Domain-specific reports of visual imagery vividness are not related to perceptual expertise

Mackenzie Sunday[1] · Rankin W. McGugin[1] · Isabel Gauthier[1]

Abstract Just as people vary in their perceptual expertise with a given domain, they also vary in their abilities to imagine objects. Visual imagery and perception share common mechanisms. However, it is unclear whether domain-specific expertise is relevant to visual imagery. Although the vividness of visual imagery is typically measured as a domain-general construct, a component of this vividness may be domain-specific. For example, individuals who have gained perceptual expertise with a specific domain might experience clearer mental images within this domain. Here we investigated whether perceptual expertise for cars relates to visual imagery vividness in the same domain, by assessing the correlations between a widely used domain-general measure of visual imagery vividness (the Vividness of Visual Imagery Questionnaire; Marks in *British Journal of Psychology*, *64*, 17–24, 1973), a new measure of visual imagery vividness specific to cars, and behavioral tests of car expertise. We found that domain-specific imagery relates most strongly to general imagery vividness and less strongly to self-reported expertise, while it does not relate to perceptual or semantic expertise.

Generating a mental image is a uniquely private experience. Because of this, it is difficult for researchers to assess visual imagery with methods other than self-report. Indeed, some scientists are skeptical as to whether visual imagery is an actual phenomenon (see Farah, 1988), despite support for the idea. Most evidence of visual imagery has come from imaging techniques, demonstrating that brain regions active during the perception of visual stimuli also show increased activation during visual imagery (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998; Johnson & Johnson, 2014; Kosslyn, Thompson, & Alpert, 1997; Lee, Kravitz, & Baker, 2012; O'Craven & Kanwisher, 2000; Reddy, Tsuchiya, & Serre, 2010; Stokes, Thompson, Cusack, & Duncan, 2009). Moreover, visual stimuli can be reliably decoded from the activity patterns in early visual regions (Albers, Kok, Toni, Dijkerman, & de Lange, 2013; Naselaris, Olman, Stansbury, Ugurbil, & Gallant, 2015). Neuronal recordings also support the idea that perception and imagery overlap (Kreiman, Koch, & Fried, 2000). Behavioral studies have revealed that mental imagery can influence perceptual judgments, furthering the idea that mental imagery is a meaningful experience with behavioral consequences (D'Ascenzo, Tommasi, & Laeng, 2014; Pearson, Clifford, & Tong, 2008).

Though the bulk of visual imagery work has focused on group-level effects, there have been attempts to investigate visual imagery at the individual level. The Vividness of Visual Imagery Questionnaire (VVIQ; Marks, 1973) has become a popular measure of self-reported individual differences in visual imagery. The VVIQ is a self-report test in which participants are asked to rate the vividness of their mental images to prompts such as "visualize a sun rising." The VVIQ tends to produce scores with high internal consistency and shows both convergent validity (as measured by high correlations between the VVIQ and other measures of imagery) and discriminant validity (as measured by a low correlation between the VVIQ and the verbal scale of the Object-Spatial Imagery and Verbal Questionnaire), making it useful for individual differences work (Campos, 2011; Campos & Pérez-Fabello, 2009; McKelvie,

✉ Mackenzie Sunday
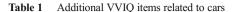 mackenzie.a.sunday@vanderbilt.edu

[1] Department of Psychology, Vanderbilt University, 226 Wilson Hall, Nashville, TN 37204, USA

Springer

1995). Because we are using the VVIQ, the present work can only speak to the relation between self-reported mental imagery's vividness and object recognition performance. Mental imagery is a complex experience, with several aspects that produce measurable effects on the brain and perception. Indeed, there is recent evidence of limited shared variance between imagery vividness and objective measures of visual imagery on subsequent binocular rivalry tasks (Bergmann, Genç, Kohler, Singer, & Pearson, 2015), though on a trial-to-trial basis, self-reported visual imagery vividness predicted rivalry priming. However, both behavioral (e.g., Lima et al., 2015, or Rodway, Gillies, & Schepman, 2006) and imaging (Cui, Jeter, Yang, Montague, & Eagleman, 2007) work has used the VVIQ to achieve correlations between reported mental imagery vividness and other measures.

