

Car expertise does not compete with face expertise during ensemble coding

Jisoo Sun, Isabel Gauthier

Department of Psychology, Vanderbilt University

This work was supported by NSF (SMA1640681) and by the David K. Wilson Chair Research Fund (Vanderbilt University). We thank Randolph Blake, Ashleigh Maxcey and Oakyoon Cha for comments on the manuscript and Melanie Kacin, Marcus Bennett and Anjali Mahapatra for help with data collection. This experiment was not preregistered, but the data are publicly accessible (<https://doi.org/10.6084/m9.figshare.12126453.v1>) and the materials are available upon request.

Correspondence concerning this article should be addressed to Jisoo Sun, Department of Psychology, Vanderbilt University, 111 21st Avenue South, Nashville, TN 37240.

E-mail: jisoo.sun@vanderbilt.edu

Main text: 3164 words

ABSTRACT

When objects from two categories of expertise (e.g., faces and cars in dual car/face experts) are processed simultaneously, competition occurs across a variety of tasks. Here, we investigate whether competition between face and car processing also occurs during ensemble coding. The relationship between single object recognition and ensemble coding is debated, but if ensemble coding relies on the same ability as object recognition, we expect cars to interfere with ensemble coding of faces as a function of car expertise. We measured the ability to judge the variability in identity of arrays of faces, in the presence of task-irrelevant distractors (cars or novel objects). On each trial, participants viewed two sequential arrays containing 4 faces and 4 distractors, judging which array was the more diverse in terms of face identity. We measured participants' car expertise, object recognition ability, and face recognition ability. Using Bayesian statistics, we found evidence against competition as a function of car expertise during ensemble coding of faces. Face recognition ability predicted ensemble judgments for faces, regardless of the category of task-irrelevant distractors. The result suggests that ensemble coding is not susceptible to competition between different domains of similar expertise, unlike single object recognition.

Keywords: Object perception, Individual differences, Competition

Our experience with objects varies depending on our environment, interests or occupation. Expertise recognizing individual objects of a category can lead to holistic processing, evidenced by difficulties processing one part of an object while ignoring other parts (Chua & Gauthier, 2020). When a task uses objects from two categories that are processed holistically, competition can occur. Competition as a function of expertise occurs in a variety of tasks, including visual search (McGugin et al., 2011), spatial judgments (Rossion et al., 2004) and discrimination under rapid presentation (McKeeff et al., 2010). This competition is attributed to overlapping neural representations for faces and other objects of expertise (Gauthier et al., 2003; McGugin et al., 2015; Rossion et al., 2007; Rossion et al., 2004). Competition occurs even when one of the two categories is task-irrelevant (e.g., McGugin et al., 2011).

The variety of tasks for which such effects are observed suggests that competition is determined by the overlap of perceptual representations, regardless of the decision required. In the cerebral functional distance framework (Kinsbourne & Hicks, 1978), processing spreads within a highly linked network. When processing objects of expertise interferes with processing faces, the representations are inferred to overlap in functional (and cerebral) space. Interestingly, all tasks in which this competition was observed require attention to individual objects or faces – but what of other tasks that do not require individuating objects? Objects often appear with regularity and redundancy, such as the books on a bookshelf or a crowd of people. People can process multiple objects at once to overcome limitations in capacity (Cohen et al., 2016; Whitney & Yamanashi Leib, 2018). Such “ensemble coding” reduces multiple items to summary statistics (e.g., mean, variance) and has been documented for low-level visual stimuli and complex objects like faces (Chong & Treisman, 2003; Haberman & Whitney, 2007; Haberman et al., 2015b; Michael et al., 2014). Many studies reveal that people extract summary statistics rapidly and

relatively accurately, in implicit judgments that are not explained by other strategic short cuts dependent on attention to single objects (Ariely, 2001; Cha & Chong, 2018; Haberman & Whitney, 2007, 2009; Phillips et al., 2018). However, the mechanisms that support these judgments, are not well understood, especially for complex objects.

Evidence of competition points to overlapping representations for recognition of faces and objects of expertise, but it is unclear whether this should be generalized to ensemble coding. Some suggest that ensemble coding depends on the same mechanisms as single object perception (Neumann et al., 2018). For instance, performance recognizing single objects is correlated with recognizing the average identity of several objects (Haberman et al., 2015a). However, differences in brain activity during tasks that require single object processing and tasks that engage ensemble coding suggest different mechanisms (Cant & Xu, 2012; Im et al., 2017).

