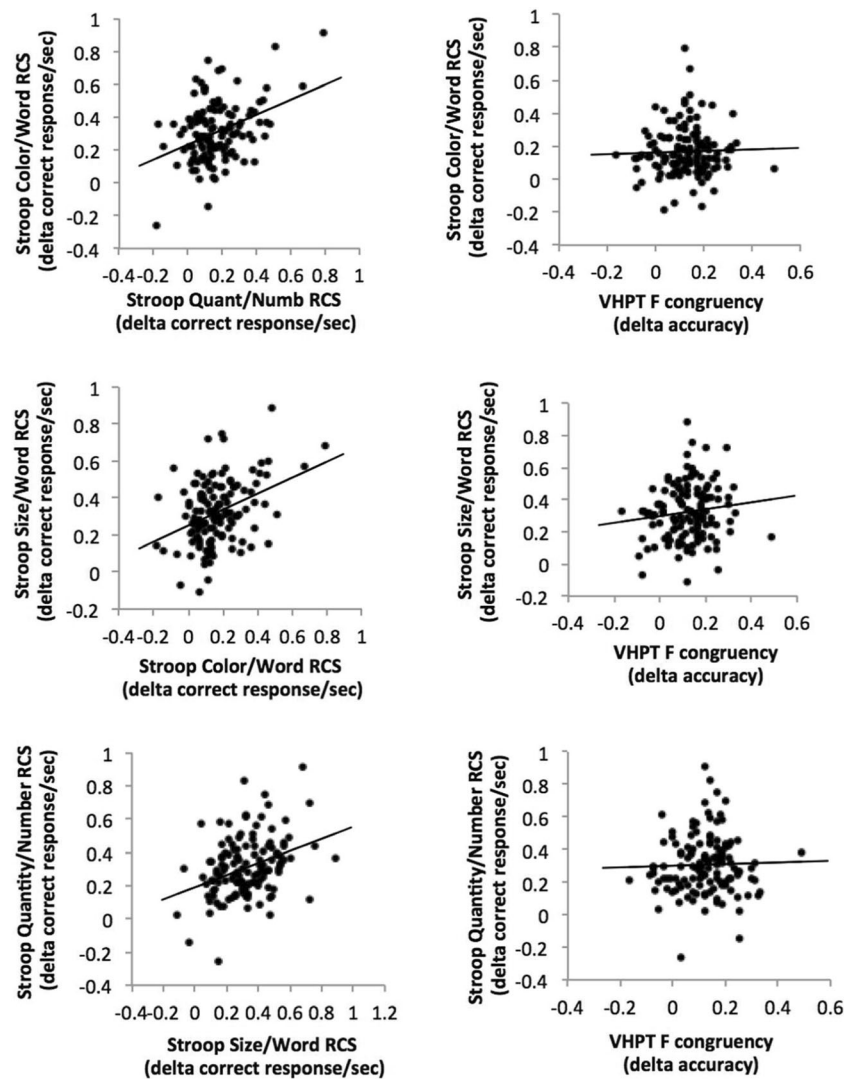
Fig. 2 On the left, scatterplots relating the RCS congruency effects in the three Stroop tasks. On the right, scatterplots relating holistic processing for faces (congruency in the VHPT-F) with RCS congruency effects in each Stroop task

required to stay on task, including the ability to suppress unwanted responses. In particular, Stroop tasks (Stroop, 1935) and Flanker tasks (Eriksen & Schultz, 1979) tap into the ability to inhibit responses triggered by task irrelevant information. On the surface, this is similar to the selective attention challenges present in variations of the composite task, including the VHPT-F (Richler et al., 2014), in which the task-relevant part is clearly indicated on each trial and interference comes from responses associated with the to-be-ignored part. This has led some to suggest that holistic processing as indexed by the congruency effect taps into the same mechanisms than Stroop or Flanker tasks (Rossion, 2013).

