Behav Res (2012) 44:587–605
DOI 10.3758/s13428-011-0160-2

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

Hugh W. Dennett · Elinor McKone · Raka Tavashmi · Ashleigh Hall · Madeleine Pidcock · Mark Edwards · Bradley Duchaine

Published online: 20 October 2011
© Psychonomic Society, Inc. 2011

Abstract Many research questions require a within-class object recognition task matched for general cognitive requirements with a face recognition task. If the object task also has high internal reliability, it can improve accuracy and power in group analyses (e.g., mean inversion effects for faces vs. objects), individual-difference studies (e.g., correlations between certain perceptual abilities and face/object recognition), and case studies in neuropsychology (e.g., whether a prosopagnosic shows a face-specific or object-general deficit). Here, we present such a task. Our Cambridge Car Memory Test (CCMT) was matched in format to the established Cambridge Face Memory Test, requiring recognition of exemplars across view and lighting change. We tested 153 young adults (93 female). Results showed high reliability (Cronbach's alpha = .84) and a range of scores suitable both for normal-range individual-difference studies and, potentially, for diagnosis of impairment. The mean for males was much higher than the mean for females. We demonstrate independence between face memory and car memory (dissociation based on sex, plus a modest correlation between the two), including where participants have high relative expertise with cars. We also show that expertise with real car makes and models of the era used in the test significantly predicts CCMT performance. Surprisingly, however, regression analyses imply that there is an effect of sex per se on the CCMT that is *not* attributable to a stereotypical male advantage in car expertise.

Keywords Object recognition · Face recognition · Agnosia · Within-class discrimination · Expertise · Sex differences

H. W. Dennett (✉) · E. McKone · A. Hall · M. Pidcock · M. Edwards
Department of Psychology (Building 39),
Australian National University,
Canberra ACT 0200, Australia
e-mail: hugh.dennett@anu.edu.au

E. McKone
Australian Research Council Centre of Excellence in Cognition and its Disorders (CCD),
Dept of Psychology,
Canberra, Australia
e-mail: elinor.mckone@anu.edu.au

R. Tavashmi
School of Health in Social Science, University of Edinburgh,
Edinburgh, UK

B. Duchaine
Department of Psychological and Brain Sciences,
Dartmouth College,
Hanover, NH, USA

In many areas of visual recognition research, it is important to have available a task testing within-class discrimination of other objects. Such a task would be useful for investigations of object recognition, and the need for such tasks has also been commonly discussed in the face-processing literature. For faces, the natural task for most humans is individual-level identification, such as "Bill" versus "Sam." Although for other objects the natural task requirement in everyday life is often basic-level categorization (e.g., car vs. table), many researchers have pointed out that where the aim is to allow a fair comparison with face recognition, an object task must, like the face recognition task, require discrimination of exemplars within a single object category (e.g., "Rex" vs. "Fido" from the category *Labrador*; Damasio, 1985; Diamond & Carey, 1986;

Gauthier, Behrmann, & Tarr, 1999). This type of comparison is theoretically important in many settings. For example, studies investigating group differences might ask whether faces and objects differ in terms of their sensitivity to variables such as inversion or contrast reversal (e.g., Aguirre, Singh, & D'Esposito, 1999; Diamond & Carey, 1986; Robbins & McKone, 2007; Yin, 1969). In single-case analyses, one might wish to know whether an individual with prosopagnosia has a face-specific disorder or, instead, suffers a broader deficit in telling apart highly similar objects of any category (e.g., Busigny, Graf, Mayer, & Rossion, 2010; Duchaine, Yovel, Butterworth, & Nakayama, 2006; Henke, Schweinberger, Grigo, Klos, & Sommer, 1998). And, in studies of individual differences in the normal population, correlations between object and face recognition can assess the degree to which these abilities depend on common processes (McKone et al., in press; Wilmer, Germine, Chabris, et al., 2010), and heritability can be assessed when twins are studied (Zhu et al., 2010).

The aim of the present article is to present a new task assessing individual-level discrimination of objects—the Cambridge Car Memory Test (CCMT)—which can be used as a comparison task for a well-established test for faces, the Cambridge Face Memory Test (CFMT; Duchaine & Nakayama, 2006). Leaving aside, for the moment, the issue of why we selected cars as the specific object class, our general rationale underlying development of an object test was as follows.

Ideally, an object comparison task should have, at a minimum, the following properties. (1) In test format, it should be closely matched to the face task of interest, in order to equate general cognitive requirements (load placed on memory, concentration, and so on); that is, it should have the same task duration, number of items, number of targets versus distractors, and so on. (2) It should have high internal reliability. High reliability of the task at the level of individual participants increases the statistical power for all types of analyses, and reliability is particularly important when the test is to be used in a clinical setting (Aiken, 2003; Kaplan & Saccuzzo, 2005)—for example, in screening for a clinical disorder in object recognition. For this purpose, we aimed for a minimum criterion of at least .80, consistent with screening instruments used in other areas of psychology (e.g., the Beck Depression Inventory; Beck, Steer, & Carbin, 1988). (3) The test should be short. For use in cognitive neuropsychology, where participants typically need to be given an extensive battery of tests, a maximum duration of 10–15 min per test is desirable. (4) Although there is no one perfect control object class for faces (if it were matched on *all* variables, it would *be* a face), the object class used should be similar to the face task of interest in at least basic perceptual requirements (e.g., all one view or requiring generalization across views). And (5) the object task should measure a theoretical construct independent of that tapped by the face recognition test (i.e., so that both tasks are not tapping merely general visual memory ability), as can be demonstrated by assessing the correlation between the object and face tasks and/or by demonstrating dissociations between the tasks on the basis of other variables.

This combination of requirements has not been met by tasks used in previous literature. For example, Robbins and McKone (2007) used a typical old–new memory format test, requiring individual-level discrimination of Labradors. This task was matched to the face test used in that study for test structure (satisfying criterion 1) and stimulus format (all same-view recognition; satisfying 4), was fairly short (5 min; satisfying 3), and demonstrated theoretical independence from face recognition on the basis of dissociation via inversion (satisfying 5). However, we have recently analyzed the internal reliability of the task, and even combining the two stimulus sets of the original study to give 30 learned Labradors and 60 test Labradors (half old, half new), Cronbach's alpha for the task was only .31 (failing criterion 2; Dawel & McKone, 2011, in preparation). In another example, Wilmer, Germine, Chabris, et al. (2010) reported an Abstract Art Test, requiring individual-level discrimination of different examples of paintings in a given style. With 50 learned items, the test has good internal reliability (α = .79, satisfying criterion 2), and demonstrates independence from face recognition (correlation with CFMT, only r = .26, satisfying 3). However, it does not satisfy criterion 1 or 4, in that the test structure (number of items, etc.) was not matched to the face task (CFMT) and, furthermore, the art images were all of intrinsically 2-D paintings, while the faces were images of 3-D real-world objects and required recognition across view and lighting change.

We chose the CFMT structure (Fig. 1) as the starting point for our new test because it produces very good reliability (α = .88, Bowles et al., 2009; α = .83, Herzmann, Danthiir, Schacht, Sommer, & Wilhelm, 2008; α = .90, Wilmer, Germine, Chabris, et al., 2010), requires only 10–15 min to administer, and has become a standard and frequently used tool in face recognition studies. This includes use in the diagnosis of potential prosopagnosia (e.g., Bate, Haslam, Tree, & Hodgson, 2008; Bowles et al., 2009; DeGutis, Bentin, Robertson, & D'Esposito, 2007; Iaria, Bogod, Fox, & Barton, 2009), in the selection of people with exceptional face recognition abilities (Woolley, Gerbasi, Chabris, Kosslyn, & Hackman, 2008), and in the study of individual differences in the typical population (e.g., Wilmer, Germine, Chabris, et al., 2010).

Turning to the issue of why we chose *cars* in particular, we had several motivations for choosing this class. First, like faces, cars are real-world objects. Second, like faces, cars are three-dimensional, and so recognition can be tested over view and lighting change. Third, like faces, our car exemplars

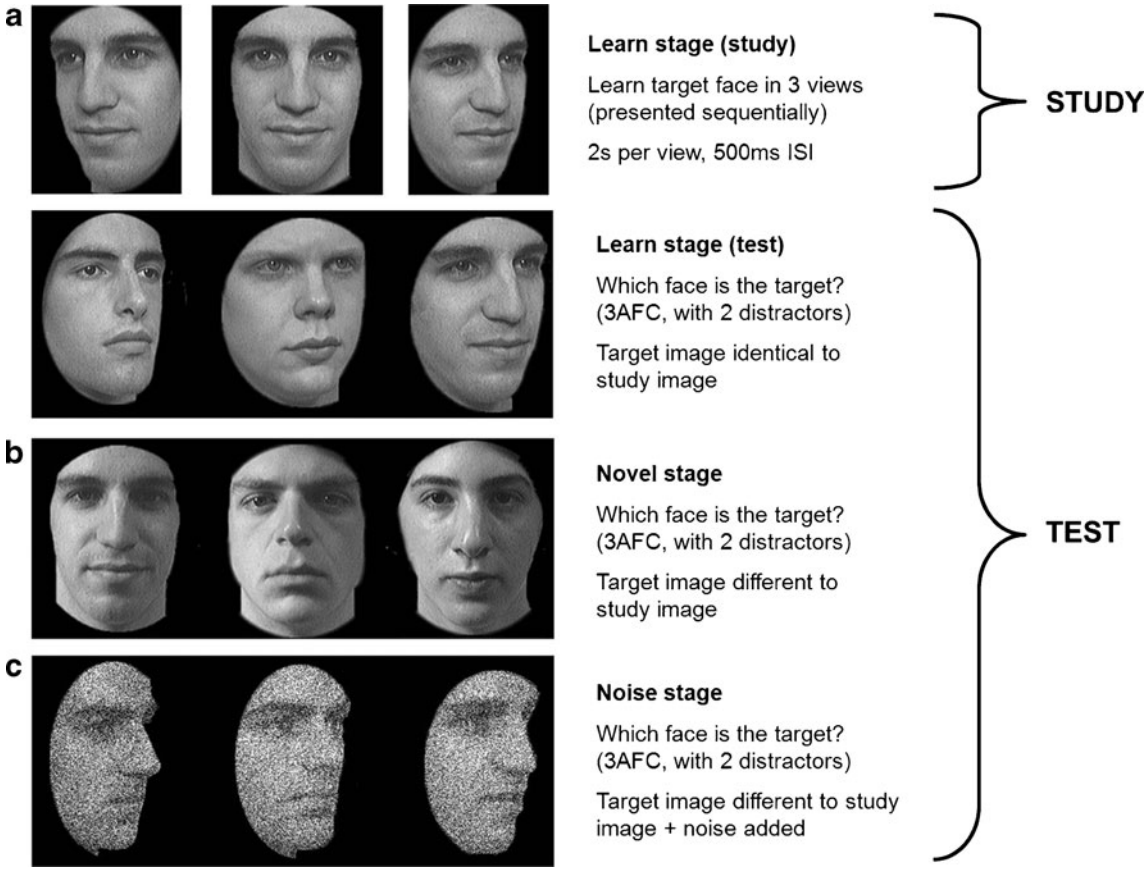


Fig. 1 The Cambridge Face Memory Test contains three stages. (**a**) *Learn* (also called the *introduction* or *same- images*) stage: Participants learn a target face in three viewpoints, then choose a learned image of the target from two matched-format distractors in 3AFC test slides; this procedure is repeated for six target faces. (**b**) *Novel* stage: On each 3AFC test slide, the target face can be any one of the six learned targets, always presented in a novel image (i.e., different viewpoint and/or lighting from the learn stage). (**c**) *Noise* stage: This is as for the novel stage, except that visual noise is added to all stimuli; the target face images are again new (i.e., different again from those used in either the novel or the learn stages)

