Heliyon 9 (2023) e14125

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Self-reported face recognition abilities moderately predict face-learning skills: Evidence from Italian samples

Serena Tagliente [a], Marcello Passarelli [b], Vitiana D'Elia [a], Annalisa Palmisano [a], James D. Dunn [c], Michele Masini [d], Tiziana Lanciano [a], Antonietta Curci [a], Davide Rivolta [a,*]

[a] Department of Education, Psychology and Communication, University of Bari Aldo Moro, Bari, Italy
[b] Institute of Educational Technology, National Research Council, Genova, Italy
[c] School of Psychology, University of New South Wales, (UNSW), Sydney, Australia
[d] Valorizzazione Innovazione Empowerment s.r.l. (VIE), Genova, Italy

ARTICLE INFO

Keywords:
Face
CFMT
PI-20
Metacognition
Face recognition
Prosopagnosia
Individual differences

ABSTRACT

Face Recognition Ability (FRA) varies widely throughout the population. Previous research highlights a positive relationship between self-perceived and objectively measured FRA in the healthy population, suggesting that people do have insight into their FRA. Given that this relationship has not been investigated in Italian samples yet, the main aim of the present work was to develop an Italian translation of the Prosopagnosia Index-20 (PI-20), a self-report measure of FRA, to investigate the relationship between PI-20 performances and an objective assessment given by the Cambridge Face Memory Test Long Form (CFMT+) in the Italian population. A sample of 553 participants filled in the PI-20 Italian version 1 or 2 (PI-20_GE or PI-20_BA) and completed the CFMT+. Results showed a negative correlation between both versions of the Italian PI-20 and CFMT+ scores, meaning that the more self-evaluations were negative, the worse they objectively performed. The same results applied to the extreme limits of the distribution (i.e., 10% of the highest and lowest PI-20 scores). Furthermore, both age and administration order of the tests were predictor variables of CFMT+ scores. Overall, our results suggest that people possess insight, although relatively limited, into their FRA.

1. Introduction

Humans widely range in their Face Recognition Ability (FRA), spanning from individuals who struggle to recognize highly familiar faces (i.e., developmental prosopagnosics - DPs) to individuals showing an extraordinary ability in recognizing newly learned or unfamiliar faces (i.e., super-recognisers – SRs) [1–6]. For clinical and applied purposes it is becoming important to understand the link between objective (i.e., test-based) and subjective (i.e., metacognitive skills) measures of FRA [7] (e.g., to justify the use of self-report questionnaires to identify people with face recognition difficulties).

A test that has been widely adopted to investigate people's subjective experience with FRA is the Prosopagnosia Index (PI-20; [8]), a self-administered questionnaire that requires rating the extent to which a person agrees or disagrees with 20 statements about their everyday experience with others' faces. High PI-20 scores correspond to greater face recognition difficulties, which potentially

* Corresponding author.
E-mail address: davide.rivolta@uniba.it (D. Rivolta).

https://doi.org/10.1016/j.heliyon.2023.e14125
Received 14 February 2022; Received in revised form 16 February 2023; Accepted 21 February 2023
Available online 26 February 2023

correspond to poor recognition performances. Indeed, to test the validity of the PI-20, self-report scores have been correlated with an "objective" task assessing FRA, the Cambridge Face Memory Test (CFMT), which consists of learning and recognizing unfamiliar faces [9]. Results show a negative correlation between PI-20 and CFMT scores [10]. In terms of performance, this implies a positive relationship between self-reported and objective performances (i.e., the lower the self-reported ability, the poorer the actual performance). This has been confirmed even in people with suspected prosopagnosia [8], indicating a reasonable relationship between subjective and objective measures of FRA.

This result has been replicated in different countries, with translated versions of the PI-20 showing consistent relationships with demographic-matched versions of the CFMT. PI-20 versions in Portuguese [11] and Polish [12] correlated negatively with the CFMT. Similarly, the Japanese version of the PI-20 [13] showed moderate correlations with both the standard (i.e., Caucasian) CFMT and the CFMT-Asian version. Estudillo and Wong [14] reported a moderate negative correlation between the CFMT-Chinese and the Mandarin version of the PI-20, but only at the lower and upper end of the FRA distribution (see Table 1 for a summary of the available literature).

Albeit the abovementioned studies provide evidence for a positive association between self-reported and objective metrics of FRA, studies adopting different self-report questionnaires have also failed to provide this evidence [15,16]. The reasons behind these inconsistencies are unclear and might be related to the questionnaire's properties (e.g., reliability) and the adopted methods. Indeed, the effects of demographic variables (e.g., age and gender) and methodological factors (e.g., the administration modality) have not always been addressed or controlled in previous research.

Given the lack of an Italian version of the PI-20, the aims of the current multi-center study were (i) to develop an Italian version of the PI-20, and (ii) to ascertain the relationship between the PI-20 and the CFMT long form (CFMT+) [5] in the Italian population. We took into consideration participants' age and gender since they could mediate the correlation between subjective and objective measures of FRA. Indeed, according to the Own-Age Bias (OAB), individuals are more accurate at remembering faces from their own as compared to a different age group [17]. Furthermore, gender differences have frequently been reported in the field of face recognition [18], highlighting greater accuracy for faces of one's gender [19]. The so-called "own-gender bias" is asymmetrical, with the bias being more commonly reported in women [19]. However, some studies showed that women exhibit better FRA than men also for male faces, thus overcoming the "own-gender bias" [1,20,21].

Most of our data were collected online in 2020 during the Covid-19 pandemic. Despite FRA being previously investigated in laboratory settings (e.g., Ref. [5]), online data collection has recently become more common (e.g. Refs. [22,23]). It remains unclear how correlations between subjective and objective measures of FRA are affected by online vs. lab administration. In our study, two Italian versions of the PI-20 have been independently developed by researchers at the University of Genova (PI-20_GE), and the University of Bari Aldo Moro (PI-20_BA).