There are several challenges to testing this hypothesis. One is that, like most measures consisting of difference scores, congruency effects in Stroop and Flanker tasks often have limited reliability (e.g., Kindt et al., 1996; Strauss, Allen, Jorgensen, & Cramer, 2005). When they do produce reliable effects, the same

paradigm (Stroop or Flanker) with different material (e.g., words, colors, arrows) does not always produce strong correlations (Salthouse & Meinz, 1995; Shilling, Chetwynd, & Rabbitt, 2002; Yehene & Meiran, 2007; Ward, Roberts, & Phillips, 2001).

Second, despite evidence that the response conflict in both the Flanker and Stroop paradigms engage similar neural substrates, including the anterior cingulate and prefrontal cortex (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Fan, Flombaum, McCandliss, Thomas, & Posner, 2003; Ridderinkhof et al., 2002), there are also differences with the Flanker task also engaging the right dorsolateral prefrontal cortex and right insula and the Stroop task also recruiting the left frontal cortex (Nee, Wager, & Jonides, 2007). Consistent with some differences in the neural mechanisms involved, individual differences across the two paradigms often fail to correlate (e.g., Bender, Filmer, Garner, Naughtin, & Dux, 2016; Stins, Polderman, Boomsma, & de Geus, 2005).

Table 5 Spearman Rho zero-order correlations among individual conditions. Accuracy is used for all variables. Values in bold are significant at an alpha of .05, uncorrected. Table is organized to facilitate comparisons of similar relations, for instance those among different Flanker conditions or different Stroop conditions. Values in boxes are the correlations among the congruent and incongruent conditions for the same task

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.VHPT-F C														
2.VHPT-F I	**0.34**													
3. Flanker arrow C	**0.19**	**0.23**												
4. Flanker arrow I	0.17	**0.23**	0.08											
5. Flanker color C	0.13	**0.19**	**0.18**	0.12										
6. Flanker color I	0.14	**0.32**	**0.20**	**0.23**	**0.46**									
7. Flanker letter C	0.08	0.10	0.15	0.08	**0.39**	**0.30**								
8. Flanker letter I	0.16	0.11	0.09	**0.29**	0.12	0.13	0.06							
9. Stroop c/w C	-0.05	0.02	-0.10	**0.19**	-0.15	0.03	0.05	**0.20**						
10. Stroop c/w I	0.08	0.04	-0.03	0.17	0.14	0.00	0.03	**0.38**	**0.27**					
11. Stroop q/n C	0.07	0.01	0.08	**0.21**	0.13	0.11	0.02	**0.24**	**0.24**	**0.24**				
12. Stroop q/n I	0.13	0.02	0.11	**0.19**	**0.19**	**0.29**	0.11	0.17	**0.22**	**0.21**	**0.18**			
13. Stroop s/w C	0.03	0.04	-0.02	0.14	0.11	0.15	0.09	0.15	0.15	0.15	0.12	**0.18**		
14. Stroop s/w I	**0.25**	**0.26**	0.12	**0.29**	**0.27**	**0.30**	0.17	**0.31**	**0.19**	**0.30**	0.12	**0.21**	**0.28**	
15. Fluid IQ	0.16	**0.23**	0.13	0.07	0.03	0.11	**0.19**	0.10	-0.04	0.13	0.06	0.03	0.04	**0.19**

Third, there is evidence that interference effects can be strongly dependent on past experience with specific stimulus dimensions (Ward et al., 2001) and that they may be mediated by independent domain-specific mechanisms (Egner, 2008).

Accordingly, any effort to consider the shared variance between holistic processing of faces and cognitive control would be difficult to interpret had we relied on a single task to index non-face cognitive control. Here, we chose to use six different tasks, three Flanker and three Stroop tasks, expecting based on the literature that there should be moderate correlations between interference in these tasks but providing a context in which to ask whether holistic processing of faces appears to behave just like another cognitive control task. In other words, any claim that the selective attention effect measured in the VHPT-F is distinct from other measures of selective attention must be evaluated against evidence that these other measures may not necessarily cohere.