Here, we investigated visual imagery in a narrow domain of expertise. It is unclear whether individuals with perceptual expertise in a particular domain can generate better mental images of objects from this domain, or whether the quality of visual imagery is independent of perceptual expertise. Perceptual expertise impacts how objects within the domain of expertise are perceived, as evidenced by increased holistic processing within the domain of expertise (Boggan, Bartlett, & Krawczyk, 2012; Bukach, Philips, & Gauthier, 2010; Busey & Parada, 2010). A few studies have reported an influence of expertise on imagery, but only when measuring differences in the ability to mentally manipulate the mental image (Bachmann & Oit, 1992; Hatano, Miyake, & Binks, 1977). Aleman, Nieuwenstein, Böcker, and de Haan (2000) found that musicians were better able to compare mentally imagined tones than were nonmusicians, and the musicians' superior performances did not extend beyond auditory to visual imagery. In the object recognition field, efforts to develop reliable measures of individual differences in different aspects of perceptual expertise are relatively new (Duchaine & Nakayama, 2006; Richler, Floyd, & Gauthier, 2014; Van Gulick, McGugin, & Gauthier, 2015). For this study, we used recently developed measures of perceptual and semantic knowledge for cars, as well as self-reports of experience with cars and a new test of the vividness of mental imagery for cars, to explore the nature of domain-specific visual imagery. Our results showed that imagery within a specific visual domain is mostly predicted by general imagery ability and not by expertise or semantic knowledge of the domain.

Method

To assess whether perceptual expertise predicts visual imagery for objects of expertise, we tested whether car expertise relates to reported car visual imagery. We first created eight new car-specific items (grouped into two sets of questions about two prompts; see Table 1), modeled after the 16 original VVIQ items (Marks, 1973). In creating these items, we purposefully

Table 1 Additional VVIQ items related to cars

Visualization prompt	Item to rate
1. Think of a familiar car that is not your own.	1. The contours of the car body, tires, and windows.
	2. The color of the car in bright sun. Any scratches or details on the car body.
	3. The car pulling into a driveway and parking.
	4. The car facing you and approaching.
2. Think of your dream car if you could have any car you wanted.	1. The car driving on a highway.
	2. The contours of the car body, tires, and windows.
	3. The car being washed.
	4. The car after driving it on a muddy road.

included only visual cues for rating, since our measure of perceptual expertise specifically assesses *visual* expertise. We included only visualization prompts that could be completed without the participants currently owning a car. In this way, we did not exclude any participants who were not currently car owners. Similar to the original VVIQ, we had participants complete the test first with their eyes open and then complete the test again, this time closing their eyes during visualization. To remain consistent with the other self-rating scales used in this study and to avoid confusion, we reversed the scale of the original VVIQ, so that higher numbers indicated a more clear and vivid mental image (as in the VVIQ-2; Marks, 1995), but kept the ratings on a 5-point scale.

To measure other aspects of car expertise, we had the participants complete (1) a perceptual car expertise test measuring the ability to learn and recognize six target cars (the Vanderbilt Expertise Test, or VET; see McGugin, Richler, Herzmann, Speegle, & Gauthier, 2012), and (2) a semantic car expertise test measuring knowledge of car model names (the Semantic Vanderbilt Expertise Test, or SVET; see Van Gulick et al., 2015). Participants also completed the VET and SVET for birds, so that we had an expertise measure to contrast with car expertise. Finally, all participants completed a questionnaire to measure self-reported expertise with both cars and birds.

The vividness of visual imagery questionnaire

The VVIQ asks participants to rate the vividness of images formed in the "mind's eye" (Marks, 1973). The questionnaire consists of four visualization prompts (a familiar face, a sun rising, a familiar storefront, and a country scene) with four items to rate for each prompt, totaling 16 rated items. The questionnaire is completed twice, once with eyes open and once with eyes closed.

The Vanderbilt expertise test

The VET is a visual learning test that includes subtests for multiple domains. In each subtest, participants learn six

exemplars from within a domain and then complete three-alternative forced choice trials. On each trial, one of the studied exemplars is presented along with two distractor images. For the first 12 trials, the exemplar images are identical to the studied exemplar images, and on the following 36 trials, the exemplar images differ from the studied exemplar image in perspective and size. Participants receive feedback on the first 12 trials. In the following 36 trials, there are two catch trials in which the incorrect choices are not from the domain being tested. In the VET-Car/Bird, participants match four target exemplars for six unique car models or bird species, respectively. All of the VET images are digitized grayscale images, and chance accuracy for the VET tests is 33 %. During the VET, we expect participants to use visual short-term and long-term memory (since the VET is a learning task), and there is evidence of an object-of-expertise advantage for visual short-term memory (Curby et al., 2009). Thus, a domain-specific visual imagery component could be relevant to the VET. If this were the case, we would expect to find a correlation between the VVIQ and VET scores. Conversely, to foreshadow our results, we did not find that these two measures relate, suggesting that performance on a car recognition test has little to no bearing on self-reported imagery vividness.