2. Method

2.1. Participants

A sample of 553 volunteers (62.7% F), aged between 18 and 72 (27.25 ± 10.28) completed the study. All participants had normal or corrected-to-normal vision. The exclusion criteria were: (1) history of neurological and psychiatric conditions (i.e., diagnoses of

Table 1
Summary table of previous findings relating to PI-20 scores and CFMT scores.

Study	PI-20	Enrollment	Order of the test	Sample					Correlation	Suspected DPs reported
				Size	Mean Age	PI-20	CFMT raw scores	CFMT %		
Shah et al. (2015)	Standard	Local participants database	PI-20 | CFMT	87 (57 F)	28.6 years	38.51 ± 8.87	57.59 ± 9.36	79.98 ± 13	$r = -.68$, $p < .001$	Yes
Gray et al. (2017), (Study 1)	Standard	University students	PI-20 | CFMT	142 (86 F)	29.23 ± 11.91	40.10 ± 9.58	58.07 ± 9.21	80.65 ± 12.79	$r = -.394$, $p < .001$	No
Gray et al. (2017), (Study 2)	Standard	University students	CFMT | PI-20	283 (177 F)	26.64 ± 13.16	41.70 ± 10.10	55.29 ± 9.29	76.80 ± 12.90	$r = -.390$, $p < .001$	No
Ventura et al. (2018)	Portuguese	University students	Not reported	123 (108 F)	20.40 ± 4.35	42.02 ± 9.26	62.06 ± 7.37	86.20 ± 10.24	$r = -.43$, $p < .0001$	No
Marschollek et al. (2019)	Polish	Social media announcements	Not reported	1270 (430 F)	28.3 years	49.6 ± 18	58.10	80.69	$r = -.42$, $p < .0001$	No
Estudillo and Wong (2021)	Mandarin	University students	Randomized	255 (188 F)	21 ± 4.2	47.89 ± 11.78	56.46 ± 8.19	78.42 ± 11.38	$r = -.35$, $p < .001$	No

mild or severe neurological or psychiatric disease); (2) history of stroke; (3) epileptic seizures; (4) brain injury with loss of consciousness; (5) consumption of psychiatric drugs during the last month; (6) drug or alcohol addiction. While 298 participants (64.1% F; mean age: 28.43; SD: 11.22) completed one version of the PI-20 (i.e., PI-20_GE) and the CFMT+, 255 participants (61.2% F; mean age: 25.87, SD: 8.89) completed the other version of the PI-20 (i.e., PI-20_BA) and the CFMT+ (see next section for the description of the questionnaires). 323 participants (65.3% F) completed the tests online (PI-20_GE: $n = 217$; PI-20_BA: $n = 106$) whereas 230 participants (59.1% F) completed the tests in lab settings (PI-20_GE: $n = 81$; PI-20_BA: $n = 149$ (see Table 2) at the University of Bari 'Aldo Moro'. The study was given ethical approval by the Ethics Committee of the Department of Education, Psychology, and Communication of the University of Bari Aldo Moro, and executed according to the Declaration of Helsinki.

2.2. Materials

2.2.1. Prosopagnosia Index-20 (PI-20)

Both versions of the PI-20 comprise 20 items concerning the self-perception of FRA (see appendix for both of the translations). As in the original version [8], people answered through a Likert Scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). Fifteen of the 20 statements are scored positively, whereas five items are reverse scored (strongly agree is scored "1" and strongly disagree is scored "5"). Higher scores on the PI-20 imply worse "self-perceived" FRA. The two Italian PI-20 versions are equivalent, with only some minor phrasing differences. The PI-20_GE was independently translated by four translators, who then compared their versions to obtain a single, agreed-upon translation. The same process was performed for the PI-20_BA. As the two groups of authors joined efforts after starting data collection, both versions were employed to maximize the sample size.

2.2.2. Cambridge face memory task long form (CFMT+) [5]

While the original CFMT [9] is widely used to diagnose prosopagnosia and assesses face recognition in the typical population (e.g., Ref. [20]), a ceiling effect occurs if performed by people with very good FRA. To overcome this problem, Russell et al. [5] developed the CFMT+, which includes an additional section (i.e, making the task more "challenging" than the original CFMT). Participants are exposed to six unfamiliar greyscale target faces and then tested on their recognition of those faces. The task includes two phases: 'study' and 'test'. Each of these phases is repeated through four stages of increasing difficulty; in the first "learn" stage, faces are the same as those saw in the training (score out of 18). In the second "novel" block, faces are seen under novel lighting and viewpoints (score out of 30). Finally, during the third "noise" stage, visual noise is overlaid on all the faces (score out of 24). The fourth section (which is included in the CFMT+ only) includes 30 images varying in pose, emotional expressions, and the amount of information available (faces are fully cropped from the external features, or presented with visible hairstyle and exposed neck; score out of 30). All pictures in the last phase contain visual noise degrading images, and the distractor identities recur more frequently than in the first three, minimizing the difference in familiarity between the studied and distractor faces. The maximum total score of the CFMT+ is 102.

2.3. Procedure

After signing informed consent, all participants were invited to complete the PI-20_GE or PI-20_BA and the CFMT+, online or in a laboratory setting. The administration order of the tests was counterbalanced: half of the sample performed the PI-20_GE or PI-20_BA and then the CFMT+, while the others completed the tests in the opposite order (see Fig. 1).

2.4. Analyses

To ascertain the relation between PI-20 scores and face recognition performance, we computed 0-order Pearson's correlations between PI-20 and both CFMT+ and CFMT scores (where CFMT scores refer to the first three sections of the test, while the CFMT+ includes all four sections). Both CFMT and CFMT+ scores can be derived from CFMT+ administration. To check for the equivalence of CFMT and CFMT+ scores in our sample, we ran a correlation between these scores, which resulted in a very high correlation ($r = 0.96$). Based on this result, only CFMT+ scores were included in the analyses, with CFMT scores available upon request. Multiple linear regressions were performed, with CFMT and CFMT+ scores as outcome variables; PI-20 scores, participants' age, and gender, and

Table 2
Sample's characteristics: size, % of females, mean age and standard deviations of subjects who completed PI-20_GE and PI-20_BA in Online and Lab conditions.