shared the same *first-order* structure (Diamond & Carey, 1986), with body, four wheels, doors, windscreen, headlights, and so on all in a fixed relationship to each other. And, most important, these similarities have led cars to be used as objects of choice by many different laboratories. Indeed, cars have been perhaps the single most common comparison objects for faces: They have—normally, in conjunction with other stimuli—been widely used in studies assessing whether cases of prosopagnosia reflect a deficit specific to faces (e.g., Bentin, Deouell, & Soroker, 1999; Busigny et al., 2010; De Haan, Young, & Newcombe, 1991; Duchaine & Nakayama, 2005; Duchaine et al., 2006; Henke et al., 1998; Sergent & Signoret, 1992), in studies of object agnosia per se (e.g., Delvenne, Seron, Coyette, & Rossion, 2004), in studies of autism spectrum disorder (Koldewyn, Weigelt, & Kanwisher, 2011; Wallace, Coleman, & Bailey, 2008; Wolf et al., 2008), and in studies that investigated differences and similarities between within-class discrimination of face and object recognition in typically developing participants (e.g., Gauthier, Curran, Curby, & Collins, 2003; Grill-Spector, Knouf, & Kanwisher, 2004; Harel, Gilaie-Dotan, Malach, & Bentin, 2010; Rossion et al., 2000).

Cars are, of course, not matched to faces on all variables. One variable on which they are not matched is the extent to which preexperimental familiarity with specific real-world exemplars could potentially affect performance. For the CFMT, all faces are novel to participants, and so effects of differences in background knowledge between participants are likely to be, at most, mild, arising only from greater or lesser familiarity with the general ethnicity of the faces. (Evidence that this can occur, even when all faces are Caucasian, is that memory for Turkish and German face stimuli interacts with the Turkish or German ethnicity of participants; Sporer, Trinkl, & Guberova, 2007). For car tasks, in contrast, car stimuli are typically closely based on real-world car models. It is theoretically plausible that a preexperimental interest in cars and, thus, knowledge of makes and models of real cars can assist observers to improve their performance on the CCMT, at least to the extent that models are similar to the style of cars used in the CCMT. This is because memory is typically improved

where material can be encoded meaningfully using preexperimental knowledge (Baddeley, 1990), and there are also potential advantages of dual visual–verbal encoding (Paivio, 1969)—that is, where participants can additionally provide themselves with a plausible verbal description of the image (e.g., "this one looks like an Audi A4, except the headlights are rounder"). Given that this issue has generally been ignored in the previous neuropsychological studies using cars, we decided here to also assess preexperimental car expertise and to examine its effects on our car task norms and its influence (or lack of influence) on other questions of interest in the study.

The present study describes the new CCMT. We report its internal reliability and its range of scores in the normal population. We also test the extent to which it taps overlapping or independent processes from face recognition (CFMT). This is theoretically important, noting that the CCMT is the first object test to match both the general cognitive requirements of the CFMT and its requirement for recognition of individual exemplars across view and image change. Next, we test the effects of sex on the CCMT. Many researchers have implicitly acknowledged the likelihood of a sex difference (male advantage) on car recognition—for example, by matching the proportion of males in a car experts group to that in car novice group (e.g., Rossion & Curran, 2010). However, not all have done so; for example, one study had more males in a car expert group than in the car novice group (Grill-Spector et al., 2004), and male prosopagnosic participants have been compared with mixed-sex control samples in two studies using cars as control stimuli (Duchaine & Nakayama, 2005; Duchaine et al., 2006). Here, we formally test for sex differences in mean and/or variance of performance on the CCMT. Finally, we investigate the role of preexperimental interest in and knowledge of cars, asking the following: whether these variables affect CCMT performance; whether independence of face and car memory disappears in individuals with high preexperimental car expertise; and whether car interest or expertise can account for sex differences on the CCMT.

Method

Participants

One hundred fifty-three young adults (18–32 years of age, M = 20.63, SD = 2.88) completed the CCMT as a component of one of two different studies. There were 60 males and 93 females; we aimed to have more than 50 in each sex group because, at this sample size, the statistical issues that arise when comparing a single case with a "small" norm sample disappear (i.e., when $N < 50$, instead of comparing a z-score with a control sample, an adjusted t-test must be used to account for underestimation of the control sample variance; Crawford & Howell, 1998). Of these 153 participants, 142 also completed the CFMT, 75 completed a multiple-choice test of knowledge of car makes and models, and 76 answered survey questions regarding their level of interest in cars and their knowledge of car makes and models. Participants were members of the Australian National University community and received either course credit for a first-year psychology course or payment ($30 for a total across sessions of 2.5 h). The vast majority (i.e., 98%) of the participants were Caucasian (i.e., the same race as the faces used in the CFMT), and those who were not had lived in Australia for most (i.e., > 90%) of their lives and were of mixed descent, with one Caucasian parent and one East Asian or South Asian parent. No participants displayed obvious signs of nonattendance to the tasks.

Apparatus

In Study 1, all the stimuli were displayed using an eMac computer with a 40-cm CRT screen at 1,152 × 864 resolution using PsyScope X software (Cohen, MacWhinney, Flatt, & Provost, 1993). In Study 2, the CCMT and the CFMT were displayed on a Dell PC with a 56-cm LCD screen at 1,680 × 1,050 resolution using standalone executable versions of the tests; and the objective test of car makes and models was displayed using an eMac, with details as for Study 1. All stimuli were viewed at approximately 60 cm, and participants were able to move their head freely.

Cambridge Car Memory Test

The CCMT (created by authors Duchaine and Tavashmi) was administered following the same procedure as the CFMT (Duchaine & Nakayama, 2006), but the stimuli were cars rather than faces. Figure 2 illustrates the general procedure of the task, using cars that are similar in appearance to those in the test but are not actual test items. Participants were required to learn six cars and discriminate the target cars from two distractor cars on each test trial.

Stimuli The stimuli (see example in Fig. 3) were modified computer-generated images of actual car makes and models, created in 3D Studio Max. To minimize memory based on single simple features, all cars were the same color, and no identifying badges, logos, and insignias were visible. The car types used included a mix of sedans, sports cars, and wagons (as both targets and distractors). It was not possible to use all cars with only a single body type, such as all sedans, because this would have made the task too difficult and, thus, unsuitable as a neuropsychological test, where a fairly high control mean is required to allow room to see impaired individuals below the control group. (Note that this

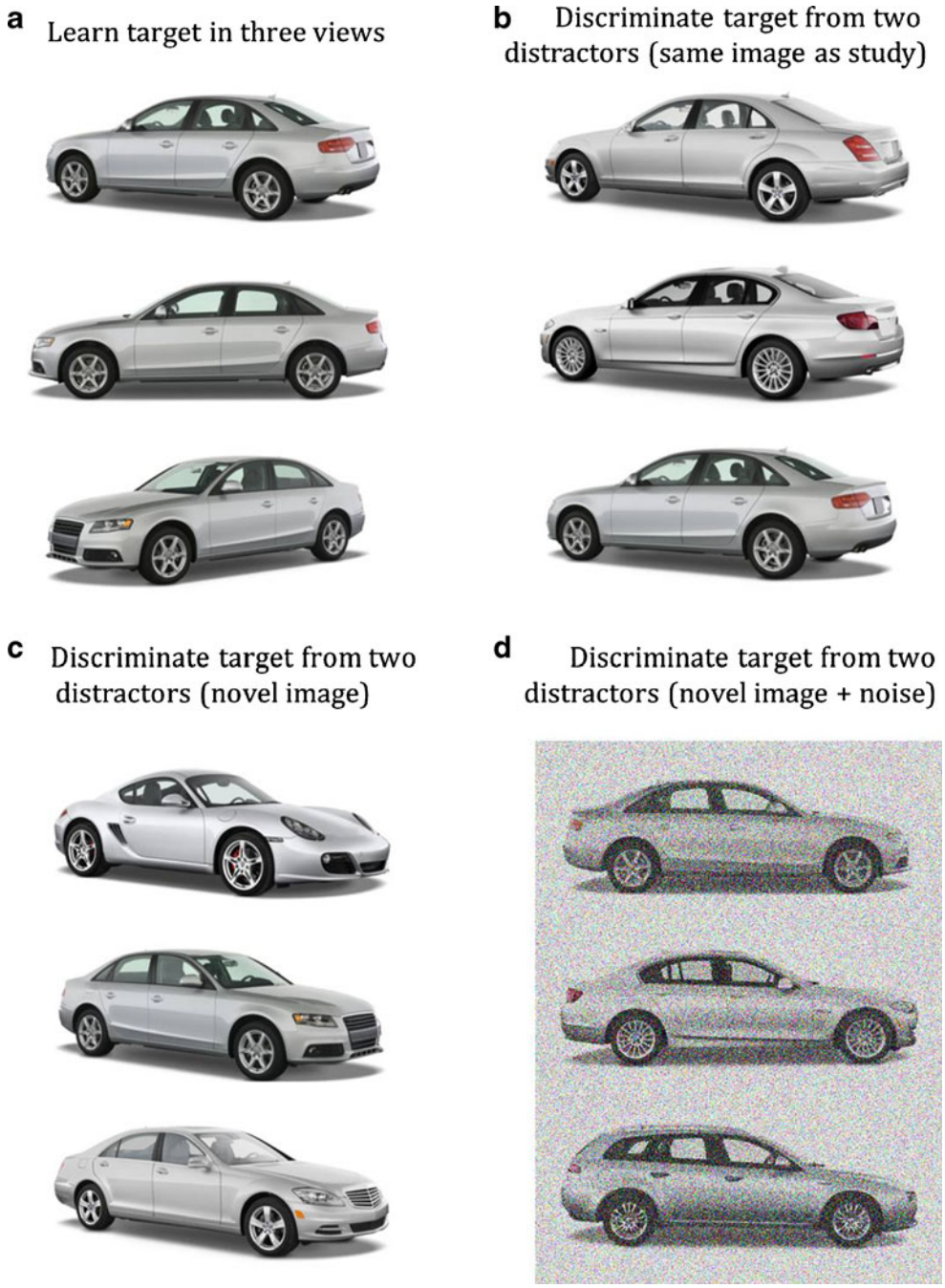


Fig. 2 Format and structure of the CCMT. (**a**) *Learn* study phase. (**b**) *Learn* test phase. (**c**) *Novel* test phase. (**d**) *Noise* test phase. **Note.** The car stimuli show in this figure are not real items used in the CCMT, but show the viewpoints used and the approximate noise appearance

Fig. 3 An Example of an actual car stimulus from the CCMT

issue applies to any object class, not only cars; see, e.g., dogs in Crookes & McKone, 2009, their Fig. 2.)

