Neuropsychologia 158 (2021) 107809

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: http://www.elsevier.com/locate/neuropsychologia

Review article

Super-Recognizers – a novel diagnostic framework, 70 cases, and guidelines for future work

Meike Ramon

Applied Face Cognition (AFC) Lab, University of Fribourg, Department of Psychology, Faucigny 2, 1700 Fribourg, Switzerland

ARTICLE INFO

Keywords:
Super-Recognizers
Face processing
Diagnostic criteria
Law enforcement
Legal cognitive neuroscience

ABSTRACT

When you hear the word *Super-Recognizer*, you may think of comic-book-hero-esque agents searching the underground to find people who went missing decades ago. Compared to this fantasy, the reality seems somewhat less exciting. Super-Recognizers (SRs) were initially reported a decade ago as a collateral while developing tests for developmental prosopagnosia. Today, the topic of SRs sparks interest from groups seeking to enhance scientific knowledge, public safety, or their monetary gain. With no immediate consequences of erroneous SR identification, there has been no pressure to establish a clear SR definition. This promotes heterogenous empirical evidence and the proliferation of unsupported claims in the media. Not only is this status quo unfortunate, it stands in opposition to the potential of special populations – both for science and application. SRs are a special population with imminent real-world value that can advance our understanding of brain functioning. To exploit their potential, I propose a needed formal framework for SR diagnosis, and introduce 70 cases identified based hereupon. These cases represent the core of a growing SR cohort, studied in my lab in the course of a long-term, multi-methodological research agenda involving academic and government collaborators. Finally, I provide recommendations for those interested in SR work, and highlight current caveats and future challenges.

1. Introduction

The term “Super-Recognizer” (SR) was first introduced by Russel et al. (2009) who reported four individuals who excelled at unfamiliar face matching (face perception), unfamiliar face memory (face recognition) and identification of celebrities based on their childhood pictures. To date, among several thousands of scientific articles published on face processing, less than 25 address the topic of SRs (Ramon et al., 2019a) (Fig. 1a).

Independently of scientific research on face processing, interest in SRs is growing. The media report about SRs in the context of law enforcement, associations offering memberships and fee-based SR “accreditations”, and companies offering SR-based services are emerging (Ramon et al., 2019a). On an international governmental level, practitioners and law enforcement bodies seek to optimize their operations through SR deployment in occupations were excellent face processing skills are pivotal, such as CCTV surveillance or criminal investigation (Cologne Police, 2016; Ramon, 2019).

This growing interest in SRs offers exciting opportunities for fundamental and applied research, which can advance our understanding of brain functioning and address societal issues, e.g. those concerning facial recognition technology (EU Agency for Fundamental Rights, 2019). Considering its implications, SR research must commit to a standard higher than the status quo. Currently, it lacks fundamental definitions and a consensus regarding diagnostic criteria. Consequently, the term SR means different things to different people – even to those within the same domain. Conceptual precision and clear definitions of SRs’ abilities are necessary to advance our understanding of the phenomenon and to meet practitioners’ expectations.

Emerging from exchanges with peers and practitioners, here I propose a formal diagnostic framework for lab-based SR identification, alongside novel and stringent criteria. I then present 70 SR cases identified using these criteria, and conclude with an outlook of my lab’s ongoing work involving various techniques. Presenting criteria, cases, and a commitment to transparent and consistent reporting of this growing cohort, the goal is to lay the foundation for future SR research, which addresses open questions concerning brain-behavior relationships with real-world implications.

E-mail address: meike.ramon@gmail.com.

https://doi.org/10.1016/j.neuropsychologia.2021.107809
Received 17 February 2021; Accepted 19 February 2021
Available online 2 March 2021