Five of our six Stroop and Flanker tasks provided minimally reliable congruency effects in accuracy and in RCS, and correlational analyses offered indications of non-negligible shared variance among these effects, especially among the three Stroop tasks and the Flanker letter task. However, there was no evidence, using either accuracy or RCS, of congruency effects in these cognitive control tasks correlating with holistic processing of faces.

We had no prediction as to whether the VHPT-F task would be more similar to the Stroop or the Flanker paradigm. In Stroop tasks, the irrelevant information is another dimension of the same stimulus, whereas in Flanker tasks, the irrelevant information is spatially separated from the target. When considering face parts, interference in the VHPT-F is more like Flanker interference (parts are separated in space), although face processing being considered holistic, it may be possible to argue that as in the Stroop, another dimension of the same object (the whole face) needs to be ignored. Zero-order correlations of mean accuracy in the individual conditions grouped performance in the VHPT-F task with accuracy in Flanker tasks more than in Stroop tasks. This is consistent with the interpretation that face part judgments are based on representations in which parts are explicitly represented, and with the suggestion that holistic processing depends on an attentional mechanism that can integrate spatially separated face parts (Chua, Richler, & Gauthier, 2014, Chua et al., 2015; Richler et al., Individual differences in object recognition

Table 6 Spearman rho zero-order correlations among individual conditions. Accuracy is used for VHPT-F and Fluid IQ, Rate Correct Score (RCS) for all other tasks. Values in bold are significant at an alpha of .05, uncorrected. Table is organized to facilitate comparisons of similar relations, for instance those among different Flanker conditions or different Stroop conditions. Values in boxes are the correlations among the congruent and incongruent conditions for the same task

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. VHPT-F C														
2. VHPT-F I	**0.34**													
3. Flanker arrow C	0.17	0.07												
4. Flanker arrow I	**0.24**	0.12	**0.72**											
5. Flanker color C	0.07	-0.01	**0.53**	**0.44**										
6. Flanker color I	**0.20**	0.08	**0.61**	**0.54**	**0.83**									
7. Flanker letter C	0.13	0.10	**0.61**	**0.52**	**0.55**	**0.56**								
8. Flanker letter I	**0.23**	0.13	**0.51**	**0.52**	**0.49**	**0.53**	**0.70**							
9. Stroop c/w C	**0.18**	0.07	**0.40**	**0.31**	**0.47**	**0.51**	**0.43**	**0.31**						
10. Stroop c/w I	**0.21**	0.10	**0.36**	**0.26**	**0.43**	**0.53**	**0.33**	**0.30**	**0.82**					
11. Stroop q/n C	0.16	0.04	**0.55**	**0.44**	**0.54**	**0.60**	**0.51**	**0.41**	**0.50**	**0.46**				
12. Stroop q/n I	**0.22**	0.10	**0.44**	**0.32**	**0.45**	**0.49**	**0.45**	**0.37**	**0.33**	**0.42**	**0.70**			
13. Stroop s/w C	0.17	0.10	**0.54**	**0.41**	**0.50**	**0.51**	**0.42**	**0.28**	**0.52**	**0.44**	**0.64**	**0.48**		
14. Stroop s/w I	**0.25**	**0.26**	**0.52**	**0.35**	**0.43**	**0.50**	**0.32**	**0.27**	**0.42**	**0.47**	**0.56**	**0.54**	**0.75**	
15. Fluid IQ	0.16	**0.23**	**0.30**	**0.31**	**0.27**	**0.45**	**0.35**	**0.38**	**0.23**	**0.28**	**0.25**	**0.29**	**0.23**	**0.31**

(submitted)), rather than on undifferentiated holistic representations (Farah et al., 1998; Tanaka & Farah, 1993). However, we note that RCS, which is arguably a more sensitive measure than accuracy, with better distributional properties, related the congruent VHPT-F scores to both Stroop and Flanker tasks equally.