PI-20 Version	Presentation mode	Sample size	F (%)	Means age and range
PI-20_GE	Online	217	67.7	27.10 (18–72)
PI-20_GE	Lab	81	54.3	31.99 (18–63)
PI-20_BA	Online	106	60.3	25.06 (18–59)
PI-20_BA	Lab	149	61.7	26.44 (18–67)
PI-20_GE and PI-20_BA	Online	323	65.3	26.43 (18–72)
PI-20_GE and PI-20_BA	Lab	230	59.1	27.83 (18–67)
PI-20_GE	Online and Lab	298	64.1	28.43 (18–72)
PI-20_BA	Online and Lab	255	61.2	25.87 (18–67)
PI-20_GE and PI-20_BA	Online and Lab	553	62.7	27.25 (18–72)

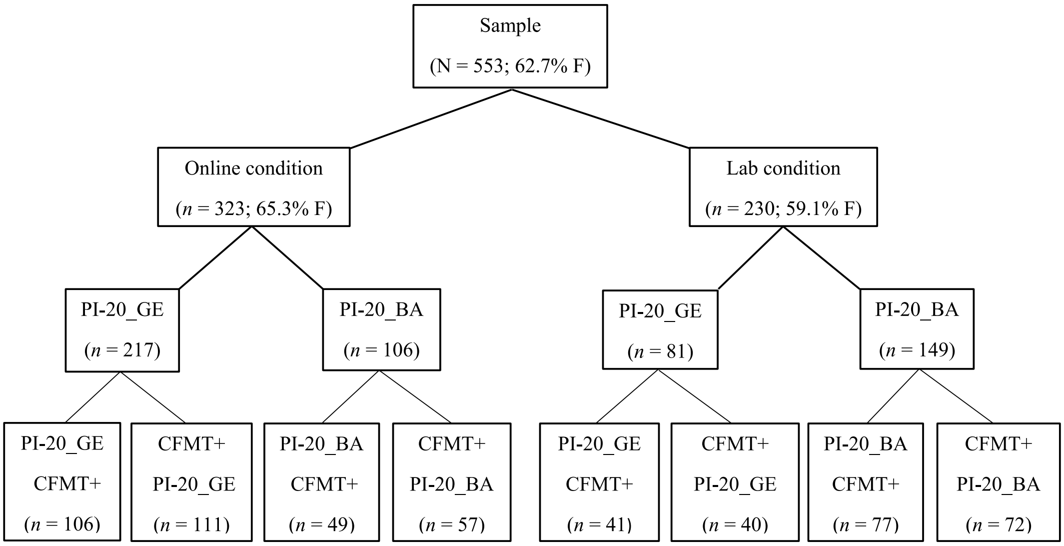

Fig. 1. Administration order of the tests both in online and lab conditions.

administration method (i.e., lab or online data collection) were included as predictors. The models also comprised the age x gender, age x PI-20 scores, and gender x PI-20 scores interactions. Additionally, the order of administration (CFMT first, or PI-20 first) was included as a control variable to check for potential order effects. Lastly, both PI-20 versions and the interactions between these and the

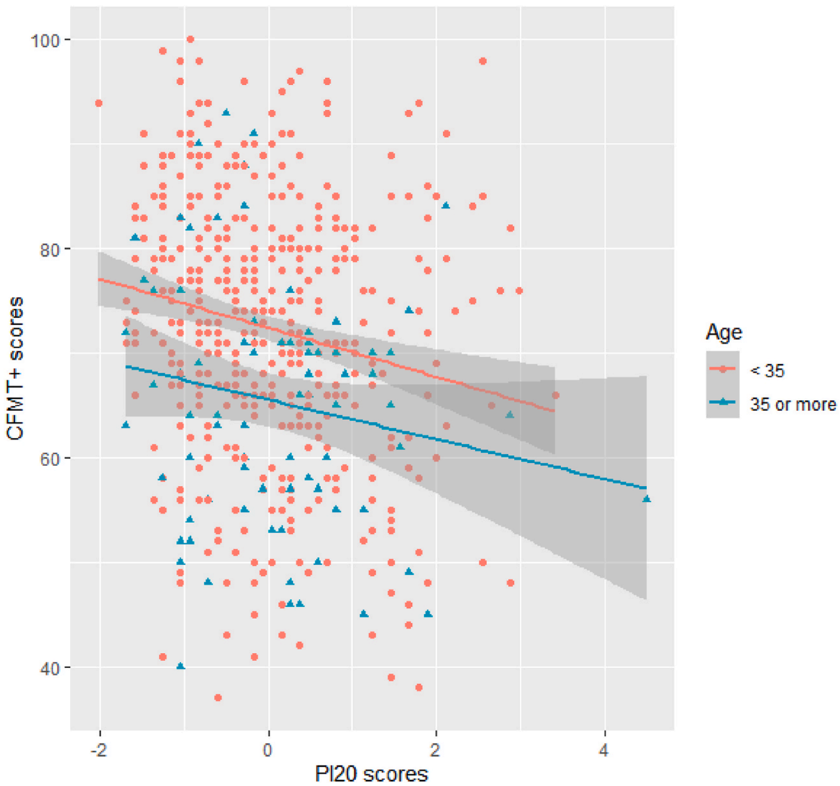

Fig. 2. Scatterplot of the PI-20 and CFMT+ scores distribution (standardized). Negative correlation between PI-20 and CFMT+ scores ($r = -0.18$, $p < .001$). Shades represent upper and lower bound confidence intervals (95%) of the correlation between CFMT+ and PI-20. Y-axis: PI-20 scores; x-axis: CFMT+ scores. Red points: <35 years scores; blue triangle: 35 or more scores. CFMT+ <35: mean = 70.82, SD = ±12.36; PI-20 < 35: mean = 40.59, SD = ± 9.05. CFMT+ ≥35: mean = 63.94, SD = 11.37; PI-20 ≥ 35: mean = 41.27, SD = ±10.07. We choose the values on the x-axis and y-axis looking at the min and max scores performed by participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

PI-20 scores were included in the model as control variables. The questionnaire version, as well as the administration order, were adopted as control variables, but they were not expected to have significant effects. Age was dichotomized to account for non-normality of the distribution (skewness = 1.93, kurtosis = 2.81). As the faces in the CFMT+ stimuli are approximately 25–35 years old, we opted for a cut-off of 35. We divided participants between those closer to the age of the CFMT+ depicted actors (and thus presumably having an advantage due to the "same-age effect") and participants older than them. The cut-off was also based on previous evidence for the decline in CFMT performances around the age of 30 [24]. Moreover, it allows us to assess OAB, as stimuli in the CFMT portray individuals "in their 20s and early 30s" [9]. All analyses were run using RStudio. All data analyzed in this study are available upon request.