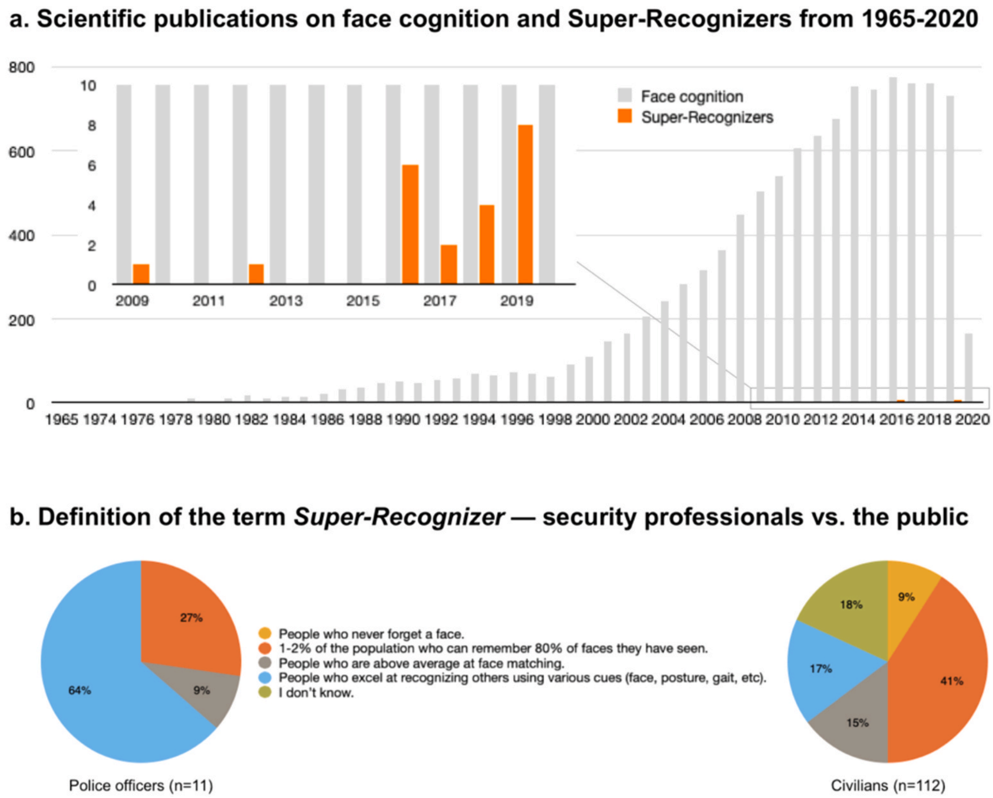

Fig. 1. Research statistics of empirical work conducted on, and variable definition of Super-Recognizersa. The large number of scientific studies published on the topic of face cognition stands in opposition to the very limited knowledge available on SRs. Searching for "((face processing) OR face recognition) OR face identification AND psychology" on pubmed.com on April 9th 2020 yielded 10820 results. **b.** With no strict scientific definition of what constitutes a SR, definitions vary between and among security professionals and the general public. Shown here are responses to the question "What is a Super-Recognizer?" provided by EU police officers attending the second consortium workshop of the European Project SafeCi - Safer Space for Safer Cities (left) and civilians (right). Note that none of the above responses are corect/false; the frist and second options represent empirically unsupported claims.

Box 1
What is a Super-Recognizer? It depends on who you ask

Empirically, SRs were originally described by Russell et al., 2009 as individuals with both superior face perception, recognition and identification. Departing from this seminal paper's standard, most researchers have used *either* above performance for face perception *or* for face recognition as a criterion for face processing superiority (Noyes et al., 2017; Bobak et al., 2016; Phillips et al., 2018). Therefore unsurprisingly, research findings in this emerging field are characterized by a large degree of heterogeneity (Ramon et al., 2019a).

In *security settings and law enforcement*, the term is used to refer to a fundamentally different and varied set of skills. Individuals that excel at *person* recognition or matching, or at detection of suspicious *behavior* have also been referred to as "SRs". Their skill is usually determined without formal testing, based on on-the-job performance relative to professional peers. I have tested three police officers, who were reported to have been identified as SRs by professional peers – those of a former special unit within the London Metropolitan Police (Cologne Police, 2016). Interestingly, while two scored highly on tests of face cognition (JD1 and NY1 listed in Table 1), the third one did not possess any special skills regarding face cognition – his scores did not meet the criteria proposed here. This coincides with his own report: he explained that, as a former undercover officer, he was skilled at detecting *suspicious behavior*.

2. "Super-Recognizer" — one label, multiple meanings

Despite the growing interest in SRs, a fundamental problem exists: there is no formal definition of what actually constitutes a SR. Therefore, unsurprisingly, *the term SR means different things to different people* (Fig. 1b; Box 1). Weak or lacking diagnostic criteria have important implications. From a scientific point of view, they decrease the likelihood of replicable research findings and hinder advancing our understanding of this interesting rare population. In applied settings, well-defined concepts are crucial for practical reasons. They are a necessary precondition for meeting practitioners' expectations regarding personnel selection and placement gains, and for the implementation of evidence-informed practices.

3. Establishing the necessary diagnostic foundation for future work

The majority of studies published since Russell et al.'s (2009) seminal paper have typically used performance on a *single* test to identify someone as a SR (Noyes et al., 2017; Bobak et al., 2016a; Phillips et al., 2018). Without agreement regarding test(s) to use, researchers have employed a range of different ones (Ramon et al., 2019a; Noyes et al., 2017) even within the same study (Phillips et al., 2018). These vary in terms of their psychometric properties, and – most crucially – usually measure *either* perception *or* recognition.

The formal diagnostic framework for SR identification introduced in Box 2 aims to lay the ground for future work and facilitate progress in SR research. I outline its general guidelines, and specific criteria according to which I classify observers as SRs. These incorporate previous recommendations (Bate et al., 2019; Devue 2019; Moreton et al., 2019;

Box 2
Novel diagnostic framework for empirical Super-Recognizer identification

General guideline	Rationale
Superiority in face cognition must be assessed at **various levels of processing** (perception, recognition), using a **range of complementary tests**.	Departing from the original report (Russell et al., 2009), which assessed face perception, recognition and identification, recent studies have relied on single tests scores (e.g., Phillips et al., 2018), or have conflated different subprocesses of face cognition as "face recogniton" (e.g., Noyes et al., 2017).
Only **empirically validated and ubiquitous tests with sufficient sensitivity** should be used for SR identification, i.e. tests that are available to others and do not suffer from ceiling effects.	Previous studies have reported individuals as alleged SR cases based on performance accuracy for the GFMT (e.g., Phillips et al., 2018). This test/measure is not sufficiently calibrated to clear two standard deviations from the mean (Burton et al., 2010) and cannot distinguish healthy controls from individuals suffering from developmental prosopagnosia (White et al., 2017) or acquired prosopagnosia (Fysh & Ramon, in preparation).
Individual SR cases must be reported using their unique anonymous **acronyms and all available information** must be **communicated**.	Consistent cross-referencing is required to enable accumulation of converging evidence across studies and to minimize information loss. This is a standard procedure in neuropsychology, where rare single cases are reported using the same acronym across studies.