There are limitations to our study. They include the fact that congruency effects in response times in cognitive control tasks were not reliable (despite significant congruency effects in RTs and reliable individual conditions) because they were strongly correlated across conditions. In addition, response times on the VHPT-F were not examined because the task is not speeded and the 3-AFC format can complicate the interpretation of speed. In addition, errors were low in the cognitive control tasks, as is often the case (e.g., Hedge et al., 2017). While difference scores in accuracy yielded significant congruency effects and medium effect sizes in most conditions, the ceiling on accuracy still limited correlations with other tasks. While we used the recently advocated (Vandierendonck, 2017) combined RCS measure to address these limitations in response times and accuracy, we acknowledge that there could be more powerful and meaningful means of combining speed and accuracy on these tasks. Future work could develop a reliable, 2-AFC speeded version of the VHPT-F, and provided versions of cognitive control tasks are

adjusted to increase error rates, response times and errors could be analyzed using a process model such as the diffusion model (Ratcliff & McKoon, 2008). This approach would provide a way to relate parameters that are more directly related to actual cognitive processes than dependent variables.

Our conclusions are also limited by the population we studied, which is a young adult population of individuals both students and from the community around the university. There is evidence that congruency effects in the composite task increase with presentation duration and are less specific to faces in older adults than in young adults (Meinhardt, Persike, & Meinhardt-Injac, 2016). It is therefore possible that the contribution of cognitive control to the VHPT-F increases in a population with reduced control ability.

Another limitation is that we found limited evidence of correlations between fluid IQ and congruency effects in our cognitive control tasks. This appears surprising given that working-memory capacity, which itself has a strong relation to fluid intelligence (Kane, Hambrick, & Conway, 2005) has been related to performance on selective attention tasks like the Stroop task (Unsworth & Spillers, 2010). Interestingly, the relation between working memory capacity and congruency effects depend on proportion of congruent trials (Kane & Engle, 2003; Hutchison, 2011), which was set to 50% here

because this is what is used in the VHPT-f. Therefore, we would not exclude the possibility that if the proportion of congruent trials was increased in all our selective attention tasks including the VHPT-F, a relation with fluid IQ might arise. Multiple levels of control are thought to support performance in selective attention tasks (Bugg & Crump, 2012).

Consequently, there is one clear sense in which our results are inconsistent with Rossion's (2013) suggestion that measuring holistic face processing as a congruency effect taps the same control mechanisms measured in the Stroop or Flanker tasks. That is, the present results suggest that to the extent that we can capture common control mechanisms in our various cognitive control tasks, they do not appear to be shared with the VHPT-F. However, given that endogenous control in complex tasks can rely on a number of mechanisms that operate at different levels and can be sensitive to a variety of task manipulations, it seems prudent to acknowledge that some other variation of these tasks could show more overlap with holistic effects in the VHPT-F. Because cognitive control measures can be implemented with an infinity of stimuli and using a variety of parameters (e.g., the proportion of congruent to incongruent trials), the answer to the question we set out to ask here may benefit from a meta-analytic approach, once many studies have included cognitive control and holistic processing of faces. We believe that the use of the VHPT-F as standard measure for holistic face processing would facilitate convergence from future studies.

To conclude, previous work has shown that holistic processing of faces as measured by the VHPT-F is stable (Richler et al., 2014). Here, we find no evidence that holistic processing of faces is accounted for by domain-general control mechanisms. Our results are consistent with other work that suggest failures of selective attention primarily reflect domain-specific mechanisms (Egner, 2008), although we found some evidence of common variance across different cognitive control tasks.