3. Results

The 0-order Pearson's correlation reports a small ($r^2 = 0.03$), albeit statistically significant, negative correlation between PI-20 and CFMT/CFMT+ scores ($r = -0.18$, $p < .001$; see Figs. 2 and 3). Pearson's correlations between PI-20 and CFMT scores were split according to gender and age, and no statistically significant difference emerged (see Table 3). The adjusted R^2 for the multiple linear regression models is 0.07. The linear models explain relatively little variance (see Table 4), and results are largely comparable for both CFMT and CFMT+ scorings of the test (i.e., including or excluding the fourth section). The most predictive variable is age, suggesting that individuals over 35 years old have lower FRA. Boxplots of CFMT+ scores according to age category and gender are shown in Figs. 4 and 5. Lastly, the administration order also shows a statistically significant effect: participants who filled in the PI-20 first scored lower at the CFMT/CFMT+ than participants who underwent the opposite administration order. However, correlations between CMFT/CFMT+ and PI-20 scores are similar for both administration orders ($r = -0.19$ for participants who took the CFMT+ first, and $r = -0.20$ for participants who took the PI-20 first; $p < .001$ in both cases). The PI-20 Italian versions are equal in reliability (Cronbach's Alpha PI-20_BA: 0.84; Cronbach's Alpha PI-20_GE: 0.83). The mean score from participants completing the PI-20_BA was 41.99, SD = 9.41 (156 females (mean scores: 41.64 = SD = 9.92) and 99 males (mean scores = 43.53, SD = 8.57), with a mean age of 25.87 (±8.89)). The mean score from participants completing the PI-20_GE was 39.58, SD = 8.88 (191 females (mean scores = 38.80, SD = 8.41) and 107 males (mean scores = 40.97, SD = 9.54), with a mean age of 28.43 (±11.22)).

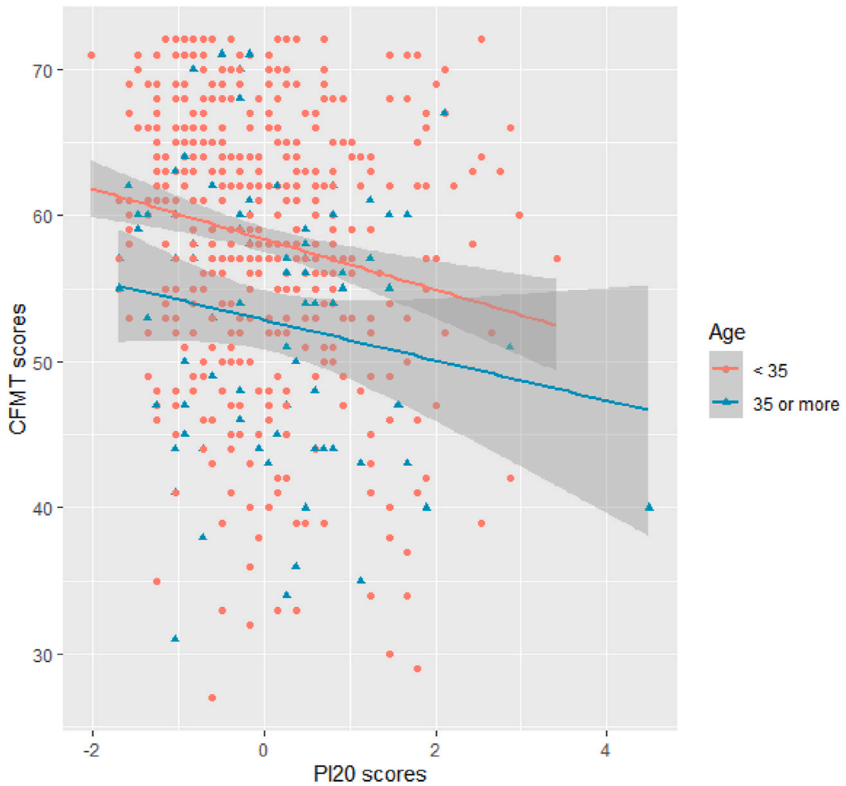

Fig. 3. Scatterplot of the PI-20 and CFMT scores distribution (standardized). Negative correlation between PI-20 and CFMT scores ($r = -0.18$, $p < .001$). Shades represent upper and lower bound confidence intervals (95%) of the correlation between CFMT+ and PI-20. Y-axis: PI-20 scores; x-axis: CFMT scores. Red points: <35 years scores; blue triangle: 35 or more scores. CFMT <35: mean = 79.39, SD = ±13.24; PI-20 < 35: mean = 40.59, SD = ± 9.05. CFMT ≥35: mean = 81.77, SD = ±13.88; PI-20 ≥ 35: mean = 41.27, SD = ±10.07. We choose the values on the x-axis and y-axis looking at the min and max scores performed by participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
Pearson's correlations between PI-20 and CFMT scores, split according to gender and age categories of participants, with 95% confidence intervals.

Gender	Age	Correlation
Women	<35	−.17 [−.28, −.06]
Women	≥35	−.10 [−.37, .19]
Men	<35	−.16 [−.31, −.01]
Men	≥35	−.29 [−.57, .05]
Both	<35	−.18 [−.27, −.09]
Both	≥35	−.17 [−.37, .05]
Women	Any	−.18 [−.28, −.07]
Men	Any	−.17 [−.30, −.03]

Table 4
Results of the multiple linear regression.