In addition to the 6 target cars, 46 different distractor cars were included in the test. Of these, 2 were shown only once, 13 were shown twice, 14 were shown thrice, 11 were shown four times, and 6 were shown five times. When distractor cars were repeated, this was sometimes in the same image and sometimes in a different image of a car. Images of cars appeared in different viewpoints and under varied lighting conditions. Car stimuli for the CCMT were

approximately 45 × 125 mm, subtending 4.3° × 11.8° of visual angle (average across cars and viewpoints).

Procedure The test was divided into three sections. In the *learn* stage (alternatively known as *introduction* or *same images*), participants were introduced to the first target car via three study images that varied in viewpoint: side-on; rotated approximately 30° left from side-on; and rotated about 30° right from side-on. Each image was presented for 3 s, with a 500-ms interstimulus interval. Following the three study images, participants responded to three test items. A test item comprised three cars, one of which was the target car that had just been introduced and two of which were distractors. With a test item, distractor and target cars in that item were matched for viewpoint and lighting conditions (this was true for all stages of the test). The test images of the target car used in the introduction were identical to the study images of the target car. The five other target cars were introduced in the same manner. After the final test item for the sixth target car, a review slide was presented for 20 s. This showed all six target vehicles simultaneously in a 2 × 3 layout, in a side-on view.

In the *novel* stage, the 30 test items consisted of a target car and two distractors presented in novel views and/or lighting conditions. This stage is designed to assess recognition of the car, rather than a specific image of that car. The trials were presented in a fixed, pseudorandom order. Following these test trials, the review slide was presented again for 20 s.

The final stage, *noise*, comprised 24 test items, which presented a target car and two distractors in novel viewpoints and/or lighting conditions with the addition of Gaussian noise over the image. The addition of noise increased the level of difficulty and disrupted ability to recognize features of the vehicles.

The test included a total of 72 test items (18 + 30 + 24). Total scores (items correct out of 72) were calculated by summing the number of correct items from the three stages.

Rationale for combining CCMT data across Study 1 and Study 2

Study 1 (authors, Hall, Pidcock, McKone, Duchaine) and Study 2 (Dennett, McKone, Edwards) both used exactly the same procedure for the CCMT itself. Also, the two studies used session structures that would be expected to produce similar levels of practice and fatigue on CCMT performance. Specifically, regarding practice, in both studies participants had no prior exposure to the cars but had first completed a CFMT-style test using faces (either the original CFMT or a new alternate version using different faces [CFMT–Australian; McKone et al., in press]) and were, therefore, somewhat familiar with the general format of the test (which was the same for the CCMT and CFMT). In both studies, there was little reason to expect fatigue to influence CCMT scores: In Study 1, the CCMT was administered straight after the CFMT-style test (which takes only 10–15 min to complete) in the first session of testing; and in Study 2, the CCMT was administered as the first experimental task in the second session of testing (the CFMT had been administered in the first session on the previous day). There were no significant differences in CCMT scores between the samples for Study 1 (N = 78, 49 female) and Study 2 (N = 75, 44 female), for either the mean [Study 1, M = 53.59 items correct, Study 2, M = 52.76; $t(151) = 0.61$, $p = .54$, $d = .10$] or the variance [Study 1, SD = 8.68; Study 2, SD = 8.00; Levene's test of equality of variances, $F(1, 151) < 1$]. Therefore, both samples were collapsed and treated as a single sample for all subsequent analyses.

Cambridge Face Memory Test

The CFMT was run using the standard procedure (Fig. 1; Duchaine & Nakayama, 2006). All faces were male and without hair. Face stimuli were approximately 60 × 40 mm (average across individual faces and viewpoints), which, at the viewing distance of 60 cm, subtended 5.7° × 3.8° of visual angle. Scoring of the CFMT is as for the CCMT—that is, in the learn phase, items correct out of 18; in the novel phase, items correct out of 30; in the noise phase, items correct out of 24; for the total score, the number of items correct out of 72.

Car expertise measures (Study 2 only)

In Study 2, participants completed two questions regarding car interest and expertise at the beginning of the session. Both of these used a 9-point Likert-type scale.

Self-reported car interest (self_rated_interest) The first question asked participants to state their "level of interest in cars," with 1 indicating "no interest at all" and 9 indicating "extremely interested."

Self-reported car knowledge (self_rated_knowledge) The second question asked participants to state their "knowledge of car makes and models from approximately 1985–2005," with 1 indicating "no knowledge at all" and 9 indicating "extremely knowledgeable." The rationale for the vintages 1985–2005 was that cars in the CCMT were in the style of this period (i.e., expert knowledge of, say, 1950s cars, would be less likely to be relevant to CCMT performance).

Objective test of preexperimental car expertise (car_knowledge) This was a multiple-choice test of knowledge of real car makes and models from approximately 1985–2005. The format of all questions was the same (see Fig. 4 for examples), with a picture of a car being shown to the participant and the

participant being asked to pick, from a list of five options, which make or model of car they thought the picture showed. The cars were approximately 50 × 140 mm, subtending approximately 4.8° × 13.1° of visual angle. Participants were able to take as long as they needed to answer each question and were encouraged to guess if they did not know the answer. For all questions, there were several plausible alternatives.

The cars were all models available in Australia and, thus, likely to have been seen on the roads by our participants. The make and model names on the test slides were those under which the cars were sold in Australia. A few of our participants had lived overseas, and these participants were allowed to provide an alternative name where relevant (e.g., the car sold under the name *Holden Astra* in Australia is sold under the name *Opel Astra* in Europe).

The test included 22 questions in total, with 14 of these asking what make the car was and 8 asking what model the car was. Scoring was percent correct across all items (chance = 20%). Internal reliability of this task was .77 (.78 for males, .77 for females).

Results

Distribution of CCMT: All participants

Figure 5a shows the frequency distribution of scores for the 153 participants included in the norm sample (i.e., combining

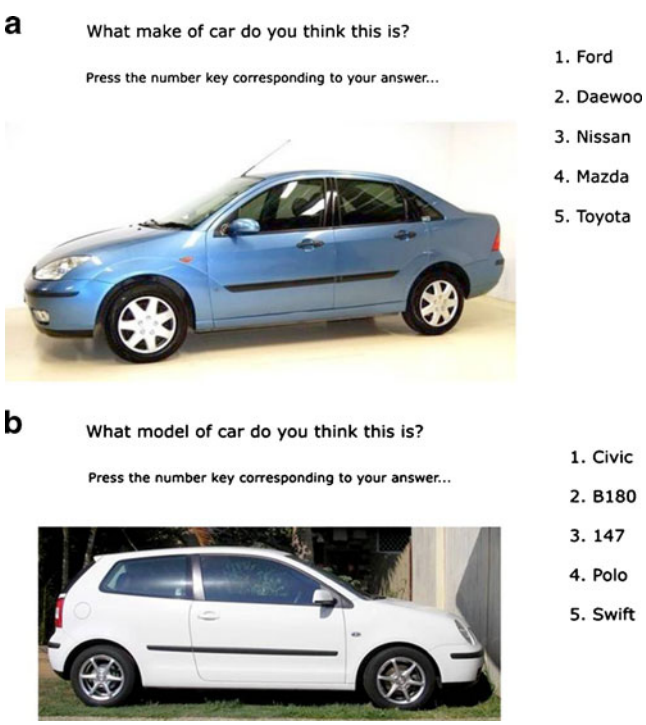


Fig. 4 Example trials from the objective test of pre-experimental knowledge about car makes (**a**) and models (**b**)

males and females). Scores were normally distributed (Shapiro–Wilk statistic = .98, $df = 153$, $p = .07$) and, importantly, showed no evidence of skew (skew = .11, $SE = .20$, $z = 0.56$, $p = .58$). This indicates that task difficulty of the CCMT was set appropriately (i.e., the test produced neither ceiling nor floor effects). Mean accuracy for the full sample was 53.18 items correct out of 72 (73.86%), with $SD = 8.33$ items. This suggests that the test should allow sufficient room for atypically poor performers to be identified below the norm-sample distribution (chance[1] score of 24 items corresponds to a z-score of −3.5), although this still requires empirical investigation using clinical populations.

Note that the three poorest performing participants in Fig. 5 (scores of 33, 33, and 34) were retained in the sample for computation of norms. This was justified on the basis that there was no evidence that their low scores had resulted from a general lack of attention to laboratory tasks or from having a "bad day." Specifically, for all three cases, performance on other tasks completed in the same session as the CCMT was comfortably within the normal range.[2]

Table 1 shows summary statistics (in raw scores; see Table 2 in the Appendix for values as percentage correct). Considering performance for each stage separately, accuracy in the learn phase approached ceiling (as is desired for this phase; i.e., a high mean of 90% correct, together with a smaller *SD* than in other stages). Mean performance on the more difficult novel and noise phases was lower, and, as was expected, the primary differentiation between participants derived from these stages (i.e., large *SD*s; see Table 1). Note that mean performance on the CCMT car task (74%) is slightly lower than that for the CFMT face task (76% in our present sample; 80% in Duchaine & Nakayama, 2006), meaning that raw scores on the two tasks (e.g., from a clinical patient) are not directly comparable.

Sex differences on the CCMT

Importantly, the means and standard deviations reported above for the full sample should not be used as norms for the test, because the CCMT displayed a very substantial sex difference in mean accuracy. Males ($M = 57.43$ items, or 79.8% correct; $n = 60$) performed significantly better than females ($M = 50.44$ items, or 70.1%; $n = 93$), $t(151) = 5.54$, $p < .001$, $d =$

[1] Note that this does not imply that people with object agnosia should necessarily score at chance level on the CCMT.