Training studies with non-face objects suggests that holistic processing is at least in part driven by the level of experience one has individuating objects from a category (Chua et al., 2014; Chua, 2017). However, the lack of correlation between the VHPT-F and face recognition ability as measured by a face learning task such as the Cambridge Face Memory Test (Duchaine & Nakayama, 2006) seems inconsistent with a role of experience in driving holistic processing of faces. But experience may be multifaceted, and is more difficult to measure than to manipulate. Recent work has shown that manipulating experience with a new category of faces can reveal brain-behavior correlations that are much more difficult to obtain using standard face recognition measures (McGugin, Ryan, Tamber-Rosenau, & Gauthier, 2017). Therefore, a combination of an individual differences approach and the manipulation of experience in a training paradigm might help uncover the sources of individual differences in holistic face processing.

References

Bender, A. D., Filmer, H. L., Garner, K. G., Naughtin, C. K., & Dux, P. E. (2016). On the relationship between response selection and response inhibition: An individual differences approach. *Attention, Perception, & Psychophysics*, *78*(8), 2420–2432.

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.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*(6758), 179–181.

Bugg, J. M., & Crump, M. J. (2012). In support of a distinction between voluntary and stimulus-driven control: a review of the literature on proportion congruent effects. *Frontiers in Psychology, 3*.

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. (2017). *Holistic processing: A matter of experience* (Doctoral dissertation, Vanderbilt University).

Chua, K. W., Richler, J. J., & Gauthier, I. (2014). Becoming a Lunari or Taiyo expert: Learned attention to parts drives holistic processing of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1174–1182.

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.

Conway, A. R., Cowan, N., Bunting, M. F., Therriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*(2), 163–183.

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.

Duchaine, B., & Nakayama, K. (2006). The Cambridge face memory test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*(4), 576–585.

Egner, T. (2008). Multiple conflict-driven control mechanisms in the human brain. *Trends in cognitive sciences*, *12*(10), 374–380.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. *Princeton, NJ: Educational testing service*.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309–331.

Eriksen, B.A., & Eriksen, C.W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143–149.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Attention, Perception, & Psychophysics*, *25*(4), 249–263.

Fan, J., Flombaum, J. I., McCandliss, B. D., Thomas, K. M., & Posner, M. I. (2003). Cognitive and brain consequences of conflict. *Neuroimage*, *18*(1), 42–57.

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. (2002). Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 431–446.

Hedge, C., Powell, G., & Sumner, P. (2017). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 1–21. https://doi.org/10.3758/s13428-017-0935-1.

Hutchison, K. A. (2011). The interactive effects of listwide control, item based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 851–860.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*(1), 66–71.

Kindt, M., Bierman, D., & Brosschot, J. F. (1996). Stroop versus Stroop: Comparison of a card format and a single-trial format of the standard color-word Stroop task and the emotional Stroop task. *Personality and Individual Differences*, *21*(5), 653–661.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological bulletin*, *109*(2), 163–203.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in cognitive sciences*, *6*(6), 255–260.

McGugin, R. W., Ryan, K. F., Tamber-Rosenau, B. J., & Gauthier, I. (2017). The role of experience in the face-selective response in right FFA. *Cerebral Cortex*, 1–14. https://doi.org/10.1093/cercor/bhx113.

Meinhardt, G., Persike, M. & Meinhardt-Injac, B. (2016). The composite effect is face-specific in young but not older adults. *Front. Aging. Neurosci.*, 8:187.

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2014). Holistic processing and reliance on global viewing strategies in older adults' face perception. *Acta psychologica*, *151*, 155–163.

Melara, R. D., & Mounts, J. R. (1993). Selective attention to Stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition*, *21*(5), 627–645.

Melnick, M. D., Harrison, B. R., Park, S., Bennetto, L., & Tadin, D. (2013). A strong interactive link between sensory discriminations and intelligence. *Current Biology*, *23*(11), 1013–1017.

Nee, D. E., Wager, T. D., & Jonides, J. (2007). Interference resolution: Insights from a meta-analysis of neuroimaging tasks. *Cognitive, Affective, & Behavioral Neuroscience*, *7*(1), 1–17.