The VET (and similar measures) has been used to define expertise in a large number of published studies and has produced measures with high internal consistencies (Bowles et al., 2009; Dennett et al., 2012; Duchaine & Nakayama, 2006; Gauthier et al., 2014; Germine, Duchaine, & Nakayama, 2011; McGugin, Van Gulick, Tamber-Rosenau, Ross, & Gauthier, 2015; Richler et al., 2011; Woolley, Gerbasi, Chabris, Kosslyn, & Hackman, 2008). Other than perceptual tests, the main alternative way to define expertise is self-report, but research suggests that self-reported expertise is a relatively poor predictor of performance on the VET and other perceptual tasks (Dennett et al., 2012; McGugin, Richler, et al., 2012). In contrast, the VET shows a great deal of validity as an expertise measure. For instance, car experts as identified by the VET achieve higher scores on a test of semantic knowledge for cars (Van Gulick et al., 2015). Moreover, good performance on the VET-Car predicts good performance on other car tasks (like a matching task), even once performance on other VET categories is regressed out (McGugin, Richler, et al., 2012). In other studies, the VET-Car shows a significant relationship (again, even after regressing out performance for other categories) with the fusiform area's response to cars (McGugin, Newton, Gore, & Gauthier, 2014). For our purposes, the VET-Car allows us to make inferences about car-specific effects, to the extent that the effects do not generalize to the VET-Bird.

The semantic Vanderbilt expertise test

The SVET was developed as a measure of semantic domain-specific knowledge independent of visual domain-specific knowledge. In the SVET, participants must choose the real subordinate-level label from two plausible-sounding foil names. For example, in the SVET-Car, participants are presented with three choices ("Volvo Focus, Mercedes-Benz C300, Mercury Alero") and asked to identify the real car ("Mercury Alero"). The test is composed of 48 experimental trials and three catch trials, in which the foil names are very obviously not plausible. Unlike the VET, the SVET is not a learning measure, as each trial is self-contained and participants only use their prior knowledge to respond. The test has produced scores with high internal consistency in all domains (Van Gulick et al., 2015). SVET scores tend to correlate more strongly with within-domain than with between-domain VET scores (Van Gulick et al., 2015).

Self-report of experience questionnaire

The self-report questionnaire is composed of seven questions, measuring various aspects of experience on a 7-point scale (with 7 corresponding to higher expertise; Gauthier et al., 2014). Responses from this questionnaire correlate with both the VET and SVET scores within the car domain, and the experience reported in this questionnaire accounts for most of the shared variance between the SVET and VET (Van Gulick et al., 2015). However, the VET and SVET both have unique variance, presumably because participants have different abilities to learn visually and learn nomenclature (Van Gulick et al., 2015).

Participants

We recruited 200 participants with Amazon Mechanical Turk to complete the VVIQ along with the new car items (125 female, 75 male; mean age = 36.55 years, range = 18–74). We contacted participants three days after they had completed the VVIQ and VVIQ-Car items to offer them the possibility to complete the VET-Car, VET-Bird, SVET-Car, and SVET-Bird (self-report experience questions were included in the SVET-Bird and VVIQ-Car items for birds and cars, respectively). Participants were compensated $0.30 for completing the VVIQ and car items, and given a bonus $2.20 if they completed both the two VET and two SVET tasks. Three participants were excluded from the VET-Bird for incorrectly answering both catch trials. One participant was excluded from the VET-Car for incorrectly answering both catch trials. Of the 99 participants who satisfactorily completed all five of the tasks, 78 were Caucasian (38 female, 40 male; mean age = 36.63 years, age range = 19–74).