[2] For the participants who scored 33 (both from Study 1), one had z-score = −1.05 and the other $z = -1.59$ on the CFMT–Australian. For the participant who scored 34 (from Study 2), her discrimination of small changes in height of the vertical bar in T-shapes (procedure described in Experiment 3 of Susilo, McKone, & Edwards, 2010) was good and stable, with slopes of psychometric curves as steep as most others and R^2 of approximately .9 for fits to these curves.

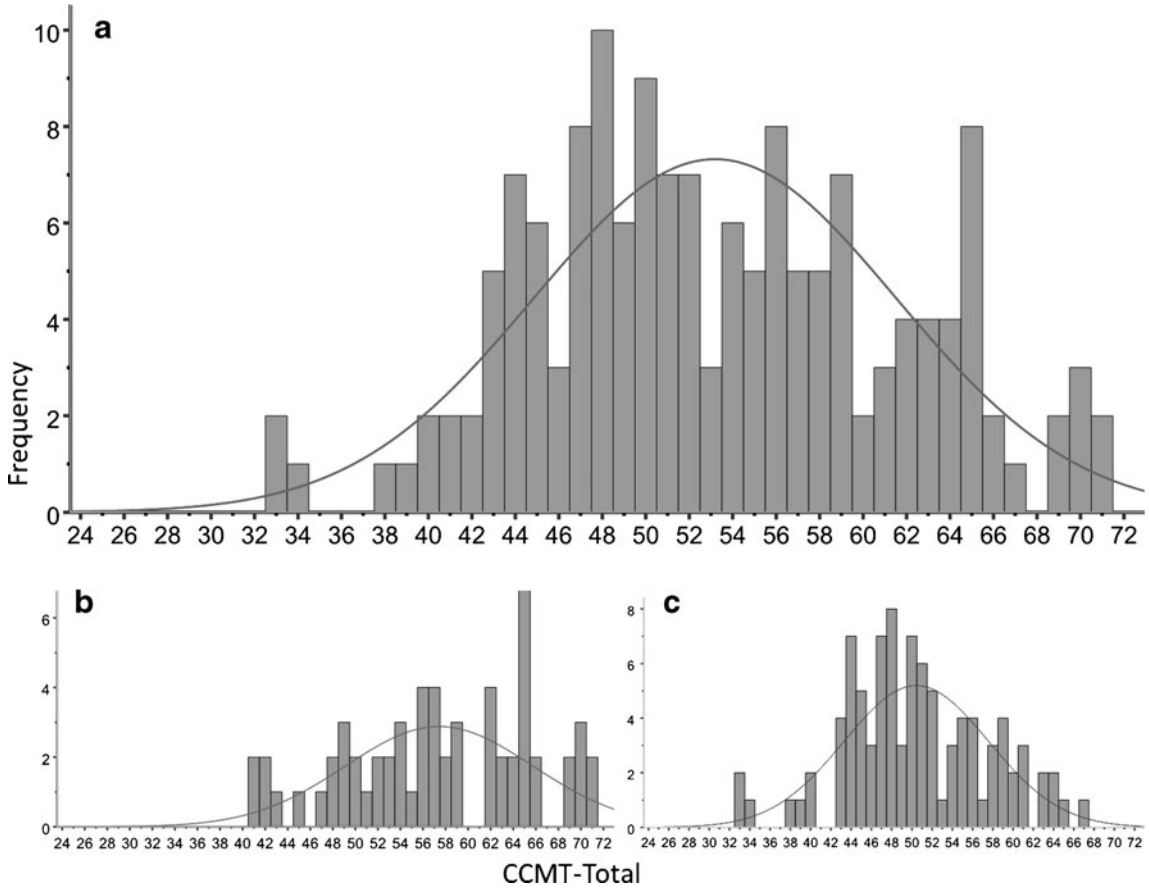


Fig. 5 (**a**) Frequency distribution of CCMT scores for the whole sample (N = 153). Chance level performance = 24, scale maximum = 72. (**b**) Frequency distribution for males (n = 60). (**c**)Frequency distribution for females (n = 93)

.92. There were no significant sex differences in variance [males, SD = 8.31 items; females, SD = 7.15; Levene's test, $F(1, 151) = 2.83$, $p = .09$]. Figure 5b, c show the frequency distribution for males and females separately. There was no significant deviation from normality for either males (Shapiro–Wilk statistic = .96, df = 60, p = .07) or females (Shapiro–Wilk statistic = .98, $df = 93$, $p = .25$). Skew was also not significant for males (statistic = −.24, SE = .31, $z = -.78$, $p = .44$) or females (statistic = .04, SE = .25, $z = .17$, $p = .86$), meaning that there were no ceiling or floor effects. While several males scored close to ceiling, none scored 100%. Females' performance was worse than males', although there may still be sufficient scope to detect atypically poor performance (chance value of 24 corresponds to $z = -3.7$). However, in case this is not sufficient, Table 1 also provides female norms for the combined learn and noise phases (the phases with highest mean accuracy): CCMT_learn + noise in females was normally distributed (Shapiro–Wilk statistic = .98, $df = 93$, $p = .11$, skew statistic = −.39, SE = .25, $z = -1.6$, $p = .11$), and chance level performance (14/42) corresponds to $z = -4.5$.

The primary implications of these results are that (1) the task is valuable for use with both male and female participants, with both sexes showing a distribution of scores suitable for individual differences studies; (2) where the aim is to use the task for diagnosis purposes, separate norms must be used for males and females (and, for females, users may want to consider using the learn-plus-noise norms); and (3) correspondingly, a prosopagnosic or other case-study individual being screened for object agnosia on the CCMT must be compared with an own-sex control sample.

Test reliability

Reliability was measured as internal consistency via Cronbach's alpha, which assesses split-half reliability for all possible splits.[3] Reliability values are shown in Table 1. For the full sample, reliability for CCMT_total was high (α = .84). Moreover, potentially consistent with the sex difference in mean performance on the task, reliability was lower for female participants (α = .75) than for male participants (α = .86). This sex difference in reliability approached statistical

[3] Note that internal reliability is a more appropriate measure than test–retest for memory tasks. This is because performance improves with practice on reexposure to the same items, and the practice effect could differ between individuals (Wilmer, Germine, Loken, et al., 2010).

Table 1 Cambridge Car Memory Test: Observed score range, means, standard deviations, and reliabilities (Cronbach's alpha) in young adults

	Observed Range		Males and Females (N=153)			Females (n=93)			Males (n=60)		
	Min.	Max.	M	SD	α	M	SD	α	M	SD	α
Age	18	32	20.63	2.88	–	20.70	2.73	–	20.52	3.12	–
CCMT_learn	9	18	16.14	1.77	.53	15.73	1.91	.53	16.78	1.30	.37
CCMT_novel	9	30	19.46	4.85	.76	17.91	4.12	.64	21.87	4.93	.81
CCMT_noise	8	24	17.58	3.32	.62	16.80	3.19	.55	18.78	3.16	.64
CCMT_novel+noise	21	54	37.04	7.36	.82	34.71	6.29	.73	40.65	7.48	.85
CCMT_learn+noise	20	41	–	–	–	32.53	4.09	.62	–	–	–
CCMT_total	33	71	53.18	8.33	.84	50.44	7.15	.75	57.43	8.31	.86

1. Maximum scores for each phase are the following: CCMT_learn (18 items correct), CCMT_novel (30), CCMT_noise (24), CCMT_novel + noise (54), CCMT_learn + noise (42), CCMT_total (72); chance performance (i.e., guessing) is always one third 1/3 of maximum (3AFC task). 2. Values converted to percentages are provided in the Appendix. 3. CCMT_learn + noise norms are provided for females only, in case, due to the poorer mean performance of females on the full test, there is insufficient room to detect atypically poor performance when using CCMT_total.

significance (z = 1.89, p = .06). Therefore, individual scores for females should be interpreted somewhat more cautiously than those for males.

We also assessed whether the scale could be improved by removing individual items. This was not the case. There were no items for which removal increased reliability above the .84 for the full scale. We also examined the correlation between each item and the overall scale. In the novel and noise phases, all individual items were positively correlated with the total, suggesting that all of these items contribute to the test's sensitivity; mean item-total correlations were novel phase, r = .27 (SD = .11), and noise phase, r = .22 (SD = .10). In the learn phase, the mean item-total correlation was lower (learn phase, r = .17, SD = .10), and some learn phase items had negative correlations with the total. This is unsurprising given the ceiling effect on performance on this section.

Combined-stage measure of across-view car recognition

Another question of some theoretical interest was the extent to which the different stages of the CCMT assess overlapping or distinct processes. This can be addressed by the size of the correlation between different stages in comparison with the upper bound correlation determined from individual-stage reliability (see Table 1; note that reliability for individual stages was lower than that for the total, as would be expected given the reduced number of trials; the upper-bound correlation is calculated as $r_{12} = \sqrt{\alpha_1} \times \sqrt{\alpha_2}$). If correlations are high relative to upper bound, this argues that stages tap similar processes. Where correlations are low, this argues that stages tap different processes.

Results showed a very strong relationship between the novel and noise stages, with the observed correlation (CCMT_novel and CCMT_noise, N = 153, r = .61, p < .001) nearly at the upper bound (r = .69). This argues that these two stages tapped essentially identical processes.[4] Both of these stages assess recognition of cars in new viewpoints and images. To assist readers who may specifically require a test of object recognition ability across view change, Table 1 presents norm statistics for these two stages combined (CCMT_novel + noise) to increase reliability (total of 54 trials; α = .82).

Test validity: Dissociations between cars and faces

Supporting the CCMT as measuring object (car) memory rather than general visual memory shared by both faces and cars, we present three results arguing that the CCMT and CFMT tapped largely different, rather than overlapping, processes.

Differential sex effects for cars and faces We have shown that males outperformed females on the CCMT (a 10.1% male advantage—i.e., 80.2% vs. 70.1%). In contrast, for faces, previous studies showed either no sex differences for male face stimuli (Lewin & Herlitz, 2002; McKelvie, Standing, St. Jean, & Law, 1993; McKone et al., in press) or a small difference in favor of females (e.g., a 3.8% female advantage in Bowles et al., 2009, and a 3.5% female advantage in Duchaine & Nakayama, 2006). In the present sample, the female advantage for faces (females, n = 93, M = 55.22, SD = 8.38; males, n = 60, M = 54.82, SD = 9.83) was only 0.6%, and this difference was not significant, $t(140)$ = 0.21, p = .83. A two-way ANOVA revealed an interaction between stimulus class (faces, cars) and sex, showing that the pattern

[4] Correlations between the learn and the later stages were significant but were lower relative to upper bound (learn with novel, N = 153, r = .46, p < .001, upper bound r = .63; learn with noise, N = 153, r = .37, p < .001, upper bound r = .57). This result is difficult to interpret: The lowered correlations could be an artifact of the moderate ceiling effect present in the learn stage.

of differences for the CCMT (i.e., a 10.1% male advantage) was significantly different from that for the CFMT (i.e., a 0.6% female advantage), $F(1, 140) = 21.81$, $p < .001$, $\eta^2 = .14$. This supports the conclusion that the two tests do not measure the same process(es).