Nunnally Jr, J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill.

Peschke, C., Hilgetag, C. C., & Olk, B. (2013). Influence of stimulus type on effects of flanker, flanker position, and trial sequence in a saccadic eye movement task. *Quarterly Journal of Experimental Psychology, 66*(11), 2253–2267.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Raven, J., Raven, J. C., Court, JH (1998). Manual for Raven's progressive matrices and vocabulary scales. Section 5: The Mill Hill vocabulary scale. Oxford, England: Oxford Psychologists Press

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., … Engle, R. W. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142*(2), 359–379.

Richler, J.J., Cheung, O.S., Wong, A. C.-N., (2009). Does response interference contribute to face composite effects? *Psychonomic Bulletin and Review*, 16(2): 258–63.

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., Mack, M. L., Palmeri, T. J., & Gauthier, I. (2011). Inverted faces are (eventually) processed holistically. *Vision Research*, *51*(3), 333–342.

Richler, J. J., Palmeri, T. J., & Gauthier, I. (2012). Meanings, mechanisms, and measures of holistic processing. *Frontiers in Psychology, 3*, 553.

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

Ridderinkhof, K. R., de Vlugt, Y., Bramlage, A., Spaan, M., Elton, M., Snel, J., & Band, G. P. (2002). Alcohol consumption impairs detection of performance errors in mediofrontal cortex. *Science*, *298*(5601), 2209–2211.

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.

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *50*(6), 297–306.

Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences*, *112*(41), 12887–12892.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*(3), 535–551.

Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia*, *40*(6), 605–619.

Stins, J. F., Polderman, J. C., Boomsma, D. I., & de Geus, E. J. (2005). Response interference and working memory in 12-year-old children. *Child Neuropsychology*, *11*(2), 191–201.

Strauss, G. P., Allen, D. N., Jorgensen, M. L., & Cramer, S. L. (2005). Test-retest reliability of standard and emotional stroop tasks: An investigation of color-word and picture-word versions. *Assessment*, *12*(3), 330–337.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Tadin, D. (2015). Suppressive mechanisms in visual motion processing: From perception to intelligence. *Vision Research*, *115*, 58–70.

Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *The Quarterly Journal of Experimental Psychology*, *46*(2), 225–245.

Thurstone, L.L. (1938). Primary mental abilities. University of Chicago Press: Chicago.

Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory and Language*, *62*(4), 392–406.

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior research methods, 49*(2), 653–673.

Van Gulick, A. E., McGugin, R. W., & Gauthier, I. (2016). Measuring nonvisual knowledge about object categories: The semantic vanderbilt expertise test. *Behavior Research Methods*, *48*(3), 1178–1196.

Wang, C-C, Ross DA, Gauthier, I. & Richler J.J. (2016) Validation of the Vanderbilt Holistic Face Processing Test. *Frontiers in Psychology* 7: 1837.

Ward, G., Roberts, M. J., & Phillips, L. H. (2001). Task-switching costs, Stroop-costs, and executive control: A correlational study. *The Quarterly Journal of Experimental Psychology: Section A*, *54*(2), 491–511.

Wetcher-Hendricks, D. (2006). Adjustments to the correction for attenuation. *Psychological methods*, *11*(2), 207–215.

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., & Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of sciences*, *107*(11), 5238–5241.

Woltz, D. J., & Was, C. A. (2006). Availability of related long-term memory during and after attention focus in working memory. *Memory and Cognition, 34*(3), 668–684.

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.

Wong, A. C. N., Palmeri, T. J., Rogers, B. P., Gore, J. C., & Gauthier, I. (2009). Beyond shape: How you learn about objects affects how they are represented in visual cortex. *PloS one*, *4*(12), e8405.

Yehene, E., & Meiran, N. (2007). Is there a general task switching ability? *Acta Psychologica*, *126*(3), 169–195.