Results

Summary statistics for each task are reported in Table 2. Among the 200 participants who completed the VVIQ and

Table 2 Summary statistics and reliabilities for each task ($N = 99$)

	Mean	SD	Cronbach Alpha
SVET-Bird	.503	.143	.802
Self-Report-Bird[a]	2.61	1.28	.878
SVET-Car	.739	.151	.851
VET-Bird	.566	.139	.836
VET-Car	.478	.127	.785
VVIQ Noncar[b]	3.72	0.655	.953
VVIQ-Car[b]	3.67	0.798	.950
Self-Report-Car[a]	3.62	1.36	.900

For each participant, the average accuracy or rating for the entire test was computed and used in all subsequent calculations. [a] 7-point scale. [b] 5-point scale

Table 4 Results of a multiple regression predicting VVIQ-Car with VVIQ-Noncar (VVIQ-NC), self-report car (SR-Car), VET-Car, and SVET-Car entered simultaneously ($N = 99$)

Model and Predictor	B	SE	t	p	partial-R
VVIQ-Car ($R^2_{adj} = 71.6$ %)					
Intercept	.064	.332	0.191	.849	–
VVIQ-NC	.959	.067	14.3	≤.0001	.828
SR-Car	.134	.034	3.97	≤.0001	.379
VET-Car	−.568	.377	−1.51	.135	−.154
SVET-Car	−.023	.340	−.681	.497	−.070

The correlations reported in the rightmost column are the partial correlations of the row predictor with VVIQ-Car, with the variance from all other predictors partialed out

the additional VVIQ-Car items, we found a strong positive correlation between the average rating of the noncar items (VVIQ-NC) and the average rating of the car items (VVIQ-Car; $r_{200} = .742$, $p < .0001$). The internal consistency of the test as a whole (both car and noncar) was high ($\alpha = .960$), as were the internal consistencies for the VVIQ-NC and VVIQ-Car items separately ($\alpha = .940$ for VVIQ-NC, $\alpha = .934$ for VVIQ-Car).

The correlations between all five tasks are shown in Table 3, as well as the correlations disattenuated for measurement error (Nunnally, 1970). Age and sex accounted for less than 1 % of the VVIQ-Car variance, and were thus disregarded.

Given that the VVIQ-Car items had high reliability, our aim was to ask what predicted which participants reported a strong ability to imagine cars. The strongest predictor of VVIQ-Car scores was the VVIQ-NC ($r_{99} = .824$, $r_{corr} = .866$). Neither of the performance tests for cars (SVET and VET) was a significant predictor of VVIQ-Car scores. Self-reported car expertise was a significant predictor of VVIQ-Car scores ($r_{99} = .349$, $r_{corr} = .377$), and this relation remained

even after partialing out both VVIQ-NC and self-report for birds ($r_{partial} = .398$).

A multiple regression on VVIQ-Car scores with VET-Car, SVET-Car, self-report for car, and VVIQ-NC items entered simultaneously as predictors accounted for 71.6 % of the VVIQ-Car variance (see Table 4). Neither VET nor SVET residuals were significant predictors of VVIQ-Car scores. The inclusion of VET, SVET, and self-report scores for birds had little impact on the fit of the model (i.e., R^2_{adj} went from .716 to .734).

Discussion

We investigated how perceptual expertise for cars relates to self-reports of imagery for cars. To do this, we created eight new car items modeled after the VVIQ. These items had high internal consistency, along with the rest of the questionnaire, and strongly correlated with the standard VVIQ items. Using two measures of car expertise (one visual and one semantic) and a self-report of experience with cars, we examined the

Table 3 Pearson's product-moment correlations between all five tasks ($N = 99$)

	1	2	3	4	5	6	7	8
1. SR-Bird	–	.333	.511	.440	.088	.001	.246	.158
2. VET-Bird	.285**	–	.545	.131	.309	.275	.283	.144
3. SVET-Bird	.429***	.446***	–	.149	.107	.268	.240	.080
4. SR-Car	.391***	.114	.127	–	.319	.341	.265	.411
5. VET-Car	.073	.250*	.085	.268**	–	.587	.195	.101
6. SVET-Car	.001	.232*	.221*	.298**	.480***	–	.229	.165
7. VVIQ-NC	.225*	.253*	.210*	.245*	.169	.206*	–	.889
8. VVIQ-Car	.144	.128	.070	.380***	.087	.148	.846***	–

The scores for the SVET and VET are the average accuracies across all trials, and the scores for the VVIQ items and self-report are the average ratings. Experimentally determined correlations are bolded in the bottom left corner, and disattenuated correlations are in the upper right corner. SR = self-report, NC = noncar. *** $p < .001$. ** $p < .01$. * $p < .05$

predictors of reports of the vividness of visual imagery for cars.