Accordingly, faces and cars may compete when dual face and car experts form ensemble representations with these categories, just as in tasks requiring attention to object identity. However, inferences based on experiments on simple features may not necessarily extend to face processing. For example, ensemble coding of facial expression can differ in qualitative ways from the processing of individual facial expression (Sun & Chong, 2020). Ensemble coding of faces may not follow the same rules as individual face recognition. This could be reflected in the absence of competition with non-face objects of expertise.

If ensemble coding relies on the same mechanisms as object recognition, we expect to see competition between cars and faces, as a function of car expertise, during ensemble coding task (although it is possible that ensemble coding relies on mechanisms distinct from object recognition, but similar effects of expertise on the functional distance between car and face representations). However, if we find Bayesian evidence against competition as a function of car

expertise during ensemble coding, this would suggest that ensemble representations, for faces or cars (or for both categories), are considerably different from those involved in individual object recognition tasks.

Because previous work that found competition between domains of expertise used natural images of faces and cars rather than morphed images (often used in ensemble coding), we chose to do the same here. This also avoids strange car images that do not exist as real models. We chose a diversity judgment task from prior ensemble coding studies (Cha et al., in press; Phillips et al., 2018), allowing us to sidestep the need for complicated assumptions about averaging faces that vary on a large number of dimensions. The ability to make such diversity judgments is correlated with judgments about the average of a set (Cha et al., submitted). We seek to test whether competition occurs as a function of car expertise during ensemble coding, rather than investigate the exact strategy used during diversity judgments. In the extreme, the use of sequential attentional strategy would increase the similarity to the task where competition has been observed, which is not supported by the results of this study.

Participants judged the diversity of the identity of arrays of faces while ignoring the presence of distractor cars (or control novel objects). We measured perceptual expertise for cars and faces and controlled for domain-general object recognition ability using novel objects.

Methods

Participants

We recruited 58 individuals to participate in the experiment and used data from 53 individuals (21 males, 32 females; mean age 19.66 years old, $SD = 2.64$) in the analyses because one person's data were not saved and four participants did not complete all tasks. A sample from

the campus population is expected to show considerable variability in car expertise (McGugin et al., 2012b). A power analysis (using G*Power 3.1, Faul et al., 2009) based on a study reporting competition between faces and cars as a function of car expertise, while cars were task-irrelevant ($r = .41$; McGugin et al., 2011) suggested that 41 participants were required to achieve 80% power at an alpha of .05. Participants received either course credit or money. They had normal or corrected-to-normal visual acuity and provided informed consent. All work was conducted under the approval of the Vanderbilt Institutional Review Board.

Face diversity judgment task (Ensemble task)

Stimuli

The experiment was conducted in MATLAB with Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). We used images of 474 faces, 296 cars, and 320 novel objects. The number of exemplars was meant to be large but it is not matched across categories, because that is not relevant to our predictions about individual differences. What matters are identical conditions across all participants.

We selected images of adult faces varying in gender, race and age but without facial hair from the Chicago Face Database (Ma et al., 2015). To minimize low-level differences, we eliminated the outside parts of the contour (Zhu & Ramanan, 2012). We converted the images into grayscale and matched their median luminance to a common value.

Car images were collected from the Motor Trend website. We selected two subsets of cars from five categories (convertible, coupe, hatchback, sedan and SUV/crossover), either roughly black or white (once changed to grayscale from their original color), and selected cars of only one set on any given trial. All images showed a front $\frac{3}{4}$ view of a car and transparent

background. The images were transformed to grayscale and resulted in a set of 147 white cars and 149 black cars.

We used novel objects called Ziggerins (Figure 1; Wong et al., 2009) as a control category that would not be familiar to participants. There were 80 different identities, each in two viewpoints. This resulted in 160 images, from which we created two sets of images – 160 in darker colors and 160 in lighter colors, to match the two sets of cars.