Specific diagnostic criteria: Achievement of high performance across a number of independent tests involving different procedures provides a reliable indicator for *genuine* face processing superiority. Following this logic and the above guidelines, an individual may be considered a SR if they exhibit superior performance for at least two of the three currently employed tests according to the cut-offs adopted based on the currently available data (cf. Table 1), placing them within the top 6% of a sample of 181 observers for the three tests CFMT+, YBT-long, and FICST (cf. Stacchi et al., 2020).[11]

Robertson and Bindemann, 2019; Young and Noyes, 2019), and developments in behavioral assessment of and individual differences in face cognition (Stacchi et al., 2020; Fysh et al., 2020). With its stricter criteria – requiring superior performance across a *collection* of demonstratedly difficult tests – this novel framework aims to decrease heterogeneity in observations caused by false positive SR diagnosis. The present guidelines can serve as a theoretically motivated and practically beneficial framework for SR work, which can be optimized in the face of new evidence. Advocating for the need to adopt only published and reproducible methods standardized under laboratory conditions, Box 3 discusses important caveats of mass online testing employed by some researchers.

Multi-level assessment with multiple tests. In line with recent suggestions (Bate et al., 2019; Ramon et al., 2019a; Devue, 2019), a *collection of tests* should be used for SR-identification. Ideally, a test battery should be administered to enable assessment of different skills or subprocesses of face cognition (e.g., perception *and* recognition). Although performance across subprocesses and tests of face cognition are *overall* generally correlated, a growing body of independent evidence demonstrates that performance at the *individual* level is not. That is, while among normal observers it is not uncommon to observe high performance for an *individual test*, such isolated superiority does not translate across a range of tests (Bate et al., 2018; Stacchi et al., 2020; Fysh et al., 2020). Therefore, observers should only be considered a SR if they achieve superior scores consistently, i.e. across *multiple tests of face cognition*.

Test sensitivity matters. Assessment of individuals' face processing abilities requires *tests that are sufficiently sensitive*, i.e. difficult enough to detect superior levels of performance. Consequently, certain tests are inappropriate to this end and should not be used for SR testing. This includes the Glasgow Face Matching Test (GFMT) (Burton et al., 2010; White et al., 2017), which suffers from ceiling effects (Burton et al., 2018) and can be accomplished by highly impaired individuals suffering from developmental prosopagnosia (White et al., 2017). Therefore, despite its past deployment (Phillips et al., 2018), the GFMT should ***not*** be used for SR-identification. The same logic applies to tests that were developed for diagnosis of impairments, such as the short (72 item) version of the Cambridge Face Memory Test (Duchaine and Nakayama, 2006), which e.g. Tardif et al. (2019) used to investigate the relationship between information usage and face cognition ability. Usage of insufficiently sensitive tests will generate findings of questionable validity due to the potential for overinclusion of individuals from the entire population (cf. Phillips et al., 2018).

Table 1
Cross-subject performance means along with percentiles and corresponding cut-offs used per test to identify superior face processing abilities.

Test	n	Mean score ± SD	Percentile	Cut-off (score)
CFMT+	367	65.42 ± 11.79	95th	85
YBT-long	292	8.89 ± 3.98	90th	1
FICST	404	8.25 ± 6.01	92.5th	15

Note that the cut-offs used here may require adjustment as new data becomes available. This is particularly important considering the age groups comprising our samples, which — despite our best efforts — due not reflect the heterogeneity that is characteristic of the population (mean age across cohorts: 30 ± 13; range: 18–80; for full details see DOI 10.17605/OSF.IO/3BUKV).

Commitment to consistent and transparent data reporting. Large and growing cohorts of SRs identified using the same criteria should be studied by independent (adversarial) research groups using complementary procedures. Critically, in order to minimize information loss, cases should be consistently referred to across research groups and studies, and all available evidence should be provided.

Consider the 70 cases presented here in the following sections. They include five SRs reported previously by Bobak et al. (2016a), as well as (Bobak et al., 2016b,c) – where they were, however, reported with *different acronyms*. Consistent referencing of cases is particularly crucial in the context of special populations with limited cases. Here, as in longitudinal studies, accumulation of evidence relies on the use of the same anonymous identifier per individual. A prime example is PS, a rare case of acquired prosopagnosia first described in 2003 (Rossion et al., 2003; Rossion, 2014; Ramon et al., 2016), now reported in over 30 papers, which have contributed significantly to our understanding of healthy and impaired face processing. Committing to transparent and consistent data reporting, I provide additional information for the SR cases introduced here, including participation in studies that are

Box 3
Tests established under controlled laboratory conditions vs. mass online testing

Tests used to identify SRs may be developed and standardized differently. Currently, online experimentation, and consequently access to data, has become trivial. Unsurprisingly, there is a continuously increasing number of tests made available online. However, we do not simply need more tests, or more data – *we need the right ones*. The potential merit of any publicly available online test will depend on the nature of its data quality assurance measures. These should be designed and implemented to tackle the varied potential serious limitations, which are particularly challenging for tests that are being standardized online. The most prominent limitations of data acquired via publicly available online tests are the lack of comparable viewing conditions and the difficulty to distinguish between data from three possible sources: earnest respondents (i.e., humans completing a test only once, to the best of their ability), as opposed to those that are unreliable (e.g. due to repeated participation, or random responses), and data that are automated, i.e. provided by non-human agents.

The problem: Identifying data from unreliable respondents and non-human agents

In the context of online data acquisition, researchers have at best limited ability to identify data from unreliable respondents. On the one hand, it is at best difficult (if not impossible) to ensure that publicly available online tests are not completed multiple times by the same individual without impinging upon data privacy and General Data Protection Regulation requirements. Another major issue concerns sustained human compliance with test instructions. This relates e.g. to observers completing a test in isolation vs. en groupe, and under consistent viewing conditions (distance, illumination, without distractions). The most challenging issue, however, concerns ensuring that a task is being performed *properly*.