Predictor	CFMT	CFMT+
(Intercept)	***β = .23 [.04, 42], t(522) = 2.38, p = .018*** *	***β = .22 [.03, .41], t(522) = 2.27, p = .023*** *
PI-20	*β = -.17 [-.29, -.06],* ***t(522) = -2.95, p = .003*** **	*β = -.18 [-.29, -.06],* **t(522) = -3.05, p = .002** **
Age (>= 35)	*β = -.64 [-.94, -.34],* ***t(522) = -4.18, p < .001*** ***	*β = -.65 [-.94, -.35],* **t(522) = -4.23, p < .001** ***
Gender (Male)	*β = -.17 [-.36, .02], t(522) = -.1.80, p = .073*	*β = -.17 [-.35, .02],* t(522) = −.1.76, p = .080
Method (online data collection)	*β = .03 [-.15, .21], t(522) = .32, p = .748*	*β = .04 [-.13, .22],* t(522) = .46, p = .646
PI-20 × Age	*β = .03 [-.18, .25], t(522) = .30, p = .768*	*β = .03 [-.19, .24],* t(522) = .24, p = .813
PI-20 × Gender	*β = -.01 [-.19, .16], t(522) = -.14, p = .888*	*β = -.00 [-.18, .17],* t(522) = −.04, p = .965
Age x Gender	*β = -.17 [-.22, .71],* ***t(522) = -2.04, p = .042*** *	*β = .28 [-.18, .74],* t(522) = 1.19, p = .236
Administration order (PI-20 first)	*β = -.17 [-.34, -.01],* ***t(522) = -2.95, p = .003*** **	*β = -.17 [-.34, -.01],* **t(522) = -2.12, p = .034** *
Version	*β = -.64 [-.21, .14],* ***t(522) = -4.18, p < .001*** ***	*β = -.03 [-.18, .17],* t(522) = −.32, p = .752

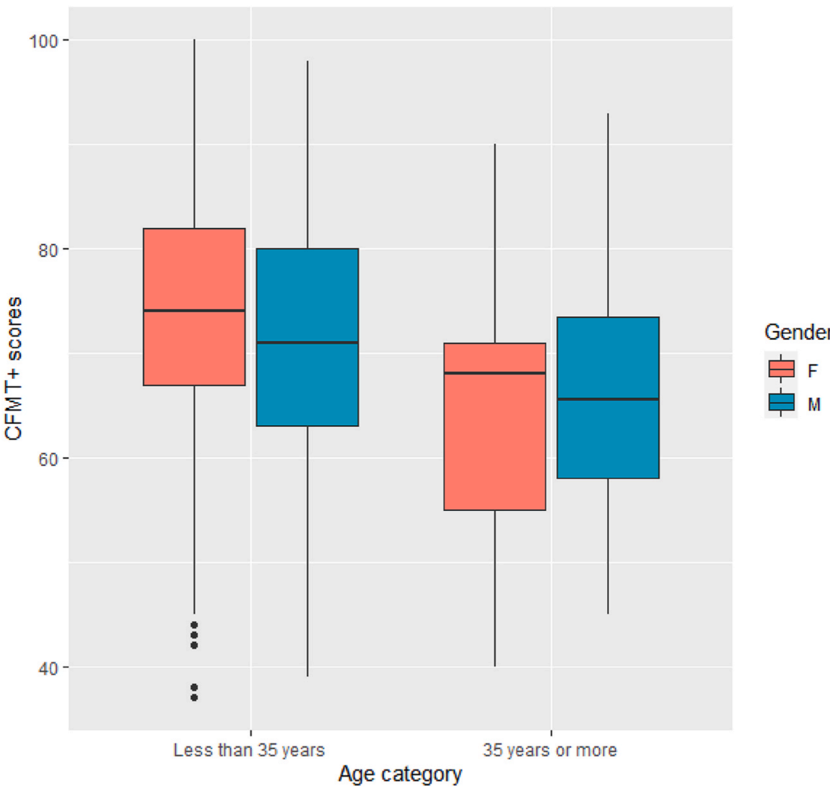

Fig. 4. Boxplots of CFMT+ scores according to age and gender of participants: <35 years (472 participants; 63.35% F); ≥35 years (81 participants; 59.26% F). Y-axis: age category; x-axis: CFMT+ scores. Red rectangle: female scores; blue rectangle: male scores. CFMT+ female <35: mean = 71.75, SD = ± 12.29; CFMT+ male <35: mean 69.21, SD = ± 12.35; CFMT+ female ≥35: mean = 63.44, SD = ± 11.31); CFMT+ male ≥35: mean = 64.68, SD = ± 11.58. We choose the values on the y-axis looking at the min and max scores performed by participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

To further assess the relationship between self-reported and measured recognition ability at the extremes of the distribution, we compared the CFMT/CFMT+ scores between the top and bottom decile of participants according to the PI-20 scores. In both cases, results were statistically significant (CFMT: $t(75.44) = -2.23$, $p = .028$; CFMT+: $t(74.23) = -2.34$, $p = .022$), with participants in the top decile for the PI-20 having lower CFMT/CFMT+ scores than participants in the bottom decile.

4. General discussion

The current study aimed to test whether the PI-20 (Italian versions) exhibits a correlation with a standard test for unfamiliar face recognition, the CFMT+. In line with previous literature [8,14], our results demonstrate a positive relationship between FRA self-evaluation and its objective assessment. The higher participants report issues in everyday experiences with faces, the lower they perform, and vice versa. Although the effect size of our correlation was small, our result suggests that typical individuals exhibit insight into their FRA.