Design and Procedure

Figure 1 shows sample arrays for each distractor condition. Participants fixated on a central spot at the start of each trial, pressed the space bar and viewed two arrays in succession, each for 700ms. Each array contained four faces and four distractors (cars or Ziggerins, depending on the block) presented along an imaginary circle centered on fixation (of radius approximately 5.25° in visual angle). The exact eccentricities of each object were spatially jittered (between 5.01° and 5.72° of visual angle). Participants chose which array contained the more diverse set of faces, regardless of the distractors (“f” key: first array and “j” key: second array). They were told that a display containing four different people is more diverse than a display that shows some repeated faces and were shown examples. Sixteen practice trials preceded the first block (with car distractors) and the second block (with Ziggerin distractors). Visual feedback (‘correct’ or ‘incorrect’) was given only for practice trials. A 30 sec break followed each block of 96 trials for a total of 384 experimental trials.

On each trial, one of two arrays contained four different faces (face-diverse array, left column in Figure 1) and the other array contained images of the faces of two people (face-repeated array, right column in Figure 1). Half of these face-repeated arrays had 2 overlapping

faces (e.g., A, A, B, and B) and the other half had two faces, with one face repeated three times (e.g., A, A, A, and B). Distractor type was blocked and trials were presented in the same order for all participants (cars in blocks 1 and 3, Ziggerins in blocks 2 and 4) to avoid order effects varying across participants. Distractors in an array were of different identities but all dark or all white in each array. Cars and Ziggerins were not consistently darker or brighter than the faces, to discourage attention based on relative luminance.

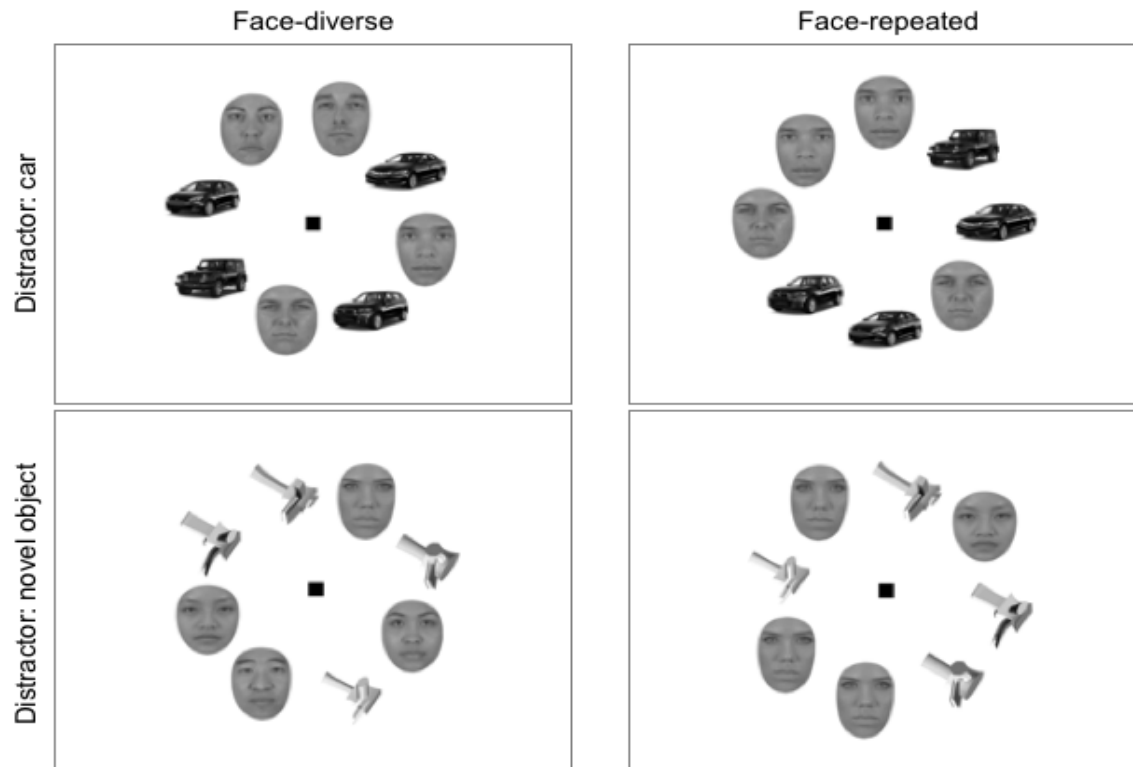


Figure 1. Example arrays from the experiment, see text for details of the task. Both cars and Ziggerins could be darker or lighter than faces, on different trials. The margins around each array are reduced for illustration purpose.