Low correlation between the CCMT and CFMT We also examined the extent to which the CCMT and CFMT use overlapping or distinct processes, using the standard psychometric logic that two tasks correlate to the extent that they share overlapping processes. Correlations between the CCMT and CFMT were moderate but significant (CCMT_total with CFMT_total: all participants, $n = 142$, $r = .37$, $p < .001$, upper-bound[5] $r = .86$; females only, $n = 93$, $r = .42$, $p < .001$, upper-bound $r = .81$; males only, $n = 49$, $r = .41$, $p = .004$, upper-bound $r = .87$). The results demonstrate that car memory and face memory relied to some extent on overlapping processes (i.e., there is a significant correlation). However, comparison of the size of the observed correlations between CCMT and CFMT to the upper-bound values indicates that car and face memory also have substantial components that are nonoverlapping. That is, despite the fact that the face and car tasks are identical in structure and stimulus format—and thus place the same demands on general factors such as memory, attention, need to generalize across views, and so forth—the correlation between the CCMT and CFMT is well below upper bound, presumably reflecting different *perceptual* components tapped by cars and faces.

Low correlation between the CCMT and CFMT even in those participants with the most expertise in cars An early fMRI study (Gauthier et al., 2000) suggested that car expertise was associated with learning to use face-selective brain regions for within-class discrimination of cars. This idea implies that, although novices use different processes for faces and cars, car experts use common processes for the two stimulus classes. In contrast, recent studies have failed to find any special association of expertise with face-selective areas when additionally examining other brain regions; instead, the typical finding is that expertise with an object class induces increased activation in a very wide range of areas and that the activation increase with expertise is larger in non-face-selective areas (including the lateral occipital complex; V1; nonvisual areas) than in face-selective areas (Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005; Moore, Cohen, & Ranganath, 2006; Op de Beeck, Baker, DiCarlo, & Kanwisher, 2006; Rhodes, Byatt, Michie, & Puce, 2004; Yue, Tjan, & Biederman, 2006; and specifically for cars, Harel et al., 2010, and Jiang et al., 2007). These newer results imply that processes used for faces and cars would remain distinct even in car experts.

We examined this issue empirically for our data set. Although it is not known whether any of the individuals in our sample would qualify as true car "experts," we examined the prediction of the Gauthier, Skudlarski, Gore, and Anderson (2000) theory that individuals who have higher *preexperimental* expertise with cars (i.e., higher scores on knowledge of car makes and models) should show stronger correlation between the CCMT and CFMT (i.e., more overlapping processes) than individuals with lower preexperimental expertise (i.e., little or no overlap in processes). There was no support for this idea. First, a multiple regression predicting the CCMT on the basis of the CFMT, car_knowledge (objective behavioral test of makes and models), and the interaction of the two variables found that the interaction term was not significant (CFMT * car_knowledge, $B = .95$, $SE = .81$), $t(71) = 1.17$, $p = .25$. This indicates no significant change in the CCMT–CFMT correlation as a function of car expertise. Second, we performed subgroup analyses for a "relative expert" group (most knowledgeable one third of the sample; specifically, car_knowledge score > 50). The correlation between the CCMT and CFMT in this subgroup was still low ($r = .28$, $n = 25$) and no higher than in the rest of the sample ($r = .29$, $n = 50$, with car_knowledge scores < 50). These results also held separately for males (experts, $r = .25$, $n = 10$; novices, $r = .38$, $n = 21$) and females (experts, $r = .13$, $n = 10$; novices, $r = .36$, $n = 34$). These results argue that individuals knowledgeable about cars still use primarily independent rather than overlapping processes for cars and faces. This supports the validity of using the CCMT in these participants.

Effects of car interest and expertise on CCMT performance

For the whole sample, the CCMT_total was significantly correlated with all our measures of preexperimental car interest and expertise: self_rated_interest ($r = .36$, $n = 76$, $p = .002$), self_rated_knowledge ($r = .49$, $n = 76$, $p < .001$), and car_knowledge ($r = .54$, $n = 75$, $p < .001$). We also examined the correlations separately for each sex. Within males, all correlations were significant (CCMT_total with self_rated_interest, $r = .37$, $n = 31$, $p = .04$; CCMT_total with self_rated_knowledge, $r = .63$, $n = 31$, $p < .001$; CCMT_total with car_knowledge, $r = .54$, $n = 31$, $p = .002$). For females, car_knowledge correlated significantly with CCMT_total ($r = .49$, $n = 44$, $p = .001$). Interestingly, for females, neither of the self-report measures correlated significantly with the CCMT (CCMT_total with self_rated_interest, $r = .18$, $n = 45$, $p = .24$; CCMT_total with self_rated_knowledge, $r = .27$, $n = 45$, $p =$

[5] Based on reliabilities from this sample (CCMT, $N = 153$) and Bowles et al. (2009; CFMT, $n = 124$ young adults).

.07). This suggests that females may be less aware of their objective knowledge or less confident in their objective ability than males.

We also explored whether the full-sample correlation between self_rated_interest and CCMT might be operating directly (e.g., because greater interest in cars causes increased attention during the CCMT) or indirectly via an intermediate effect on knowledge of makes and models (i.e., greater interest in cars causes one to develop greater car knowledge, and it is the increased knowledge that benefits CCMT performance). A test of the corresponding mediation model (using SPSS syntax provided by Preacher & Hayes, 2004) supported the latter hypothesis (see Fig. 6). The initially significant effect of self_rated_interest on CCMT performance, $t(73) = 3.39$, $p = .001$, was reduced to nonsignificance, $t(73) = 1.65$, $p = .10$, by adding car_knowledge as a mediating variable. And the indirect effect of self_rated_interest via car_knowledge was significant ($z = 2.86$, $p = .004$). Thus, the association between car interest and CCMT observed in this study appears to have been mediated largely by car expertise.

Can differences in car interest/expertise explain the sex difference in CCMT means? Stereotypically, males are considered more interested in, and knowledgeable about, cars than are females. One question of interest, therefore, is whether the sex difference in mean CCMT performance could be explained by sex differences in preexperimental car interest and/or expertise. Perhaps surprisingly, results provided little support for this hypothesis and also clearly indicated a strong direct effect of sex on CCMT independently of the other factors. These conclusions were derived from three types of analyses.

First, we examined sex differences on the means of the car interest and expertise measures. Only if a male advantage is actually present in our sample on these tasks would it be feasible to suggest that they could be an origin of the sex difference on CCMT. All interest and expertise measures showed a trend toward a male advantage. This was significant for self_rated_interest (males, $M = 4.32$, $SD = 1.85$, $n = 31$; females, $M = 3.24$, $SD = 1.69$, $n = 45$), $t(74) = 2.63$, $p = .01$, $d = .61$, but recall that we have shown above that interest per se is not a direct predictor of CCMT performance. For knowledge of makes and models, the male advantage was *not* significant on either self_rated_knowledge (males, $M = 3.00$, $SD = 1.65$, $n = 31$; females, $M = 2.53$, $SD = 1.22$, $n = 45$), $t(74) = 1.43$, $p = .16$, $d = .33$, or car_knowledge[6] (males, $M = 46.4\%$ correct, $SD = 21.4$, $n = 31$; females, $M = 38.0\%$, $SD = 20.0$, $n = 44$), $t(73) = 1.74$, $p = .09$, $d = .41$. Although we would not want to rule out a significant male advantage in car knowledge emerging with a larger sample, the fact that it was nonsignificant here suggests that it is an unlikely explanation of the large, highly significant male advantage seen on the CCMT for the same subset of 75 participants who completed the car knowledge test (males: $n = 31$, $M = 56.61$, $SD = 8.43$; females: $n = 44$, $M = 49.91$, $SD = 6.46$), $t(73) = 3.90$, $p < .001$, $d = .91$.

Second, we used mediation analyses to test for direct versus indirect effects of sex on CCMT. The results are shown in Fig. 7. There were no significant indirect effects of sex via self_rated_interest (Fig. 7a), via self_rated_knowledge (Fig. 7b), or via car_knowledge (Fig. 7c). At the same time, the direct effect of sex on CCMT remained clearly significant in all three mediation models (all ps < .01).

Third, we used multiple regression analyses to ask whether sex still had an independent effect on CCMT once all three interest/expertise variables were accounted for. The four predictors (sex, self_rated_interest, self_rated_knowledge, and car_knowledge) were entered simultaneously in the model as predictors of CCMT. The results (see Fig. 8) showed no unique variance explained by self_rated_interest (consistent with our Fig. 6 mediation analysis) but significant unique variance (i.e., significant semipartial correlation with CCMT) attributable to each other factor, including, most relevantly, sex.

Overall, these results argue that sex had a clear independent effect on CCMT performance, over and above any sex difference than can be attributed to a male advantage on car interest and/or expertise. Indeed, the evidence that the male advantage on CCMT could even be *partially* attributed to male advantages on the other variables was weak at best (nonsignificant trend for male advantage on car expertise, and no significant indirect effects of sex on CCMT operating via any of the three interest/expertise variables).

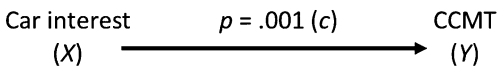
A formula for adjusting CCMT scores for preexperimental expertise

Under some circumstances, a researcher may wish to have access to CCMT scores that are adjusted for preexperimental car expertise (i.e., CCMT independent of car knowledge). For example, in an individual-differences study examining overlap in processes for different object classes, one might wish to know the extent to which memory for cars, independent of car knowledge, correlates with memory for some other nonface object class (potentially also independent of preexperimental knowledge of that class).