Consider a classic test of observers' ability to simultaneously match facial identity. On each trial observers decide whether both faces depict the same identity by pressing one of two possible keys ("a" or "l"). To ensure compliance across the entire experiment, researchers implement "attention checks", e.g. by including infrequent "catch" trials requiring an independent response ("press the spacebar any time you see an upside-down face"). Observers' data are only included if performance for this control condition is above a certain threshold. Despite offering a degree of control, in my opinion this represents an insufficient quality assurance measure. Two observers meeting the attention check criterion may exhibit comparably low identity matching performance for different reasons. Observer 1 finds the identity matching task extremely difficult, while Observer 2 responds at random for "non-catch" trials. Both observers' data are admitted, strengthening the impression of a task being challenging.

Finally, automated responses from non-human agents represent an additional concern. Continuing with the previous example, data from Observer 1 and Observer 2 may also not be distinguishable from responses provided by bots. The issue of automated data raises an important question: If you are implementing online tests, should you be protecting against automated responses? If you do decide to protect against this, how sophisticated do you make the countermeasures? Unfortunately, to date I have never experienced a publicly available test of face cognition that requires a simple confirmation of the well-known "I am not a robot" statement.

The solution: Responsible test development

Naturally, researchers are aware of these issues and *some* try to decrease their potential impact. However, to most researchers the impact of unsuccessfully mitigating unreliable or automated responses remains abstract. Considering the interest of law enforcement in SR research, however, the stakes are much higher. If individual observers' test performance determines decisions regarding personnel deployment in security-relevant areas, or the importance of a witness's testimony, the impact becomes real.

With these issues in mind, as an advisor to and collaborator of different law enforcement agencies, I have practiced due diligence by adopting a simple and parsimonious three-step solution. First, tests administered for SR identification must have been standardized in reasonably sized, heterogeneous cohorts tested in person, *in the lab* or under controlled conditions (Stacchi et al., 2020; Fysh et al., 2020). After initial lab-based normative data acquisition, online versions of new tests are developed to determine whether both provide comparable data (cf. Fysh et al., 2020; Ramon and Docherty, in preparation). Second, it is *my* responsibility as a researcher to ensure that viewing conditions are comparable across observers; simply informing observers to *not* use a mobile device for completing a test is insufficient. Thus, it is critical that researchers implement screen calibration procedures *and* disable access from inappropriate devices. Again, unfortunately, these simple procedures have not been implemented in publicly available tests of face cognition (e.g., Dunn et al., 2020; Davis, 2019). Finally, measures to minimize the impact of unreliable respondents must be in place. For these reasons, I have do not implement publicly accessible online tests. Rather, I have personally tested each of the SRs reported here *individually* – either via remote, or in-person testing sessions. While undoubtedly time consuming, I deem this preferable and necessary to ensure high quality data. Additionally, in-person testing is arguably the most valuable source of insight for further work.

ongoing or currently in preparation.

Regarding communication of *all available* information, consider PG1. He participated in Philips et al.'s (2018) study, where individuals were reported as having been identified as SRs based on performance "above a superrecognizer benchmark" on "any [face recognition test reported] in the superrecognizer literature; e.g., Cambridge Face Memory Test extended […] or Glasgow Face Matching Test (GFMT)". In reality, PG1 was identified as a top performer in an online test "battery of 10 or more multi-ethnic face tests measuring four components" (Davis, 2019). Specifically, PG1 was informed that he ranked #87 among the top 500 performers (documentation provided in a personal communication) identified as SRs among *several millions* of observers. Thus, referring to PG1 as identified via above average performance on (any) *one* face cognition test (Phillips et al., 2018) represents a gross underestimation

[1] Importantly, note that the 181 observers are a subset from an initially larger group. These 181 individuals continued voluntarily participation, providing single-subject data *for a range of five* tests in total. These included the CFMT+, YBT-long and FICST, as well as the Expertise in Facial Comparison Test (EFCT) and Person Identification Challenge Test (PICT) (White, Phillips, Hahn, Hill & O'Toole, 2015). Thus, it is possible that these 181 individuals more likely to possessed elevated face cognition skills, as compared to those that declined to complete these further tests.Regardless of any such potential skewing towards more proficient data available for the CFMT+, YBT-long and FICST, note that top performers identified via these tests did not necessarily excel in the EFCT and PICT (cf. Stacchi et al., 2020). The lack of sensitivity of the EFCT and FICST is further evidenced by their inability to distinguish between typical observers and developmental and acquired prosopagnosics (White et al., 2017; Fysh & Ramon, in preparation).

Table 2
Super-Recognizer cases identified using the proposed novel diagnostic framework. Provided are demographic information (sex, age, handedness), performance across three tests of face cognition, and information regarding cases' inclusion in studies involving different methods.