Category-specific tests

To measure car expertise, we used the Cambridge Car Memory Test (CCMT; Dennett et al., 2012) and the Vanderbilt Car Matching Test (VCMT; Sunday et al., 2018). To measure domain-general object recognition, we used two tasks with two different novel object categories (both different from the Ziggerins in the main task): the Novel Object Memory Tests with

Sheinbugs (NOMT; Richler et al., 2019) and a Sequential Matching Task with Greebles (Richler et al., 2019). Face recognition ability was measured using the Cambridge Face Memory Test (CFMT+; Russell et al., 2009). The details for each test are provided elsewhere (see Reference in Table 1). The CCMT, CFMT+ and NOMT are learning tests with similar formats. The VCMT was included to demonstrate that we have sufficient variability in performance with cars to produce a correlation across two different car tasks, and the Sequential matching test helped to provide a better estimate of domain-general object recognition, independent of task demands (Richler et al., 2019).

	Car expertise		Domain-general ability		Face
Name	Cambridge Car Memory Test (CCMT)	Vanderbilt Car Matching Test (VCMT)	Novel Object Memory Test (NOMT)	Sequential Matching Task	Cambridge Face Memory Test (CFMT+)
Reference	Dennett et al. (2012)	Sunday et al. (2018)	Richler et al. (2019)	Richler et al. (2019)	Russell et al. (2009)
Brief description	Select one of 6 studied cars (the same throughout the whole test) in a three-alternative forced-choice. 72 trials.	Select one of 2 studied cars (different for each trial) in a three-alternative forced-choice. 97 trials.	Select one of 6 studied novel objects (the same throughout the whole test) in a three-alternative forced-choice. 48 trials.	Same/Different judgment on pairs of sequentially presented novel objects, separated by a pattern mask. 180 trials.	Select one of 6 studied faces (the same throughout the whole test) in a three-alternative forced-choice. 102 trials.
Mean	0.70	0.66	0.50	1.35	0.66
SD	0.13	0.11	0.11	0.54	0.11
Cronbach's α	0.86	0.83	0.64	0.91	0.87

Table 1. Description of category specific tests that we used in the experiment. Mean accuracy and the SD were reported for all tasks except the Sequential matching task. Mean d' and SD were reported for the sequential matching task.

All measurements achieved good reliability (Cronbach's α for Ensemble task: 0.94, see Table 1 for other tasks). The CCMT and the VCMT were correlated ($r = .57$, 95% CI: [.35, .73]),

suggesting a good range of car expertise. Both tests tap into the domain of car recognition but differ in aspects that are not of interest (the specific cars and task constraints). Performance on these tasks was z-scored and averaged to produce a highly reliable index of car expertise (CAR, weighted reliability = .90, Wang & Stanley, 1970). The NOMT with Sheinbugs and the Matching task with Greebles were also related ($r = .36$, 95% CI: [.06, .59]). This moderate correlation is in line with prior work (Richler et al., 2019). While the two tasks tap into domain-general object recognition, they were selected to differ in aspects that are not of interest (the object categories and task constraints). Performance on these tasks was z-scored and averaged to produce a highly reliable single index of domain-general Object Recognition ability (OBJECT, weighted reliability = .83, Wang & Stanley, 1970). Because of a programming error, 7 participants' NOMT data were not saved and the NOMT results for 3 participants were at chance and with performance no better on the easiest 12 trials than the hardest 12 trials so these scores were taken to reflect a misunderstanding of instructions. In those cases, the z scores in the Matching task alone was used as the OBJECT score.