By far the best predictor of vivid visual imagery for cars was general vivid visual imagery reports. Interestingly, those individuals who reported experience with and interest in cars also tended to report better imagery for cars, but this was entirely independent of the actual quality of their perceptual skills for cars or their knowledge of car models. This could indicate that, perhaps counterintuitively, an expert cannot necessarily generate a more clear and vivid image of an object within their domain of expertise. However, given how poorly self-reports of expertise tend to predict perceptual skills (McGugin, Richler, et al., 2012), it may also be that it is difficult to predict the quality of our own visual imagery in a given domain, as we have virtually no information about others' mental images. In other words, similar to what is found in perception, experts may be unaware that their domain-specific mental images are more vivid than those of novices. There is some evidence that participants are able to assess the vividness of their own visual imagery when comparing the vividness of their own imagery between trials (Pearson, Rademaker, & Tong, 2011), but only for the imagery of simple patterns, not imagery within an expertise domain. Cars represent a domain where self-reports of experience show a modest but significant relationship with visual skills (here, $r_{99} = .268$; see also Van Gulick et al., 2015), whereas this relation is often absent in other domains (McGugin, Richler, et al., 2012; Van Gulick et al., 2015). Thus, it is interesting that even in this domain, the portion of variance in self-reports that predicts visual imagery scores is independent from the one that predicts perceptual performance for cars.

Our findings have practical applications for research in mental imagery. They suggest that when measuring reported visual imagery ability, even if one is interested in a specific domain, not much is gained by obtaining domain-specific ratings—domain-general questions such as those in the VVIQ and domain-specific questions about cars appear to provide redundant information. However, our intuition is that only someone with very good perceptual knowledge of cars will be able to create distinct mental images for different cars, but the ability to create visual images of different cars was not evaluated here.

Of course, it is possible that perceptual or semantic car knowledge predicts the quality of visual imagery in a way that is not accessible to participants. Because we used self-report measures of visual imagery vividness, we cannot speak to the relation between objective measures of visual imagery and perceptual expertise. Luckily this is a question that may be addressed using functional magnetic resonance imagery. Responses on the VVIQ have been found to relate to activity in early visual cortex (Cui et al., 2007). In contrast, performance on perceptual tasks with cars relates to activity in higher-level visual areas, including the fusiform face area (McGugin, Gatenby, Gore, & Gauthier, 2012). Future work could investigate the reliability of neural signals during visual imagery of different cars in each of these areas, and the relation with the behavioral measures we collected here, as well as with objective measures of visual imagery (e.g., binocular rivalry priming; see Pearson et al., 2008). One possibility is that while domain-general VVIQ scores predict individual differences in low-level visual areas, perceptual expertise could predict the representation quality in the fusiform face area.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology, 23,* 1427–1431.

Aleman, A., Nieuwenstein, M. R., Böcker, K. B., & de Haan, E. H. (2000). Music training and mental imagery ability. *Neuropsychologia, 38,* 1664–1668.

Bachmann, T., & Oit, M. (1992). Stroop-like interference in chess players' imagery: An unexplored possibility to be revealed by the adapted moving-spot task. *Psychological Research, 54,* 27–31.

Bergmann, J., Genç, E., Kohler, A., Singer, W., & Pearson, J. (2015). Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cerebral Cortex*. doi:10.1093/cercor/bhv186. **Advance online publication**.

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,* 37–42.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L.,…Yovel, G. (2009). Diagnosing prosopagnosia: Effects of ageing, sex, and participant–stimulus ethnic match on the Cambridge Face Memory Test and Cambridge Face Perception Test. *Cognitive Neuropsychology, 26,* 423–455.

Bukach, C. M., Philips, W. S., & Gauthier, I. (2010). Limits of generalization between categories and implications for theories of category specificity. *Attention, Perception, & Psychophysics, 72,* 1865–1874. doi:10.3758/APP.72.7.1865

Busey, T. A., & Parada, F. J. (2010). The nature of expertise in fingerprint examiners. *Psychonomic Bulletin & Review, 17,* 155–160. doi:10.3758/PBR.17.2.155

Campos, A. (2011). Internal consistency and construct validity of two versions of the Revised Vividness of Visual Imagery Questionnaire 1. *Perceptual and Motor Skills, 113,* 454–460.

Campos, A., & Pérez-Fabello, M. J. (2009). Psychometric quality of a revised version Vividness of Visual Imagery Questionnaire. *Perceptual and Motor Skills, 108,* 798–802.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research, 47,* 474–478.

Curby, K. M., Glazek, K., & Gauthier, I. (2009). A visual short-term memory advantage for objects of expertise. *Journal of*

Experimental Psychology: Human Perception and Performance, 35, 94–107. doi:10.1037/0096-1523.35.1.94

D'Ascenzo, S., Tommasi, L., & Laeng, B. (2014). Imagining sex and adapting to it: Different aftereffects after perceiving versus imagining faces. *Vision Research, 96,* 45–52.