	Code	Demographics			Face perception		Face Recognition	*Additional data acquired to date/participation in other studies*				
					FICST score (piles + errors-2) Max: 0	YBT long raw score Max: 35	CFMT + raw score Max:102	Fixation biases[1]	Infor-mation content[2]	Stimulus memora-bility[3]	Criminal investi-gation[4]	Neural represen-tations[5]
1	AC1	m	25	l	0	21	97				…	
2	AH1	f	43	r	3	16	94				…	…
3	AK1	f	53	r	3	17	89					
4	AM1*	f	30	r	0	17	99	…	…	…	…	…
5	AM2	f	57	r	1	18	84				…	
6	AW1	f	29	l	0	16	87					
7	BV1	m	45	r	3	19	90					
8	CB1†	f	31	r	2	20	87	…	…	…	…	
9	CB2	m	31	r	0	24	93				…	
10	CF1	f	57	r	2	23	95				…	
11	CH1‡	m	31	r	0	21	97				…	
12	CH2	f	48	r	1	16	96				…	
13	CV1	m	36	r	1	12	95					
14	CW1‡	m	26	r	5	21	100				…	
15	DF1‡	m	22	r	5	22	97				…	
16	DL1	m	48	b	0	16	97				…	
17	DS1	m	32	r	2	17	85					
18	DZ1	f	51	r	0	15	90				…	
19	EM1	f	32	r	7	24	89				…	
20	FR1	f	49	r	3	23	90					
21	FW1	m	31	r	1	22	97	…	…	…	…	…
22	GK1‡	m	37	r	10	16	100				…	
23	GP1	m	47	r	7	15	99	…	…	…	…	
24	GV1	m	33	r	0	13	97				…	
25	IS1	f	40	r	3	18	93					
26	JD1§|	m	38	r	0	19	97				…	
27	JN1‡	f	40	r	0	24	97				…	
28	JG1	m	43	r	0	15	99				…	
29	JS1	f	45	r	1	22	80				…	
30	JW1	f	26	r	3	25	99				…	
31	KW1◊	f	35	r	0	19	84				…	
32	LE1	f	29	r	2	21	86					
33	LG1	f	38	r	3	21	90				…	
34	LH1	f	45	r	2	15	85				…	
35	MB1	m	33	r	0	20	97	…	…	…	…	
36	MB2	f	44	r	0	18	96	…	…	…	…	…
37	ML1	m	44	r	5	19	99				…	
38	MM1↓	m	43	r	1	22	97				…	
39	MW1	m	42	r	1	20	93		…	…	…	…
40	NC1	f	40	r	7	17	92	…	…	…	…	…
41	NB1	f	38	r	13	23	92				…	
42	NY1§Δ	m	33	r	0	5	93				…	
43	OS1	m	45	r	2	16	87			…	…	
44	OW	m	47	b	1	20	87					
45	PG1#	m	32	r	0	20	97				…	
46	PK1	f	55	r	0	18	91				…	
47	PT1	f	33	r	0	17	78	…	…	…	…	
48	PT2	m	46	r	5	22	97				…	
49	RF1	f	40	r	0	29	92				…	…
50	RH1	m	53	b	0	20	97					
51	RG1	m	41	r	1	21	97					
52	SB1	m	32	r	5	17	93				…	
53	SB2	f	26	b	3	26	99					
54	SC1°	m	38	r	0	19	89				…	
55	SC2	f	25	r	0	18	88					
56	SH1	m	34	r	0	19	94				…	
57	SM1	f	41	r	0	20	98				…	
58	SM2	m	36	r	0	15	96				…	
59	SS1	m	48	r	0	20	101				…	
60	SS2	m	33	r	0	24	97					
61	TB1	f	46	r	0	7	96					
62	TH1	m	36	r	0	18	84				…	
63	TL1	m	48	r	0	22	93				…	
64	TK1	f	42	r	3	16	91					
65	TR1	f	27	r	0	17	87				…	
66	TU1	m	28	r	2	23	97				…	
67	UC1	m	42	r	0	20	92	…	…	…	…	…
68	UP1	f	37	r	6	18	91				…	

(*continued on next page*)

Table 2 (*continued*)

	Code	Demographics			Face perception		Face Recognition	*Additional data acquired to date/participation in other studies*				
					FICST score (piles + errors-2) Max: 0	YBT long raw score Max: 35	CFMT + raw score Max:102	Fixation biases[1]	Infor-mation content[2]	Stimulus memora-bility[3]	Criminal investi-gation[4]	Neural represen-tations[5]
69	VZ1	m	23	r	6	21	90	···	···	···		
70	YR1	f	42	r	1	9	95					

Asterisks provide references to previous reporting in the media or scientific publications; ··· indicates inclusion in additional/previous studies currently in preparation ([1]Linka et al., 2021; [2]Nador et al., 2021; [3]Nador er al., submitted; [4]Ramon, 2019; [5]Faghel-Soubeyrand et al., 2020). *Highest ranking individual among 500 performers identified as SRs among several millions of observers reported to have completed the online test battery (Davis, 2019); †https://www.3sat.de/wissen/nano/super-recognizer-100.html; ‡Reported by Bobak et al. (2016a) using the same code, albeit without the accompanying number added here; §https://www.presseportal.de/blaulicht/pm/12415/3237269; |https://www.deutschlandfunkkultur.de/gesichtserkenner-bei-der-polizei-schau-mir-auf-die-nase.2147.de.html?dram:article_id=390549; Δhttps://www.ksta.de/koeln/-ich-schlage-jeden-computer–wie-ein-super-recognizer-der-koelner-polizei-arbeitet-28162628-seite3; ◊https://www.annabelle.ch/leben/gesellschaft/bin-ich-ein-super-recognizer-49278; ↓Reported in a Bachelor thesis conducted in the German Police; #Ranked 87 on among the top 500 performers identified as SRs among several millions of observers reported to have completed the online test battery (Davis, 2019) and participated in a recently published study (Phillips et al., 2018); °Reported in a Diploma thesis conducted in the Swiss Police. All observers' data, as well as normative data can be found on the accompanying Open Science Framework public repository (DOI 10.17605/OSF.IO/3BUKV).

of his skills. Someone who ranks #87 among several millions of observers should also not be treated similarly to, and much less grouped together with, observers who achieved high performance on merely a *single test*. While unfortunately not uncommon, the practice of partial data reporting bears detrimental potential for any field of research.