We provide a formula for expertise-predicted CCMT scores, using as predictors car_knowledge (i.e., score from our objective knowledge of car makes and models test expressed as a percentage), self_rated_knowledge, and sex,

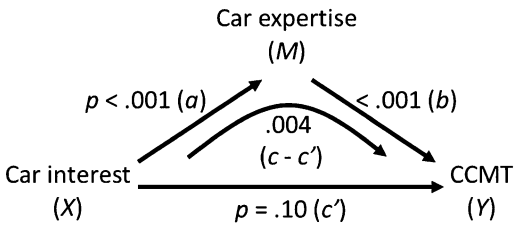
[6] Nor did variances differ significantly between the male and female samples: Levene's test, $F(1, 73) < 1$, $p = .574$.

a Model 1: Direct effect of car interest on CCMT performance



Path	B	SE	t	p
c	1.63	.48	3.39	.001

b Model 2: Includes indirect path via car expertise



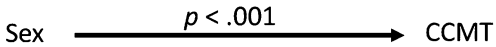
a	4.74	1.23	3.85	< .001
b	.18	.04	4.40	< .001
c'	.78	.47	1.65	.10

Sobel test for indirect effect

	Value	SE	z	p
c - c'	.85	.30	2.86	.004

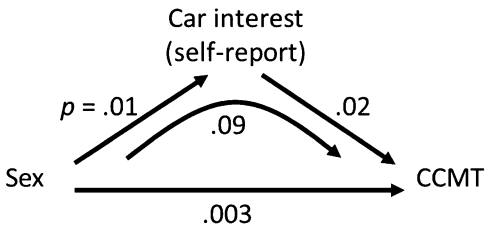
Fig. 6 (**a**) Raw bivariate association between car interest and CCMT performance. (**b**) Results of mediation model testing whether this correlation could be operating indirectly via car expertise. The indirect effect (curved arrow) is strongly significant, while the direct effect (straight line at the bottom of the triangle plot) is not significant

a Model 1: Direct effect of sex on CCMT performance



Path	B	SE	t	p
c	-3.30	.85	-3.86	< .001

b Model 2: Indirect path via self-rated car interest

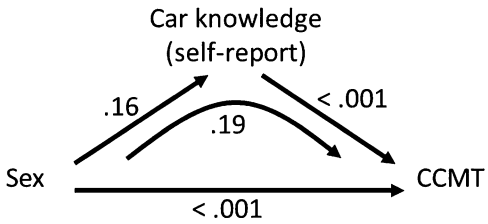


a	-.54	.21	-2.63	.01
b	1.13	.47	2.42	.02
c'	-2.68	.86	-3.11	.003

Sobel test for indirect effect

	Value	SE	z	p
c - c'	-.61	.36	-1.71	.09

c Model 3: Indirect path via self-rated car knowledge

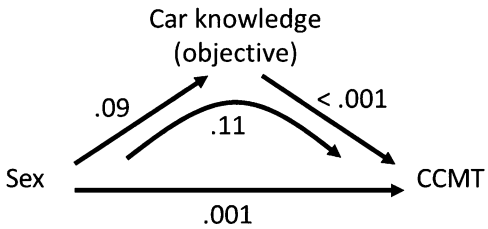


a	-.23	.16	-1.42	.16
b	2.44	.54	4.56	< .001
c'	-2.73	.77	-3.55	< .001

Sobel test for indirect effect

	Value	SE	z	p
c - c'	-.57	.43	-1.33	.19

d Model 4: Indirect path via objective car knowledge



a	-4.20	2.42	-1.74	.09
b	.18	.04	5.09	< .001
c'	-2.58	.76	-3.41	.001

Sobel test for indirect effect

	Value	SE	z	p
c - c'	-.77	.48	-1.62	.11

Fig. 7 Mediation models investigating whether the male advantage on CCMT (**a**) could be mediated by: car interest (**b**); self-report of knowledge of real car makes and models (**c**); or objective ability to identify makes and models (**d**). In all cases, the indirect effect of sex of CCMT (curved arrow) is non-significant, while the direct effect (straight line at the bottom of each triangle plot) is strongly significant. Thus, sex appears to predict CCMT performance independently of pre-experimental interest/expertise in cars. **Notes.** Values on each path in the diagrams are *p*-values. The paths a, b, c' and (c-c') are as defined in Fig. 6

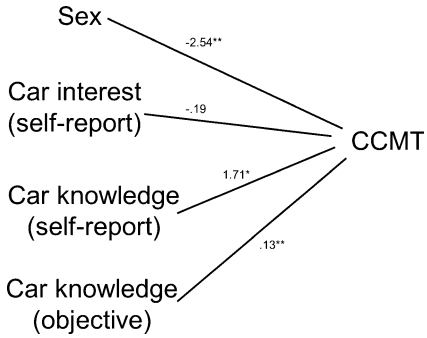


Fig. 8 Multiple regression analysis entering the four variables on the left- hand side of the figure as independent variables, and CCMT as the outcome variable. Path weights are semi-partial correlation coefficients, representing the unique portion of CCMT variance attributable to each IV. Results show that sex predicts CCMT performance independently of car interest and knowledge (i.e., the semi-partial correlation between sex and CCMT is significant) $*p < .05$. $**p < .01$

but not self_rated_interest (a nonsignificant predictor; see Fig. 8). This formula is

$$\begin{aligned}\mathrm{CCMT}_{\mathrm{predicted}} = 48.72 &+ [.14*\mathrm{car_knowledge}] \\ &+ [1.57*\mathrm{self_rated_knowledge}] \\ &+ [-3.6 * \mathrm{sex}],\end{aligned}$$

where sex is coded as 1 = male, 2 = female, and the standard deviation of the residuals is

$$SD_{\mathrm{resid}} = 6.54.$$

Note that, in deriving this formula using multiple regression, we first checked that (1) there were no significant interactions between variables in a multiple regression model that included interactions and (2) the standard deviation of the residuals was constant across levels of expertise.

Discussion

The CCMT provides a valuable test of within-class object recognition performance. In the introduction, we outlined several desirable characteristics for a nonface comparison task to be used in conjunction with the widely used CFMT face task. The CCMT meets these requirements. First, it is closely matched to the CFMT in format, structure, general cognitive requirements (i.e., test length, number of to-be-learned items, etc.), and basic stimulus properties (particularly the requirement for recognition across view and lighting changes). Second, it has high internal reliability (α = .84). Third, our results demonstrate that the CCMT measures a theoretically separate construct from that measured by the CFMT, showing only a moderate correlation between CCMT and CFMT performance (including in both car novices and those more knowledgeable about cars). Finally, the CCMT has a score range that is potentially useful for detecting clinical impairments in object recognition ability (chance score $z = -4.02$ for males, $z = -3.7$ females; $z = -4.5$ for females on the CCMT_learn + noise subtest) and is certainly appropriate for investigating individual differences within the normal population.

Our results have a number of practical and theoretical implications.

Practical implications

Using the CCMT to isolate the face-specific component of visual memory There are a number of theoretical settings in which researchers may wish to isolate the face-specific component of visual memory. Testing both the CFMT face task and the matched CCMT car task in the same participant(s) offers the opportunity to achieve this.

One way to remove variance in face memory scores due to general visual memory is to take multiple regression (Cohen, Cohen, West, & Aiken, 2003) or ANCOVA (Tabachnick & Fidell, 2007) approaches. In the context of an individual-differences study, CCMT scores would be included as a second predictor (i.e., a suppressor variable) of CFMT performance, removing variance associated with general memory ability and, thus, potentially revealing a clearer relationship with the primary theoretical predictor (e.g., the extent to which fixations on the eye region predict *face* memory performance). In studies using group differences, entering CCMT as a covariate using ANCOVA would achieve the same aim (e.g., testing whether a group with autism spectrum disorder has poorer *face* memory than does a control group).

Another way to remove variance, which is common to both CCMT and CFMT scores, is to use a *difference* score, calculated by subtracting the participant's CCMT score from her/his CFMT score. The shared cognitive components common to both tests will be absent from the difference measure. This type of approach has been used previously with theoretical benefit by Zhu et al. (2010), who investigated heritability of face recognition ability by comparing intraclass correlations within groups of monozygotic (MZ) and dizygotic (DZ) twins: While intraclass correlations for tests of either same-image face recognition alone or same-image object recognition alone did not differ significantly between MZ and DZ twins, a faces-minus-objects difference score did differ between the groups (indicating heritability of face-specific memory). Given the internal reliability of α = .88 for CFMT (Bowles et al., 2009) and α = .84 for CCMT (present study) and the correlation between them of r = .37 (present study), the reliability of the CFMT-minus-CCMT difference score[7] is

[7] The formula for calculating the reliability of a difference score is as follows: $r_{\mathrm{diff}} = [\frac{1}{2}(r_{11} + r_{22}) - r_{12}] / (1 - r_{12})$, where r_{11} is the reliability of the first measure, r_{22} is the reliability of the second measure, and r_{12} is the correlation between the first and second measures (Kaplan & Saccuzzo, 2005).

α = .78. These reliabilities compare favorably with those for the face and house tasks used by Zhu et al.

Sex differences, norms, and sex matching Our results showed a large advantage for males, as compared with females, on CCMT performance. Where the test is used in a neuropsychological setting, this implies that an individual's score (e.g., from a potential object agnosic or prosopagnosic individual) should not be evaluated with reference to norms obtained from mixed-sex samples. Instead, females should be compared with a female control sample, and likewise, males should be compared with a male control sample.

The sex difference also implies that, in group-based studies comparing car experts with car novices, it is important to match sex across groups. This is particularly the case because the present study argues for an effect of sex on car memory that is *independent* of the effects of preexperimental expertise.

The sex difference on the CCMT leads to some practical inconvenience. However, we note that choosing a stimulus class other than cars would not necessarily avoid the problem. There is a very limited number of real-world object classes for which exemplars can be selected that even approximately share a first-order configuration (e.g., houses do not). Of these, cars are not the only class likely to have sex differences in interest (e.g., for horses, females stereotypically show a greater level of interest than do males).

Converting CCMT reliability to a 95% CI The internal reliability of a task can be used to calculate a 95% CI around an individual's score on the task. That is, higher reliability corresponds to lower measurement error and, thus, a narrower 95% CI. Using the formula provided by Ley (1972), which takes as input the task's mean, standard deviation, and reliability, an individuals' raw score on the CCMT has a 95% CI of ±7 items correct for females and ±6 for males (wider for females due to lower reliability for females). These values are of particular relevance in neuropsychological settings; for example, an individual with a raw score falling in a "clinical" range on the CCMT (typically defined as the lowest 2% of the population) can, in fact, be *reliably* concluded to have performance this poor only if the *upper bound* of his or her 95% CI falls below the clinical cutoff.

Applicability of the present norms The present article provides specific norms for males and females on the CCMT. To which populations can these norms be taken as applicable? The key issues are the age of participants and the potential effects of testing country and/or era on familiarity with the style of cars.