Results

	Ensemble (Car)	Ensemble (Novel Object)	CAR	OBJECT
Ensemble (Novel Object)	$r = .91$ [.85, .95] BF = 8.938e+17	-		
CAR	$r = .07$ [-.20, .33] BF = 0.19	$r = .11$ [-.17, .37] BF = 0.23	-	
OBJECT	$r = .12$ [-.16, .38] BF = 0.24	$r = .14$ [-.14, .40] BF = 0.28	$r = .32$ [.05, .54] BF = 2.26	-
FACE	$r = .51$ [.28, .69] BF = 281.14	$r = .46$ [.21, .65] BF = 51.18	$r = .26$ [-.01, .50] BF = 1.01	$r = .14$ [-.14, .39] BF = 0.27

Table 2. Correlations among abilities depending on object categories and tasks. BF indicates BF_{10} . The numbers in the square brackets indicate the 95% confidence interval of r .

We used JASP statistics software (JASP Team, 2019) for Bayesian analyses. Table 2 shows the zero-order correlations across tasks. The average performance of two face diversity judgment conditions was similar, without conclusive evidence for or against a difference (car distractor $M = 0.75$, $SD = 0.08$, novel object distractor $M = 0.75$, $SD = 0.09$, $BF_{10} = .60$).

Performance in the two face diversity conditions was highly correlated since the task-relevant aspects of the task are identical. While OBJECT and CAR recognition abilities did not predict face diversity judgments, FACE recognition ability did, most likely because faces were task-relevant. To assess whether car distractors competed with faces during the ensemble task, we measured the correlation between the face diversity task performance with car distractors and CAR, controlling for performance in the face diversity task with novel object distractors. This correlation was small, with Bayesian support favoring the null ($r = -.07$, 95% CI: [-.33, .21], $BF_{10} = 0.19$, $BF_{01} = 5.26$ – See Figure 2). There was substantial (Jeffreys, 1961) support for no

competition from cars during ensemble coding of faces, as a function of car expertise. Similar results were obtained when OBJECT was also controlled to remove any possible influence of domain-general object recognition measured by CAR ($r = -.06$, 95% CI: $[-.33, .21]$, $BF_{10} = 0.19$, $BF_{01} = 5.33$). When contrasted with the several examples of competition between these domains during single object perception, the results suggest that ensemble coding of face and car recognition are more distant in functional space than are individual recognition in the same domains.

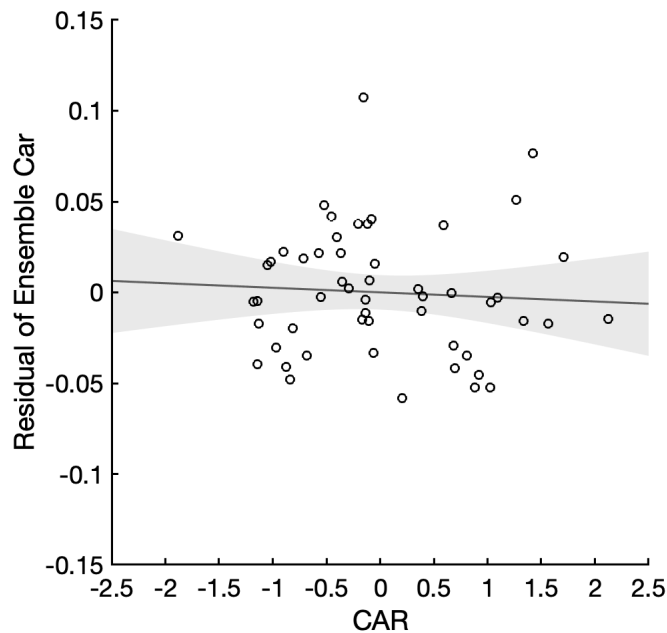


Figure 2. The relationship between CAR and the residual of car distractor condition in ensemble task after controlling the novel object distractor condition. The shaded area indicates the 95% confidence interval.

Discussion

We investigated possible competition between faces and cars during ensemble perception, as a function of car expertise. We found evidence against competition, with results supporting a null effect of car distractors as a function of car expertise, during face diversity judgments.

Importantly, this result was obtained in the context of several robust correlations – between the two distractor conditions for the face diversity task, between the two car tasks operationalizing car expertise and between performance in face diversity judgments and face recognition ability. These effects increase confidence that we could detect competition between faces and cars, had it occurred in this ensemble task. While we did not investigate the strategies used by participants during diversity judgments, relying on attention to individual faces would have only increased reliance on face recognition mechanisms that have previously resulted in competition with cars.