Criteria for Super-Recognizer identification. I propose conservative diagnostic criteria involving challenging tests of face cognition (see Box 2). These criteria align with the seminal report of SRs (Russel et al., 2009) and, following the aforementioned guidelines, involve sensitive and appropriate tests for SR identification that are either well-established (Russell et al., 2009), or recently standardized under laboratory conditions (Stacchi et al., 2020; Fysh et al., 2020).

Testing for SR-identification incorporates assessment of different subprocesses of face cognition via two novel and challenging tests of unfamiliar face matching: the Yearbook Test (long version) (YBT-long) (Stacchi et al., 2020; Bruck et al., 1991) and the Facial Identity Card Sorting Test (FICST) (Stacchi et al., 2020; Fysh et al., 2020; Jenkins et al., 2011). Assessing face perception across superficial image variations and considerable age-related changes in appearance as they occur in real-life settings, both represent ecologically relevant tests of face cognition. These tests are complemented with the most frequently used test of face recognition under controlled settings: the long version of the Cambridge Face Memory Test (CFMT+) (Russel et al., 2009).

Using these tests, an individual's performance profile is considered superior if they excel in at least two of these three measures relative to the *current best estimates* across tests, which are derived from data that is presently available (cf. DOI 10.17605/OSF.IO/3BUKV). Taking into consideration differences in (i) task difficulty, (ii) performance range, and (iii) available sample sizes, the cut-offs provided in Table 1 were used. Using these cut-offs, 11 out of 181 observers (i.e., 6%) who completed the CFMT+, YBT-long and FICST (Stacchi et al., 2020) would have been identified as possessing superior face cognition skills[1]. This aligns with the theoretically rooted notion that SRs may exhibit distinct performance profiles: superiority may be confined to face perception, or may additionally encompass face recognition. Importantly, observations must be made under controlled or laboratory conditions given the serious limitations of currently available mass online tests (Box 3).

4. 70 Super-Recognizer cases identified using the proposed diagnostic framework

Applying these guidelines and criteria to data obtained through in-person physical or remote individual testing, I have identified a group of 70 individuals (law enforcement professionals and civilians) that possess exceptional face processing skills. Table 2 reports their demographic data alongside individual test scores. This heterogeneous group of SRs comprises 33 females and 37 males (average age: 39 ± 8 years; range: 22 to 57 years) with highly varied backgrounds and occupations. Contact to these individuals was established in three different ways: While most cases contacted me seeking assessment of their face processing abilities, others had been previously published as SRs (e.g. by Bobak et al., 2016a, (Bobak et al., 2016b,c); Phillips et al., 2019) based on superior performance for only one of the aforementioned tests, namels the CFMT+ (Russell et al., 2009). The third party was formed by individuals who were referred to me through their colleagues within various international police agencies. Most of the 70 individuals reported here had previously completed some other form of online testing (Davis, 2019), or had participated in previous studies, which did not necessarily report the full range of their known capacities (Phillips et al., 2018; Faghel-Soubeyrand et al., 2020). Table 2 provides information regarding previous testing, reports in the media or scientific studies, and participation in other studies (ongoing, or in preparation). Studying this large and growing group of SR cases, my lab's goal is to offer an increasingly comprehensive view of the mechanisms underlying superior face processing.

Individual observers' test scores and correlations between performance across tests. Fig. 2 demonstrates the 70 cases' superior level of performance for each individual test relative to previously reported groups of demographically heterogenous control observers (Stacchi et al., 2020; Fysh et al., 2020). With few exceptions, for each test SRs fall within the highest performance range. According to the adopted criteria an observer must achieve superior performance on *at least two of three* tests of face cognition, due to which SRs occupy one extreme of the 3-dimensional test space.

Similarity between performance profiles as a function of overall proficiency. Finally, the 70 SRs and 181 controls who completed all three tests (Stacchi et al., 2020) were ranked based on their overall proficiency. Observers' individual performance profiles (i.e., their unique combinations of standardized scores across the three tests) were then subject to Spearman correlations. Fig. 2 shows the resulting matrix of performance (dis)similarity. Each cell represents the similarity of performance across three tests, for a given pairwise combination of observers. Reflecting the adopted criteria for SR-identification, performance profiles among SRs are more similar compared to controls. Considering performance similarity as a function of overall performance rank (Fysh et al., 2020) one could anticipate higher performance similarity among low performers (due to *generally* lower abilities). Among highest performing controls, two scenarios are plausible: their performance profiles could be (i) highly similar due to a *general* proficiency for face cognition; (ii) relatively less similar if patterns across assessed subprocesses/tests are idiosyncratic. As demonstrated in Fig. 2, there is no systematic relationship between performance rank and similarity in