Regarding age, our CCMT norms are derived from 18- to 32-year-olds. The norms will not apply to individuals younger than this range, since within-class discrimination object memory improves substantially across childhood and adolescence (e.g., Crookes & McKone, 2009). For populations older than 32 years, we are not aware of any published systematic investigation of the effect of aging on within-class discrimination memory for nonface objects. However, two studies (Bowles et al., 2009; Germine, Duchaine, & Nakayama, 2010) have looked at CFMT performance across the lifespan, with both finding that performance begins to drop noticeably below age 20 performance levels at about 50 years of age. Given many other findings that memory declines with aging, we would not recommend using the present norms for participants older than approximately 50 years; whether the norms can be used for participants between 32 and 50 years is unknown. We also note that for elderly participants, the control performance might drop too low for there to be room to diagnose clinical impairments; under these circumstances, a potential solution is to improve performance by running the learn phase twice (for both the case and the age-matched control groups).

Another important point concerns the applicability of our norms in terms of observer familiarity with the types of makes and models of cars used in the CCMT. Makes and models of cars vary across different countries, but in fact, the basic styles are fairly consistent across the world. That is, the styles of cars used in the CCMT would be unlikely to confer much of an advantage to observers in, say, the U.S. over those in China. Moreover, while styles of cars change with time, these changes have been fairly slow within the last few decades, such that cars made today would not look particularly strange to someone who was familiar only with cars up to and including those made in the late 1980s. The cars used in the CCMT resemble those made within the last 20 years or so, and therefore it is likely that—barring any sudden drastic changes in car design—the CCMT cars will look reasonably familiar to most people who have been exposed to modern cars for the next decade or so at least.

Interpretation and appropriateness of using expertise-adjusted norms Our results have provided a regression equation for predicting CCMT scores on the basis of sex and level of car expertise. Although this may be useful in some settings, we caution against automatic application of this formula to calculate "norms" for the CCMT in neuropsychological studies.

To illustrate some of the issues, consider the following case involving a participant whose within-class discrimination of cars we wish to test (e.g., a prosopagnosic being tested for "face-pure" vs. "object-general" deficits). Imagine that this person (1) shows a CCMT score in the normal range with respect to the *full-sample* own-sex norms (see Table 1) but that (2) his or her CCMT score is more than 2 *SD*s poorer than the *expertise-predicted* CCMT score (i.e.,

poorer than predicted CCMT *minus* 2 * SD_{resid}), because (3) the person has high preexperimental expertise with cars. In this case, it is necessary to decide which set of results to believe: Is this person impaired at car recognition, as suggested by finding 2, or not impaired at car recognition, as suggested by findings 1 and 3?

In some situations, there can be a sensible resolution of this apparent conflict. For example, if the person is an *acquired* case of prosopagnosia, with recent brain injury, the results imply that premorbid visual knowledge of car makes and models has been retained but the person cannot use this premorbid knowledge normally to assist learning of new cars. In a case of *developmental* prosopagnosia, however, the use of expertise-adjusted norms to claim that someone highly knowledgeable about the visual appearance of cars is "impaired" at car discrimination would seem nonsensical: If they were genuinely impaired at discriminating cars, how could they have learned to tell apart makes and models so successfully? We thus suggest that, in developmental cases, our full-sample norms, rather than expertise-adjusted norms, would typically be the most appropriate.

Implications of our findings for testing object classes other than cars Our results for the CCMT, using cars, provide an important proof of the concept that the CFMT-style format can be successfully adapted for nonface objects. We have shown that the CCMT loses little in internal reliability, as compared with the CFMT, while still requiring only 10–15 min to administer. This argues that it is worth the effort for researchers to develop CFMT-style tests for objects other than cars.

No single object class makes a perfect control for faces, and it would be ideal to eventually have available CFMT-style tests for multiple different classes. In the diagnosis of object agnosia, it is necessary to test patients using more than one class of object. Similarly, the conclusion that a prosopagnosic individual is a "pure" case of prosopagnosia (i.e., impaired at faces and not objects) would always be strongest if the individual shows normal within-class discrimination memory for a series of other object classes (e.g., cars, houses, office chairs, flowers, etc.). Moreover, in normal observers, we have reported a dissociation (low correlation) here between *one* object class (cars) and faces, but this does not prove that processing of objects per se differs from the processing of faces on CFMT-style tests: It could be that the correlation between two different nonface object classes would be as low as between faces and each object class.

Also of potential relevance to other object classes are our findings regarding the effects of preexperimental expertise. Cars are far from the only object class on which individuals will differ in their level of preexperimental familiarity. Houses of the type used in a given experiment might be common in some world locations, but not in others. Individuals who work in offices are likely to be more familiar with real-world office chairs than are those who work in retail or on construction sites (thus making office chair memory potentially associated with socio-economic status). Individuals differ in their familiarity with dogs and horses. And so on. We expect that researchers would readily accept the theoretical idea that preexperimental familiarity should improve performance on learning and remembering novel exemplars for that class. Yet preexperimental expertise in the object class has typically not been measured in previous neuropsychological or individual-difference studies. For example, while neuropsychological assessments typically take into consideration factors such as age, background, and so forth, it has not been standard to compare an individual with acquired brain injury with a control sample of other individuals with the same level of premorbid knowledge of the object class (see Barton, Hanif, & Ashraf, 2009, for an exception). Our present results argue that there can be value in assessing preexperimental expertise for the object class being tested, and we have discussed the pros and cons of using expertise-adjusted "norms" in different cases.

Theoretical implications

By what mechanisms do car interest and expertise exert their influence on car memory? A priori, readers might have expected that car interest would improve CCMT performance at least partly because interest would increase attention to cars *during the test itself*. Interestingly, our results did not support this interpretation. Indeed, a mediation model implied the apparent effect of interest operated *only* via preexperimental knowledge of real-world makes and models of cars: That is, interest causes people to increase their *pre*experimental knowledge of cars, and this, in turn, causes better CCMT performance in the laboratory.

Considering the likely mechanism by which preexperimental knowledge of makes and models improves CCMT performance, one difference between individuals with more and less car expertise is their familiarity with the aspects of cars that are most informative for delineating makes and models. Knowledge of relevant attributes could improve memory performance on the CCMT in two ways. First, it could direct attention to the most discriminating attributes of the CCMT cars, allowing participants to allocate the available learning time appropriately. That is, participants with more expert knowledge would be more likely to attend to the attributes of the stimuli that would usefully assist subsequent recognition. Second, it could assist individuals with more expert car knowledge to create a richer and more meaningful verbal description of the CCMT stimuli, thus enhancing their memory performance via

verbal mediation (Paivio, 1969). There was anecdotal evidence that participants attempted to use such verbal mediation strategies, even where these would not be successful. For instance, one participant with very little car knowledge, upon seeing the first test three-alternative forced choice (3AFC) item, exclaimed, "oh, it appears my description 'silver and bubbly' won't be very useful!" (because all three cars in the 3AFC test display could, to a car novice, reasonably be described as "silver and bubbly"). A final way in which preexperimental knowledge of car makes and models could improve CCMT performance is via knowledge of specific makes and models that appear in the CCMT. Some of the CCMT items are derived from real makes and models. Some participants stated that they recognized some of the models used in the CCMT and, moreover, that they believed that this knowledge aided their performance on the task (e.g., via remembering the model name).

Theoretical origin of the sex difference We found a large sex difference in mean CCMT performance, favoring males. In terms of the explanation of this effect, our results show that males have more car interest, that interest in cars predicts better car knowledge, and that knowledge of cars in turn predicts CCMT performance. However, our results also argue that it is unlikely that this is the only reason underlying the male advantage. This is because (1) the trend toward a male advantage on the car *knowledge* measures was not significant, (2) sex was an *independent* predictor of CCMT performance in a multiple regression, and (3) indeed, sex was in fact a *stronger* independent predictor of CCMT performance than was self-reported car interest, self-reported car knowledge, or objective car knowledge.

Therefore, it appears that there was an effect of sex per se, over and above any effect of interest or expertise (at least as far as these factors were assessed by our measures). The origin of this sex effect is not yet clear. One speculation is that it could arise from the typical male advantage in mental rotation (Delgado & Prieto, 1996; Halpern, 2004; Halpern & Wright, 1996). The CCMT requires participants to recognize cars that are presented in novel viewpoints, and this ability might potentially benefit from mental rotation skill. Consistent with this interpretation, an absence of mental rotation demands may have contributed to a previous lack of a male advantage in a car old–new test that required recognition of identical images (Duchaine & Nakayama, 2005). In that study, in a sample of consisting primarily of graduate students between 21 and 32 years of age, the mean *d'* for 10 men was 2.67 (*SD* = 0.60), while it was 2.91 (*SD* = 0.60) for 8 women. Similarly, in a sample drawn from the community at large, 35 to 45 years of age, average *d'* was 2.13 (*SD* = 0.75) for 10 men and 2.02 (*SD* = 0.80) for 13 women.

If differences in mental rotation ability account for some of the sex difference on the CCMT, an interesting implication would be that this same mental rotation process would apparently not be applied when *faces* are recognized across views (which shows, if anything, a reverse trend of a slight female advantage).

How overlapping are the processes tapped by face and object processing? In psychometrics, the correlation between two variables relative to upper bound is taken to assess the extent to which those two variables tap the same versus different processes. Previously, the correlation between the CFMT and memory for nonface visual stimuli had been reported to be $r = .26$, using an Abstract Art Test (Wilmer, Germine, Chabris, et al., 2010). However, the Abstract Art Test was only partially similar in format to the CFMT: Like the CFMT, each test trial was 3AFC, and the test required within-class discrimination; but, unlike the CFMT, there was a different number of trials, there was only one learning image per target item, the different exemplars did not share a common first-order configuration (Diamond & Carey, 1986), and the test assessed only same-image recognition rather than recognition of new images across view and lighting change.

Results from the CCMT provide the first correlational test of the extent to which the CFMT shares overlapping processes with object recognition when the general task requirements are much more closely matched. Our finding was that, despite this matching, the CFMT face task still was correlated only modestly with the CCMT car task. The correlation ($r = .37$, $n = 142$) was perhaps slightly higher than that previously found for the CFMT with the Abstract Art Test ($r = .26$, $n = 3{,}044$), but this was not a significant difference, $z = 1.41$, $p = .16$. Moreover, the correlation between the CFMT and the CCMT was still well below the upper-bound correlation ($r = .86$). These results are consistent with findings that the CCMT also shows a low correlation with the CFMT–Australian face task ($r = .21$, upper-bound $r = .86$; McKone et al., in press). The low CCMT correlations with face tasks contrast with much higher correlations found between different face sets in CFMT format tests ($r = .61$ for correlation of CFMT with CFMT–Australian; McKone et al., in press). The low face–car association is of theoretical importance in that it indicates that face and car performance tapped processes that were, to a large extent, distinct. Furthermore, given the similarity in format, structure, and so forth, the fact that the correlation is well below the upper bound argues that there are nontrivial components of performance on these two tests that are underpinned by distinct *perceptual* processes. The present study did not assess the nature of these perceptual differences. However, two possibilities are holistic/configural processing for faces but not objects

(e.g., Diamond & Carey, 1986; Robbins & McKone, 2007; Tanaka & Farah, 1993; Yin, 1969) and/or differential reliance on mental rotation.