Competition between faces and cars as a function of car expertise arises in a variety of other tasks (Gauthier et al., 2003; McKeeff et al., 2010), including in a visual search task where faces and cars were presented on the screen simultaneously and only faces were task relevant (McGugin et al., 2011). These conditions are very similar to our task, making the evidence against competition all the more surprising. As a result, we suggest that the representations supporting ensemble coding of faces are not the same as those used in individual face recognition, inconsistent with some prior conclusions (e.g., Neumann et al., 2018).

It is difficult to know why something does not occur, of course. During our task, task-irrelevant cars may not have been formed as ensemble representations. There is evidence in the decoding of EEG signals for ensemble processing during presentation of face arrays without any explicit judgment (Roberts et al., 2019), but irrelevant cars may not be processed in the same manner during explicit ensemble coding of another category. Alternatively, the cars may have been processed as ensembles without competing with ensemble processing of faces. It would be useful to explore paradigms where ensemble processing of both categories is task-relevant. Competition from task-irrelevant cars during visual search for faces suggest that cars should be processed in some way (McGugin et al., 2011), leading us to conclude that while we do not

know how cars were processed here, ensemble perception of faces differs from single face perception, where task-irrelevant cars *did* compete with task-relevant faces.

While we found that normal cars did not compete here with faces as a function of car expertise, it is possible that other types of car images could compete to process faces. For instance, Curby and Gauthier (2014) observed competition between faces and car parts in a new configuration. Further studies are needed to explore the role of configuration in the competition between domains of expertise during ensemble coding.

Our result dovetails conclusions from studies suggesting different mechanisms for single object perception and ensemble perception (Cant et al., 2015; Sun & Chong, 2020). For instance, Garner interference, which indicates that two features of an object are not processed independently, was observed between shape and texture during ensemble coding, while no interference between the same features was obtained when single objects were attended (Cant et al., 2015). The same task-irrelevant features or items may interfere or not with processing of task-relevant objects, depending on whether the task requires attention to objects or summary statistics.

Studies of the neural mechanisms involved in ensemble coding support the notion of a distinct mechanism. fMRI studies implicate the lateral occipital cortex in the processing of single objects and the parahippocampal place area in ensemble processing (Cant & Xu, 2012). The right dorsal visual stream was recruited when people judged the average emotion of multiple faces, whereas judgments about the facial expression of a single face engaged ventral regions including the fusiform gyrus (Im et al., 2017). The neural mechanisms of ensemble perception, just like those of individual object recognition, may depend on the specific features or object categories, but they appear to also differ from those of individual object recognition when tasks are

contrasted for the same category. Even when ensemble and individual judgments about face identity are strongly correlated, the EEG time courses of the two types of processing differ, taking longer to accumulate and peak for ensemble processing (Roberts et al., 2019). In contrast, the recognition of individual faces and individual cars as a function of expertise overlap in the fusiform face area (Gauthier et al., 2000; McGugin et al., 2012a), an area that appears to be a locus of competition when both categories are presented simultaneously (McGugin et al., 2015). Car and face processing in dual experts could be more functionally distant (Kinsbourne & Hicks, 1978) during ensemble coding than during object recognition.

Finally, we found that performance in face diversity judgments was highly correlated with face recognition ability, in both distractor conditions. These results are consistent with other work (Alvarez & Oliva, 2008; Haberman et al., 2015a; Roberts et al., 2019; Sun & Chong, 2020) in suggesting that the precision of ensemble representations depends to some extent on the precision of the representations of individual items. Demonstrating this correlation at the level of individual subjects confirms that the precision of ensemble representations *in a given individual* depends on the precision of their individual item representations. In sum, while the correlation between single face recognition ability and ensemble coding of faces suggests a dependency, our result indicates that ensemble representations of faces are also farther in cerebral functional space than individual face representations are from the representation of task-irrelevant cars.

Acknowledgment

This work was supported by NSF (SMA1640681) and by the David K. Wilson Chair Research Fund (Vanderbilt University). We thank Randolph Blake, Ashleigh Maxcey and Oakyoon Cha

for comments on the manuscript and Melanie Kacin, Marcus Bennett and Anjali Mahapatra for help with data collection.

Open Practice Statement

This experiment was not preregistered, but the data are publicly accessible

(<https://doi.org/10.6084/m9.figshare.12126453.v1>) and the materials are available upon request.