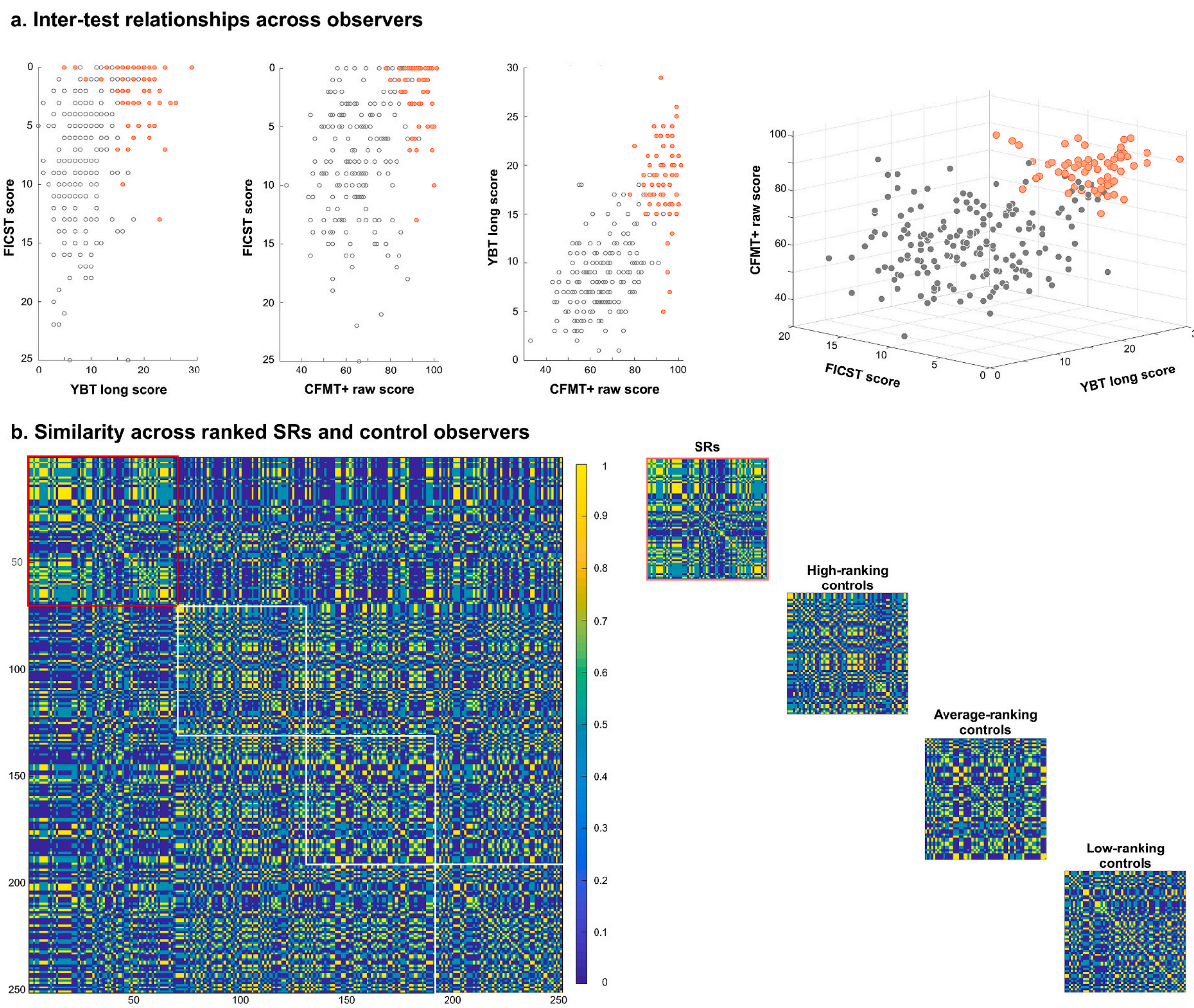

Fig. 2. Relationship between controls' and Super-Recognizers' performance across three tests of face cognition. a. The 2D plots (left) depict the relationships between individual observers' scores (controls in gray; SRs in orange) for all pairwise combinations of the three tests. The 3D plot (right) demonstrates that SRs, identified through the proposed criteria, are located at the high performing extreme of this multidimensional space. The three tests reported involve two tests of face perception, measured via facial identity matching with the YBT long (Stacchi et al., 2020; Bruck et al., 1991) and FICST (Stacchi et al., 2020; Jenkins et al., 2011), and one test of face recognition, measured via old/new memory for experimentally learned identities with the CFMT+ (Russell et al., 2009). Note that higher scores indicate superior performance for the YBT long and CFMT+; a FICST score of 0 indicates perfect performance. Control data plotted here were published previously (Stacchi et al., 2020); SRs' individual test scores are reported in Table 2. Note that the left 2D plot shows data from 218 controls who completed the YBT long and FICST; the remaining plots depict data from 181 controls who completed all three tests. **b.** The performance similarity matrix (PSM) demonstrates the similarity of behavioral performance profiles between all pairs of individual observers. The 70 SRs and 181 controls were ranked according to their overall performance across all tests. Then, performance similarity was computed for all observers' unique triplet of z-standardized test scores. The large PSM (left) displays the relationship between performance displayed by *all* observers, while PSMs on the right zoom into sub-groups. Each matrix cell represents the performance similarity for a given pair of observers based on their performance across the three tests. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

overall performance profiles among controls and SRs.

5. Discussion

Leveraging the informative value of Super-Recognizers as a special population. Case studies have provided the foundation of research in cognition and neuroscience, with unique cohorts playing a special role. The majority of these studies have focused on impairments and rare clinical syndromes (Rosenbaum et al., 2014; Medina et al., 2017). With few exceptions (Dresler et al., 2017; Shakenshaft et al., 2015), considerably less focus has been directed towards the opposite extreme, i.e. individuals exhibiting *high* performance levels. However, in light of a recent surge in interest in inter-individual variability in healthy and clinical cohorts (e.g., Finn et al., 2015; Gabrieli et al., 2015; Sui et al., 2020), interest in special, *high-performing* populations is also likely to increase. SRs are one example already receiving increasing attention. Bearing in mind both scientifc and applied interests, the goal of this paper was to increase the informative value of SR research. To achieve this goal, it addresses the current lack of a diagnostic framework that is fundamental in pursuit of a deeper understanding of human brain-behavior relationships.