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,* 587–605. doi:10.3758/s13428-011-0160-2

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,* 576–585.

Farah, M. J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review, 95,* 307–317. doi:10.1037/0033-295X.95.3.307

Gauthier, I., McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Van Gulick, A. E. (2014). Experience moderates overlap between object and face recognition, suggesting a common ability. *Journal of Vision, 14*(8), 7. doi:10.1167/14.8.7

Germine, L. T., Duchaine, B., & Nakayama, K. (2011). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition, 118,* 201–210.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: Direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience, 10,* 1563–1573.

Hatano, G., Miyake, Y., & Binks, M. G. (1977). Performance of expert abacus operators. *Cognition, 5,* 47–55.

Johnson, M. R., & Johnson, M. K. (2014). Decoding individual natural scene representations during perception and imagery. *Frontiers in Human Neuroscience, 8,* 59. doi:10.3389/fnhum.2014.00059

Kosslyn, S. M., Thompson, W. L., & Alpert, N. M. (1997). Neural systems shared by visual imagery and visual perception: A positron emission tomography study. *NeuroImage, 6,* 320–334. doi:10.1006/nimg.1997.0295

Kreiman, G., Koch, C., & Fried, I. (2000). Imagery neurons in the human brain. *Nature, 408,* 357–361.

Lee, S. H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage, 59,* 4064–4073.

Lima, C. F., Lavan, N., Evans, S., Agnew, Z., Halpern, A. R., Shanmugalingam, P.,…Warren, J. E. (2015). Feel the noise: Relating individual differences in auditory imagery to the structure and function of sensorimotor systems. *Cerebral Cortex, 25,* 4638–4650.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64,* 17–24. doi:10.1111/j.2044-8295.1973.tb01322.x

Marks, D. F. (1995). New directions for mental imagery research. *Journal of Mental Imagery, 19,* 153–167.

McGugin, R. W., Gatenby, J. C., Gore, J. C., & Gauthier, I. (2012). 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,* 17063–17068. doi:10.1073/pnas.1116333109

McGugin, R. W., Newton, A. T., Gore, J. C., & Gauthier, I. (2014). Robust expertise effects in right FFA. *Neuropsychologia, 63,* 135–144. doi:10.1016/j.neuropsychologia.2014.08.029

McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision Research, 69,* 10–22. doi: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,* 2610–2622. doi:10.1093/cercor/bhu060

McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: A critical quantitative review and plea for direction. *Journal of Mental Imagery, 19,* 1–106.

Naselaris, T., Olman, C. A., Stansbury, D. E., Ugurbil, K., & Gallant, J. L. (2015). A voxel-wise encoding model for early visual areas decodes mental images of remembered scenes. *NeuroImage, 105,* 215–228.

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

O'Craven, K. M., & Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *Journal of Cognitive Neuroscience, 12,* 1013–1023.

Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology, 18,* 982–986. doi:10.1016/j.cub.2008.05.048

Pearson, J., Rademaker, R. L., & Tong, F. (2011). Evaluating the mind's eye: The metacognition of visual imagery. *Psychological Science, 22,* 1535–1542. doi:10.1177/0956797611417134

Reddy, L., Tsuchiya, N., & Serre, T. (2010). Reading the mind's eye: Decoding category information during mental imagery. *NeuroImage, 50,* 818–825. doi:10.1016/j.neuroimage.2009.11.084

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

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. doi:10.1167/14.11.10

Rodway, P., Gillies, K., & Schepman, A. (2006). Vivid imagers are better at detecting salient changes. *Journal of Individual Differences, 27,* 218–228.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience, 29,* 1565–1572. doi:10.1523/JNEUROSCI.4657-08.2009

Van Gulick, A. E., McGugin, R. W., & Gauthier, I. (2015). Measuring nonvisual knowledge about object categories: The Semantic Vanderbilt Expertise Test. *Behavior Research Methods.* doi:10.3758/s13428-015-0637-5. **Advance online publication**.

Woolley, A. W., Gerbasi, M. E., Chabris, C. F., Kosslyn, S. M., & Hackman, J. R. (2008). Bringing in the experts how team composition and collaborative planning jointly shape analytic effectiveness. *Small Group Research, 39,* 352–371.