Participants younger than 35 performed better than older participants, which could reflect both the OAB and the decline in FRA associated with aging [24]. Unexpectedly, participants who performed the PI-20 before completing the CFMT+ obtained lower recognition scores (CFMT/CFMT+) than those who filled the tests in the opposite order. Several features could influence the self-perception of FRA. For instance, receiving feedback on an objective performance could inform individuals about their actual FRA level. Bobak et al. [15] developed the Stirling Face Recognition Scale (SFRS) for one's everyday experience in face recognition (predominantly adapted from the PI-20; [8]). They administered SFRS and CFMT+ to 96 students who previously received or did not receive feedback on their performance, with results showing moderate insight into FRA for naïve participants and higher accuracy for those who were informed of their (exceptional) performance. In our study, participants did not receive any feedback on their CFMT+ performance. This implies they were not using any prior knowledge from objective assessments to complete the PI-20. Indeed, self-evaluations of FRA were not affected by the administration order of the tests. On the other hand, although we did not expect any order effect, the administration order was a significant predictor of CFMT+ scores (i.e., participants that completed the PI-20 first showed lower CFMT+ scores). One potential interpretation is that answering questions on personal experiences with FRA (PI-20) before a performance task negatively influences the objective (CFMT+) scores. This might be based, for instance, on performance anxiety, or one's tendency to align his/her performance to self-evaluations. We could argue that the self-evaluation completed first may have driven (i.e., negatively impacted) participants' CFMT+ scores based on the "risk" for discrepancies between perceived and actual abilities (i.e., a self-protection strategy against objective disconfirmations of one's self-evaluations) [25].

Given the recent rise of online testing in research, we believe that tests' administration modality should be considered. Germine et al. [26] found that tests' presentation modality (i.e., online or laboratory) does not influence CFMT performances. Consistently with their work, the online vs. laboratory testing in our study did not affect CFMT+ scores. However, Petersen and Leue [27] demonstrated that CFMT+ online testing results in significantly higher scores than laboratory testing. A possible explanation here is that lab vs. online participants were differentially motivated to participate in the study. Indeed, lab participants were students who participated to gain credit points. We could assume that irrespective of the administration modality, all participants in our sample took part in the study on the basis of a genuine interest in the topic.

Demographic variables could also play a role in FRA. According to the OAB, individuals are more accurate at remembering faces from their own age group as compared to a different one [17], which in turn could affect individuals' insight into their abilities. Germine et al. [24] showed that FRA peaks around the age of 30. Accordingly with this, Susilo et al. [28] reported that ~30 years old participants perform better in the CFMT than ~20 years old participants. This age pattern was also confirmed by a different test format [29]. In line with previous literature, younger people (<35 years old) in our sample were more accurate in CFMT+ than participants older than 35 years, although their degree of FRA self-awareness was the same. This result might be interpreted in light of the *late maturation hypothesis* of face recognition, based on both environmental (e.g., the quantity of faces one is exposed to in daily life based on age) and genetic (i.e., protracted expression of genes) factors.

Further research is needed to investigate which (and to which extent) individuals' features could influence the relationship between objective and subjective abilities in face recognition (for instance, socioemotional variables and personality traits). These variables may interact with age, which represents the strongest CFMT+ scores predictor in our study, and consequently, impact the relationship between objective and subjective measures of FRA.

Our results provide evidence for a correlation between the CFMT+ and the PI-20. Despite the small effect size implying caution in its interpretation, our results indicate that higher face recognition performances correspond to good self-perceived FRA. Importantly, we assessed the difference between CFMT+ scores at the extremes of the distribution (i.e., by comparing the top and bottom percentiles according to the PI-20 scores). In line with the observed correlations, a significant difference in CFMT+ performances emerges at the extreme ends of the PI-20 distribution, showing that participants in the top decile for the PI-20 have lower CFMT+ scores and vice versa.

More in general, one might argue that self-reports and behavioral objective measures assess partially overlapping constructs. However, they are designed for the assessment of different processes. While behavioral measures tap people's maximal performance (i.e., by encouraging people to do their best), self-report measures tap people's evaluation of their typical performance (i.e., how they usually behave) [30]. Moreover, different factors could affect objective performances (e.g., fatigue, motivation, etc.), and subjective reports (e.g., desirability bias) [10]. Moreover, self-evaluations seem to be particularly affected by systematic biases [31].

As for our PI-20 translations, our results show they are equivalent. However, since our translations have not been validated, we cannot provide recommendations on which version should be adopted. However, the adoption of two independent translations providing the same results represents a strength of our work, as compared to analogous studies available for other languages. Future

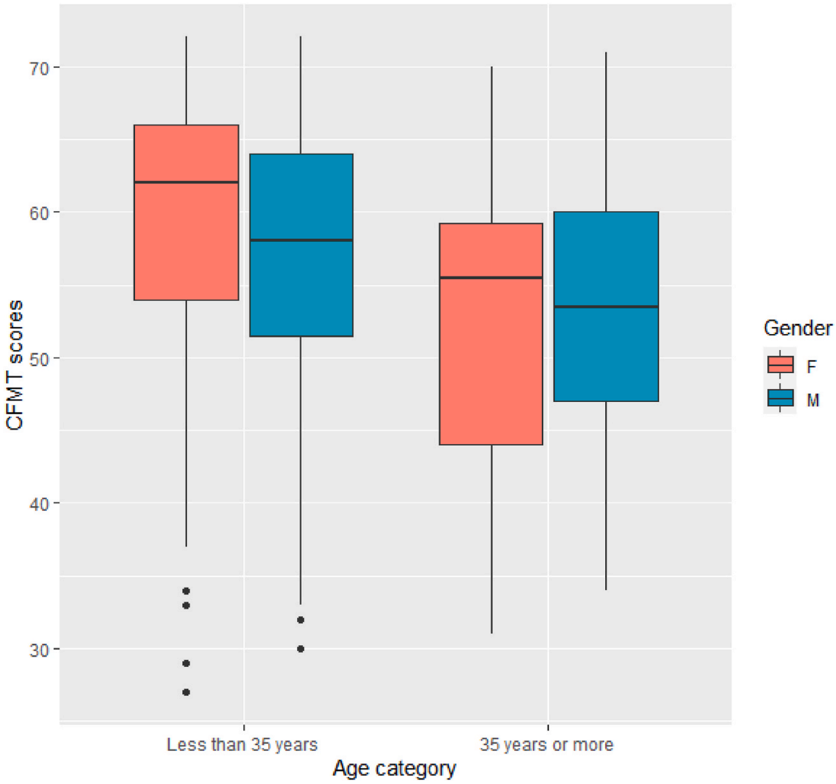

Fig. 5. Boxplots of CFMT scores according to age and gender of participants: <35 years (472 participants; 63.35% F); ≥35 years (81 participants; 59.26% F). Y-axis: age category; x-axis: CFMT+ scores. Red rectangle: female scores; blue rectangle: male scores. CFMT female <35: mean = 81.96, SD = ± 12.99; CFMT male <35: mean = 79.01, SD = ± 13.11; CFMT female ≥35: mean = 72.83, SD = ± 13.41; CFMT male ≥35: mean 73.57, SD = ± 12.45. We choose the values on the y-axis looking at the min and max scores performed by participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

research could validate an Italian version of PI-20 to be used for both clinical and research purposes with Italian native speakers.