References

- Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, 19(4), 392-398.
<https://doi.org/10.1111/j.1467-9280.2008.02098.x>
- Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, 12(2), 157-162. <https://doi.org/10.1111/1467-9280.00327>
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10(4), 433-436.
<https://doi.org/10.1163/156856897X00357>
- Cant, J. S., Sun, S. Z., & Xu, Y. (2015). Distinct cognitive mechanisms involved in the processing of single objects and object ensembles. *Journal of Vision*, 15(4), 1-21.
<https://doi.org/10.1167/15.4.12>
- Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *Journal of Neuroscience*, 32(22), 7685-7700.
<https://doi.org/10.1523/JNEUROSCI.3325-11.2012>
- Cha, O., Blake, R., & Gauthier, I. (in press). The role of category- and exemplar-specific experience in ensemble processing of objects. *Attention, Perception & Psychophysics*.
- Cha, O., Blake, R. & Gauthier, I (submitted). Judgments of average and variability within object ensembles rely on a common ability.
- Cha, O., & Chong, S. C. (2018). Perceived average orientation reflects effective gist of the surface. *Psychological Science*, 29(3), 319-327.
<https://doi.org/10.1177/0956797617735533>
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 43(4), 393-404. [https://doi.org/10.1016/S0042-6989\(02\)00596-5](https://doi.org/10.1016/S0042-6989(02)00596-5)

- Chua, K.-W., & Gauthier, I. (2020). Domain-specific experience determines individual differences in holistic processing. *Journal of Experimental Psychology: General*, 149(1), 31-41. <https://doi.org/10.1037/xge0000628>
- Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences*, 20(5), 324-335. <https://doi.org/10.1016/j.tics.2016.03.006>
- Dennett, H. W., McKone, E., Tavashmi, R., Hall, A., Pidcock, M., Edwards, M., & Duchaine, B. (2012). The Cambridge Car Memory Test: A task matched in format to the Cambridge Face Memory Test, with norms, reliability, sex differences, dissociations from face memory, and expertise effects. *Behavior Research Methods*, 44(2), 587-605. <https://doi.org/10.3758/s13428-011-0160-2>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience*, 6(4), 428-432. <https://doi.org/10.1038/nn1029>
- Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, 3(2), 191-197. <https://doi.org/10.1038/72140>
- Haberman, J., Brady, T. F., & Alvarez, G. A. (2015a). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, 144(2), 432-446. <https://doi.org/10.1037/xge0000053>

- Haberman, J., Lee, P., & Whitney, D. (2015b). Mixed emotions: Sensitivity to facial variance in a crowd of faces. *Journal of Vision*, 15(4), 1-11. <https://doi.org/10.1167/15.4.16>
- Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, 17(17), R751-R753. <https://doi.org/10.1016/j.cub.2007.06.039>
- Haberman, J., & Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 718-734. <https://doi.org/10.1037/a0013899>
- Im, H. Y., Albohn, D. N., Steiner, T. G., Cushing, C. A., Adams, R. B., & Kveraga, K. (2017). Differential hemispheric and visual stream contributions to ensemble coding of crowd emotion. *Nature Human Behaviour*, 1(11), 828-842. <https://doi.org/10.1038/s41562-017-0225-z>
- JASP Team (2019). JASP (Version 0.11.1) [Computer software]
- Jeffreys, H. (1961). *Theory of Probability*. Oxford: Clarendon Press.
- Kinsbourne, M., & Hicks, R. E. (1978). Functional cerebral space: A model for overflow, transfer and interference effects in human performance. In J. Requin (Ed.), *Attention and performance* (pp. 345–362). New York: Academic Press.
- Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, 47(4), 1122-1135. <https://doi.org/10.3758/s13428-014-0532-5>
- McGugin, R. W., Gatenby, J. C., Gore, J. C., & Gauthier, I. (2012a). High-resolution imaging of expertise reveals reliable object selectivity in the fusiform face area related to perceptual performance. *Proceedings of the National Academy of Sciences*, 109(42), 17063-17068. <https://doi.org/10.1073/pnas.1116333109>