Finally, of theoretical interest is our finding that the low correlation between the CFMT and CCMT did not increase as individuals become more knowledgeable about cars. This implies that the perceptual processes tapped by faces and cars are no more overlapping in (relative) car experts than they are in car novices. This result is inconsistent with the fMRI results of Gauthier and co-workers (2000) and is consistent with more recent fMRI studies (e.g., Calvo-Merino et al., 2005; Harel et al., 2010; Moore et al., 2006; Op de Beeck et al., 2006; Rhodes et al., 2004; Yue et al., 2006; see also the behavioral evidence in McKone, Kanwisher, & Duchaine, 2007).

Acknowledgements We thank Mary Broughton for testing some participants in Study 1. This research was funded by Australian Research Council Grant DP0984558 to E.M. and M.E. and by ESRC Grant RES-061-23-0400 to B.D. These funding bodies had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the manuscript for publication. Copies of tasks are available to researchers and clinicians from bradley.c.duchaine@dartmouth.edu for the CCMT and from elinor.mckone@anu.edu.au for the car makes and models knowledge test.

Appendix

Table 2 Cambridge Car Memory Test: Observed score range, means, standard deviations, and reliabilities (Cronbach's alpha) in young adults, expressed as percentage correct

	Observed Range		Males and Females (N=153)			Females (n=93)			Males (n=60)		
	Min.	Max.	M	SD	α	M	SD	α	M	SD	α
Age	18	32	20.63	2.88	–	20.70	2.73	–	20.52	3.12	–
CCMT_learn	50	100	89.69	9.82	.53	87.39	10.59	.53	93.24	7.24	.37
CCMT_novel	30	100	64.88	16.15	.76	59.70	13.75	.64	72.89	16.43	.81
CCMT_noise	33	100	73.23	13.82	.62	70.00	13.30	.55	78.26	13.18	.64
CCMT_novel+noise	38.89	100	68.59	13.63	.82	64.28	11.65	.73	75.28	13.85	.85
CCMT_learn+noise	47.62	97.62	–	–	–	77.45	9.73	.62	–	–	–
CCMT_total	46	98.61	73.87	11.58	.84	70.06	9.93	.75	79.77	11.55	.86

References

Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *NeuroReport, 10*, 189–194.

Aiken, L. R. (2003). *Psychological testing and assessment* (11th ed.). Boston: Allyn & Bacon.

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Boston: Allyn & Bacon.

Barton, J. J. S., Hanif, H., & Ashraf, S. (2009). Relating visual to verbal semantic knowledge: The evaluation of object recognition in prosopagnosia. *Brain, 132*, 3456–3466.

Bate, S., Haslam, C., Tree, J. J., & Hodgson, T. L. (2008). Evidence of an eye movement-based memory effect in congenital prosopagnosia. *Cortex, 44*, 806–819.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77–100.

Bentin, S., Deouell, L. Y., & Soroker, N. (1999). Selective visual streaming in face recognition: Evidence from developmental prosopagnosia. *NeuroReport, 10*, 823–827.

Bowles, D. C., McKone, E., Dawel, A., Duchaine, B., Palermo, R., Schmalzl, L., et al. (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.

Busigny, T., Graf, M., Mayer, E., & Rossion, B. (2010). Acquired prosopagnosia as a face-specific disorder: Ruling out the general visual similarity account. *Neuropsychologia, 48*, 2051–2067.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah: Erlbaum.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers, 25*, 257–271.

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist, 12*, 482–486.

Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition, 111*, 219–247.

Damasio, A. R. (1985). Prosopagnosia. *Trends in Neurosciences, 8*, 132–135.

Dawel, A. & McKone, E. (2011). *Holistic face perception is associated with individual differences in face recognition ability*. Manuscript in preparation.

De Haan, E. H. F., Young, A. W., & Newcombe, F. (1991). Covert and overt recognition in prosopagnosia. *Brain, 114*, 2575–2591.

DeGutis, J. M., Bentin, S., Robertson, L. C., & D'Esposito, M. (2007). Functional plasticity in ventral temporal cortex following cognitive rehabilitation of a congenital prosopagnosic. *Journal of Cognitive Neuroscience, 19*, 1790–1802.

Delgado, A. R., & Prieto, G. (1996). Sex differences in visuospatial ability: Do performance factors play such an important role? *Memory & Cognition, 24*, 504–510.

Delvenne, J.-F., Seron, X., Coyette, F., & Rossion, B. (2004). Evidence for perceptual deficits in associative visual (prosop) agnosia: A single-case study. *Neuropsychologia, 42*, 597–612.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology. General, 115*, 107–117.

Duchaine, B. C., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience, 17*, 249–261.

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.

Duchaine, B., Yovel, G., Butterworth, E., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology, 23*, 714–747.

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neuroscience, 11*, 349–370.

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6*, 428–432.

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*, 191–197.

Germine, L. T., Duchaine, B., & Nakayama, K. (2010). Where cognitive development and aging meet: Face learning ability peaks after age 30. *Cognition, 118*, 201–210.

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience, 7*, 555–562.

Halpern, D. F. (2004). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current Directions in Psychological Science, 13*, 135–139.

Halpern, D. F., & Wright, T. M. (1996). A process-oriented model of cognitive sex differences. *Learning and Individual Differences, 8*, 3–24.

Harel, A., Gilaie-Dotan, S., Malach, R., & Bentin, S. (2010). Top-down engagement modulates the neural expressions of visual expertise. *Cerebral Cortex, 20*, 2304–2318.

Henke, K., Schweinberger, S. R., Grigo, A., Klos, T., & Sommer, W. (1998). Specificity of face recognition: Recognition of exemplars of non-face objects in prosopagnosia. *Cortex, 34*, 289–296.

Herzmann, G., Danthiir, V., Schacht, A., Sommer, W., & Wilhelm, O. (2008). Toward a comprehensive test battery for face cognition: Assessment of the tasks. *Behavior Research Methods, 40*, 840–857.

Iaria, G., Bogod, N., Fox, C. J., & Barton, J. J. S. (2009). Developmental topographical disorientation: Case one. *Neuropsychologia, 47*, 30–40.

Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., VanMeter, J., & Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron, 53*, 891–903.

Kaplan, R. M., & Saccuzzo, D. P. (2005). *Psychological testing: Principles, applications, and issues* (6th ed.). Belmont: Thomson Wadsworth.

Koldewyn, K., Weigelt, S., & Kanwisher, N. G. (2011, May). *Is face recognition selectively impaired in children with ASD?* Paper presented at the 10th Annual International Meeting for Autism Research, San Diego.

Lewin, C., & Herlitz, A. (2002). Sex differences in face recognition—Women's faces make the difference. *Brain and Cognition, 50*, 121–128.

Ley, P. (1972). *Quantitative aspects of psychological assessment*. London: Duckworth.

McKelvie, S. J., Standing, L., St. Jean, D., & Law, J. (1993). Gender differences in recognition memory for faces and cars: Evidence for the interest hypothesis. *Bulletin of the Psychonomic Society, 31*, 353–421.

McKone, E., Hall, A., Pidcock, M., Palermo, P., Wilkinson, R. B., Rivolta, D., . . . O'Connor, K. B. (in press). Face ethnicity and measurement reliability affect face recognition performance in developmental prosopagnosia: Evidence from the Cambridge Face Memory Test–Australian. *Cognitive Neuropsychology.*

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences, 11*, 8–15.

Moore, C. D., Cohen, M. X., & Ranganath, C. (2006). Neural mechanisms of expert skills in visual working memory. *Journal of Neuroscience, 26*, 11187–11196.

Op de Beeck, H. P., Baker, C. I., DiCarlo, J. J., & Kanwisher, N. G. (2006). Discrimination training alters object representations in human extrastriate cortex. *Journal of Neuroscience, 26*, 13025–13036.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*, 241–263.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*, 717–731.

Rhodes, G., Byatt, G., Michie, P. T., & Puce, A. (2004). Is the fusiform face area specialized for faces, individuation, or expert individuation? *Journal of Cognitive Neuroscience, 16*, 189–203.

Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition, 103*, 34–79.

Rossion, B., & Curran, T. (2010). Visual expertise with pictures of cars correlates with RT magnitude of the car inversion effect. *Perception, 39*, 173–183.

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., et al. (2000). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects: An electrophysiological account of face-specific processes in the human brain. *NeuroReport, 11*, 69–72.

Sergent, J., & Signoret, J.-L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex, 2*, 375–388.

Sporer, S. L., Trinkl, B., & Guberova, E. (2007). Matching faces: Differences in processing speed of out-group faces by different ethnic groups. *Journal of Cross-Cultural Psychology, 38*, 398–412.

Susilo, T., McKone, E., & Edwards, M. (2010). Solving the upside-down puzzle: Why do upright and inverted face aftereffects look alike? *Journal of Vision, 10*(13), 1–16.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon.

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

Wallace, S., Coleman, M., & Bailey, A. (2008). Face and object processing in autism spectrum disorders. *Autism Research, 1*, 43–51.

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

Wilmer, J. B., Germine, L., Loken, E., Guo, X. M., Chatterjee, G., Nakayama, K., et al. (2010). Response to Thomas: Is human face recognition ability entirely genetic? *Proceedings of the National Academy of Sciences, 107*, E101.

Wolf, J. M., Tanaka, J. W., Klaiman, C., Cockburn, J., Herlihy, L., Brown, C., et al. (2008). Specific impairment of face-processing abilities in children with autism spectrum disorder using the Let's Face It! skills battery. *Autism Research, 1*, 329–340.

Woolley, A. W., Gerbasi, M. E., Chabris, C. F., Kosslyn, S. M., & Hackman, J. R. (2008). Bringing in the experts. *Small Group Research, 39*, 352–371.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology. Human Perception and Performance, 81*, 141–145.

Yue, X., Tjan, B. S., & Biederman, I. (2006). What makes faces special? *Vision Research, 46*, 3802–3811.

Zhu, Q., Song, Y., Hu, S., Li, X., Tian, M., Zhen, Z., et al. (2010). Heritability of the specific cognitive ability of face perception. *Current Biology, 20*, 137–142.