One limitation of our study is the lack of potential SR or DP diagnoses. Indeed, this would have required the adoption of multiple tests [32]. Based on previous evidence for DPs and SRs' insight into their FRA, atypical populations' objective performances might differentially correlate with self-evaluation scores than in the typical population [15,16], potentially based on prior knowledge of their objective performances affecting their insight in FRA. Other potential limitations in the present work, as well as in other studies in the field (e.g., Ref. [10]), are represented by the adoption of only one task to assess objective FRA (i.e. a test of unfamiliar faces learning), and the testing modality (i.e., online for the majority of the data). Despite CFMT+ being considered a gold standard in this field [33], face recognition represents a multifaceted process [34]. Future research should replicate our results by adopting different and/or multiple tests (e.g., a task for famous face recognition, as in Ref. [11]), and potentially conducting experiments in controlled laboratory settings.

5. Conclusions

The current study provides evidence for a relationship between self-perceived and objective FRA in Italian samples. Despite the small effect size posing caution for interpretation, our study has some strong points, particularly the large sample size and the consistent results on the extremes of the PI-20 scores' distribution, which allow us to assume that people may have insight into their ability to recognize faces, in line with most of the literature on the typical population [11,12]. Further replication is needed to test our result with different subjective and objective measures of FRA.

6. Production notes

APC waived - waiver code: Heliyon - Contractual payment discounts.

Author contribution statement

Serena Tagliente, Marcello Passarelli, Michele Masini, Tiziana Lanciano, Antonietta Curci: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data.

Vitiana D'Elia, Annalisa Palmisano: Conceived and designed the experiments; Performed the experiments.

James D. Dunn; Conceived and designed the experiments; Performed the experiments; Wrote the paper.

Davide Rivolta: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interest's statement

The authors declare no competing interests.