- McGugin, R. W., McKeef, T. J., Tong, F., & Gauthier, I. (2011). Irrelevant objects of expertise compete with faces during visual search. *Attention, Perception, & Psychophysics*, 73(2), 309-317. <https://doi.org/10.3758/s13414-010-0006-5>
- McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012b). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision research*, 69, 10-22. <https://doi.org/10.1016/j.visres.2012.07.014>
- McGugin, R.W., Van Gulick, A.E., Tamber-Rosenau, B.J., Ross, D.A., & Gauthier, I. (2015). Expertise effects in face selective areas are robust to clutter and diverted attention but not to competition. *Cerebral Cortex*, 25(9), 2610-2622. <https://doi.org/10.1093/cercor/bhu060>
- McKeef, T. J., McGugin, R. W., Tong, F., & Gauthier, I. (2010). Expertise increases the functional overlap between face and object perception. *Cognition*, 117(3), 355-360. <https://doi.org/10.1016/j.cognition.2010.09.002>
- Michael, E., De Gardelle, V., & Summerfield, C. (2014). Priming by the variability of visual information. *Proceedings of the National Academy of Sciences*, 111(21), 7873-7878. <https://doi.org/10.1073/pnas.1308674111>
- Neumann, M. F., Ng, R., Rhodes, G., & Palermo, R. (2018). Ensemble coding of face identity is not independent of the coding of individual identity. *Quarterly Journal of Experimental Psychology*, 71(6) 1357–1366. <https://doi.org/10.1080/17470218.2017.1318409>
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437-442. <https://doi.org/10.1163/156856897X00366>

- Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). Perceiving groups: The people perception of diversity and hierarchy. *Journal of Personality and Social Psychology*, *114*(5), 766-785. <https://doi.org/10.1037/pspi0000120>
- Richler, J. J., Tomarken, A. J., Sunday, M. A., Vickery, T. J., Ryan, K. F., Floyd, R. J., ... & Gauthier, I. (2019). Individual differences in object recognition. *Psychological Review*, *126*(2), 226-251. <https://doi.org/10.1037/rev0000129>
- Roberts, T., Cant, J. S., & Nestor, A. (2019). Elucidating the Neural Representation and the Processing Dynamics of Face Ensembles. *Journal of Neuroscience*, *39*(39), 7737-7747. <https://doi.org/10.1523/JNEUROSCI.0471-19.2019>
- Rossion, B., Collins, D., Goffaux, V., & Curran, T. (2007). Long-term expertise with artificial objects increases visual competition with early face categorization processes. *Journal of Cognitive Neuroscience*, *19*, 543–555. <https://doi.org/10.1162/jocn.2007.19.3.543>
- Rossion, B., Kung, C. C., & Tarr, M. J. (2004). Visual expertise with nonface objects leads to competition with the early perceptual processing of faces in the human occipitotemporal cortex. *Proceedings of the National Academy of Sciences*, *101*, 14521–14526. <https://doi.org/10.1073/pnas.0405613101>
- Russell, R., Duchaine, B., & Nakayama, K. (2009). Super-recognizers: People with extraordinary face recognition ability. *Psychonomic Bulletin & Review*, *16*(2), 252-257. <https://doi.org/10.3758/PBR.16.2.252>
- Sun, J., & Chong, S. C. (2020). Power of averaging: Noise reduction by ensemble coding of multiple faces. *Journal of Experimental Psychology: General*, *149*(3), 550-563. <https://doi.org/10.1037/xge0000667>

Sunday, M. A., Lee, W. Y., & Gauthier, I. (2018). Age-related differential item functioning in tests of face and car recognition ability. *Journal of Vision*, 18(1), 1-17.

<https://doi.org/10.1167/18.1.2>

Wang, M. W., & Stanley, J. C. (1970). Differential weighting: A review of methods and empirical studies. *Review of Educational Research*, 40(5), 663-705.

<https://doi.org/10.3102/00346543040005663>

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology*, 69, 105-129. <https://doi.org/10.1146/annurev-psych-010416-044232>

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. <https://doi.org/10.1111/j.1467-9280.2009.02430.x>

Zhu, X., & Ramanan, D. (2012, June). *Face detection, pose estimation, and landmark localization in the wild*. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition. <https://doi.org/10.1109/CVPR.2012.6248014>