References

[1] B.C. Duchaine, K. Nakayama, Developmental prosopagnosia: a window to content-specific face processing, Curr. Opin. Neurobiol. 16 (2006) 166–173, https://doi.org/10.1016/j.conb.2006.03.003.
[2] D. Rivolta, A. Woolgar, R. Palermo, M. Butko, L. Schmalzl, M.A. Williams, Multi-voxel pattern analysis (MVPA) reveals abnormal fMRI activity in both the â€œcoreâ€ and â€œextendedâ€ face network in congenital prosopagnosia, Front. Hum. Neurosci. 8 (2014), https://doi.org/10.3389/fnhum.2014.00925.
[3] D. Rivolta, R.P. Lawson, R. Palermo, More than just a problem with faces: altered body perception in a group of congenital prosopagnosics, Q. J. Exp. Psychol. 70 (2017) 276–286, https://doi.org/10.1080/17470218.2016.1174277.
[4] J. Tardif, X. Morin Duchesne, S. Cohan, J. Royer, C. Blais, D. Fiset, B. Duchaine, F. Gosselin, Use of face information varies systematically from developmental prosopagnosics to super-recognizers, Psychol. Sci. 30 (2019) 300–308, https://doi.org/10.1177/0956797618811338.
[5] R. Russell, G. Chatterjee, K. Nakayama, Developmental prosopagnosia and super-recognition: No special role for surface reflectance processing, Neuropsychologia 50 (2012) 334–340, https://doi.org/10.1016/j.neuropsychologia.2011.12.004.
[6] R. Russell, B. Duchaine, K. Nakayama, Super-recognizers: people with extraordinary face recognition ability, Psychon. Bull. Rev. 16 (2009) 252–257, https://doi.org/10.3758/PBR.16.2.252.
[7] K.A. Dalrymple, R. Palermo, Guidelines for studying developmental prosopagnosia in adults and children, WIREs Cogn. Sci. 7 (2016) 73–87, https://doi.org/10.1002/wcs.1374.
[8] P. Shah, A. Gaule, S. Sowden, G. Bird, R. Cook, The 20-item prosopagnosia index (PI20): a self-report instrument for identifying developmental prosopagnosia, R. Soc. Open Sci. 2 (2015), 140343, https://doi.org/10.1098/rsos.140343.
[9] B. Duchaine, K. Nakayama, 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 (2006) 576–585, https://doi.org/10.1016/j.neuropsychologia.2005.07.001.
[10] K.L.H. Gray, G. Bird, R. Cook, Robust associations between the 20-item prosopagnosia index and the Cambridge Face Memory Test in the general population, R. Soc. Open Sci. 4 (2017), 160923, https://doi.org/10.1098/rsos.160923.
[11] P. Ventura, L.A. Livingston, P. Shah, Adults have moderate-to-good insight into their face recognition ability: further validation of the 20-item Prosopagnosia Index in a Portuguese sample, Q. J. Exp. Psychol. 71 (2018) 2677–2679, https://doi.org/10.1177/1747021818765652.
[12] Karol Marschollek, Marta Nowakowska-Kotas, Paweł Marschollek, Julia Marschollek, Gerald Drożdż, Jerzy Drożdż, Sławomir Budrewicz, Developmental prosopagnosia in Poland: an analysis of online-conducted population based study, EMET 6 (2020) 57–64, https://doi.org/10.15503/emet2019.57.64.
[13] S.F. Nakashima, M. Ukezono, R. Sudo, M. Nunoi, S. Kitagami, M. Okubo, R. Toriyama, Y. Morimoto, Y. Takano, Development of a Japanese version of the 20-item prosopagnosia index (PI20-J) and examination of its reliability and validity, J. Pshicho. 90 (2020) 603–613, https://doi.org/10.4992/jjpsy.90.18235.
[14] A.J. Estudillo, H.K. Wong, Associations between self-reported and objective face recognition abilities are only evident in above- and below-average recognisers, PeerJ 9 (2021), e10629, https://doi.org/10.7717/peerj.10629.
[15] A.K. Bobak, V.R. Mileva, P.J. Hancock, Facing the facts: naive participants have only moderate insight into their face recognition and face perception abilities, Q. J. Exp. Psychol. 72 (2019) 872–881, https://doi.org/10.1177/1747021818776145.
[16] R. Palermo, B. Rossion, G. Rhodes, R. Laguesse, T. Tez, B. Hall, A. Albonico, M. Malaspina, R. Daini, J. Irons, S. Al-Janabi, L.C. Taylor, D. Rivolta, E. McKone, Do people have insight into their face recognition abilities? Q. J. Exp. Psychol. 70 (2017) 218–233, https://doi.org/10.1080/17470218.2016.1161058.
[17] J.S. Anastasi, M.G. Rhodes, An own-age bias in face recognition for children and older adults, Psychon. Bull. Rev. 12 (2005) 1043–1047, https://doi.org/10.3758/BF03206441.
[18] A. Palmisano, F. Bossi, C. Barlabà, F. Febbraio, R. Loconte, A. Lupo, M.A. Nitsche, D. Rivolta, Anodal tDCS effects over the left dorsolateral prefrontal cortex (L-DLPFC) on the rating of facial expression: evidence for a gender-specific effect, Heliyon 7 (2021), e08267, https://doi.org/10.1016/j.heliyon.2021.e08267.
[19] A. Herlitz, J. Lovén, Sex differences and the own-gender bias in face recognition: a meta-analytic review, Vis. Cognit. 21 (2013) 1306–1336, https://doi.org/10.1080/13506285.2013.823140.
[20] D.C. Bowles, E. McKone, A. Dawel, B. Duchaine, R. Palermo, L. Schmalzl, D. Rivolta, C.E. Wilson, G. Yovel, Diagnosing prosopagnosia: effects of ageing, sex, and participant–stimulus ethnic match on the Cambridge face memory test and Cambridge face perception test, Cogn. Neuropsychol. 26 (2009) 423–455, https://doi.org/10.1080/02643290903343149.
[21] C. Østergaard Knudsen, K. Winther Rasmussen, C. Gerlach, Gender differences in face recognition: the role of holistic processing, Vis. Cognit. 29 (2021) 379–385, https://doi.org/10.1080/13506285.2021.1930312.
[22] L.P. Satchell, J.P. Davis, E. Julle-Danière, N. Tupper, P. Marshman, Recognising faces but not traits: accurate personality judgment from faces is unrelated to superior face memory, J. Res. Pers. 79 (2019) 49–58, https://doi.org/10.1016/j.jrp.2019.02.002.
[23] J.P. Davis, L.D. Bretfelean, E. Belanova, T. Thompson, Super-recognisers: face recognition performance after variable delay intervals, Appl. Cognit. Psychol. 34 (2020) 1350–1368, https://doi.org/10.1002/acp.3712.
[24] L.T. Germine, B. Duchaine, K. Nakayama, Where cognitive development and aging meet: face learning ability peaks after age 30, Cognition 118 (2011) 201–210, https://doi.org/10.1016/j.cognition.2010.11.002.
[25] M.D. Alicke, C. Sedikides, Self-enhancement and self-protection: what they are and what they do, Eur. Rev. Soc. Psychol. 20 (2009) 1–48, https://doi.org/10.1080/10463280802613866.
[26] L. Germine, K. Nakayama, B.C. Duchaine, C.F. Chabris, G. Chatterjee, J.B. Wilmer, Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments, Psychon. Bull. Rev. 19 (2012) 847–857, https://doi.org/10.3758/s13423-012-0296-9.
[27] L.A. Petersen, A. Leue, Extraordinary face recognition performance in laboratory and online testing, Appl. Cognit. Psychol. 35 (2021) 579–589, https://doi.org/10.1002/acp.3805.

[28] T. Susilo, L. Germine, B. Duchaine, Face recognition ability matures late: evidence from individual differences in young adults, J. Exp. Psychol. Hum. Percept. Perform. 39 (2013) 1212–1217, https://doi.org/10.1037/a0033469.
[29] J.D. Dunn, S. Summersby, A. Towler, J.P. Davis, D. White, UNSW Face Test: a screening tool for super-recognizers, PLoS One 15 (2020), e0241747, https://doi.org/10.1371/journal.pone.0241747.
[30] J. Dang, K.M. King, M. Inzlicht, Why are self-report and behavioral measures weakly correlated? Trends Cognit. Sci. 24 (2020) 267–269, https://doi.org/10.1016/j.tics.2020.01.007.
[31] X. Zhou, R. Jenkins, Dunning–Kruger effects in face perception, Cognition 203 (2020), 104345, https://doi.org/10.1016/j.cognition.2020.104345.
[32] S. Bate, E. Portch, N. Mestry, R.J. Bennetts, Redefining super recognition in the real world: skilled face or person identity recognizers? Br. J. Psychol. 110 (2019) 480–482, https://doi.org/10.1111/bjop.12392.
[33] M. Stantic, R. Brewer, B. Duchaine, M.J. Banissy, S. Bate, T. Susilo, C. Catmur, G. Bird, The Oxford Face Matching Test: a non-biased test of the full range of individual differences in face perception, Behav. Res. 54 (2021) 158–173, https://doi.org/10.3758/s13428-021-01609-2.
[34] A.M. Chernorizov, Y.P. Zinchenko, J. Zhong-qing, A.V. Petrakova, Face cognition in humans: psychophysiological, developmental, and cross-cultural aspects, Psych. Rus. 9 (2016) 37–50, https://doi.org/10.11621/pir.2016.0404.