Filling the diagnostic void. The present diagnostic framework for SR identification represents a departure from existing procedures in three important ways. Firstly, in line with Russell et al.'s (2009) original report, it involves assessment of different sub-processes of face cognition (perception and recognition) using tests that are both ecologically valid and highly sensitive (Russell et al., 2009; Stacchi et al., 2020; Fysh et al., 2020; Bruck et al., 1991; Jenkins et al., 2011). Secondly, this framework includes an expressed commitment to consistent and transparent case reporting to ensure amassing findings and minimizing information loss. Finally, the requirement of individuals exhibiting superior performance across *multiple* tests enables the distinction between different

Box 4
Recommendations for those involved in Super-Recognizer work

Civilians/potential Super-Recognizers

- Hold those interested in your skills accountable. You have the right to request your data, and demand evidence and transparency from researchers, companies, and associations.
- Do not pay for testing, travel, training, accreditations, etc. Your *data* is valuable enough and you should not be expected to both offer it, and pay for it, too. Contrary to claims made by fee-based associations, there is no scientific evidence that SRs "lose their ability" without training. Moreover, there is no empirical data on training efficacy for SRs, whose superior skills are observed without any formal training.
- Currently no-one can promise employment as a SR. SRs are not an officially recognized professional group, and there is no officially accepted body that can certify SR status.

Government practitioners

- Ask experts — *several independent* ones. Not all experts share the same knowledge and opinions. Practice due diligence by gathering different perspectives, requesting publicly available and peer-reviewed scientific evidence to support claims and documentation of procedures, and by vetting for potential conflicts of interest.
- Remember that *you* are the expert in *your field.* No-one understands the requirements of your job better than you do. Independent expert support is meant to compliment your skills and improve *your procedures.* Beware of one-fits-all-solutions; they may well be the hammer being sold by a person to whom everything looks like a nail.

Researchers

- Do not support empirically unjustified claims. The most prominent example concerns the prevalence of SRs, often referred to as being 1–2% of the population. In the absence of accepted diagnostic criteria or objective measures, such estimates cannot be made and are therefore unsupported.
- Address critical knowledge gaps and adapt procedures based on evidence and needs. Next on my agenda are those requiring large SR cohorts and inter-disciplinary efforts enabled via large-scale governmental collaborations: *How reliable are novel, sensitive testing tools? What are the neural mechanisms of SRs' high performance? Do empirically measured inter-individual differences predict real-life proficiency?*

Box 5
Practical steps to maximize the information gain of inter-disciplinary work

In 2016 the Fribourg Cantonal Police approached me to seek assistance from lab-identified SRs in the context of ongoing criminal investigations. Below I outline some key considerations that I developed through this experience. They guide my inter-disciplinary work with various international government agencies, who are interested in SRs as a potential means to counteract terrorism, investigate criminal acts, or increase safety in public spaces.

- ***Clear definitions and expectations.*** As outlined, peoples' definition of SRs and procedures to identify them are highly varied (Fig. 1b; Box 1). Thus, it is critical to first establish a common conceptual basis to avoid misunderstandings further down the line. Then, expectations directed towards to-be-identified individuals should be clearly defined. For example, an agency may be interested in identifying individuals that excel at *general person processing*, which can be achieved via various sources of information-not necessarily restricted to facial information. Clarity regarding concepts and expectations is fundamental for successful personnel selection and deployment.
- ***Appropriate diagnostic tools.*** Lab-based tests are usually designed to address specific research questions, and therefore involve highly controlled conditions. Such control may actually be undesirable in applied settings, where personnel selection tools must ensure efficient and successful deployment of individuals of interest. Thus, instead of simply adopting a one-fits-all approach, procedures should be aligned with the specific roles that individuals are expected to perform. If appropriate tools are non-existent, or as novel or changing areas of intended deployment emerge, bespoke diagnostic procedures may need to be developed.
- ***Evaluate transferability from the lab to real life***. Even if tools for personnel selection are *designed* to match real-world professional requirements, this does not mean that they are in fact able predict on-the-job performance, or that their deployment will lead to measurable performance gains. To this end a final step is required: pre/post evaluation of personnel selection and subsequent deployment. Funded through governmental resources, it should be an obligatory step, which is critical in order to provide information decision-makers require to implement evidence-based policies.
- ***Commitment to transparent data reporting — regardless of the outcome.*** Given collaborations between academics and enterprises promoting "SR services", and in light of governments' expressed interest in SR deployment, now — more than ever — access to unbiased data is critical. As a researcher with no corporate or financial conflicts of interest, I have both the freedom and obligation to demand the right to publish data, which I have acquired, regardless of their outcome. In my experience, practitioners appreciate findings that can help improve their operations, and they are more than willing to share them with their international peers. Communicating the results obtained in the context of inter-disciplinary collaborations with government is fundamental for establishing evidence-informed practices.

Box 6
Super-Recognizer identification within the Berlin Police: How adapting can create innovation and opportunities for science

The expectations expressed by law enforcement professionals (see Fig. 1b) might prompt the question: Do researchers need to adapt their definition(s) to meet practitioners' interest? Note that the professionals, whose responses are reported in Fig. 1b, operate(d) in a range of different areas related to public safety. Their expectations towards SRs contrast with the original SR report describing individuals with superior face cognition abilities (Russell et al., 2009), but resemble the description of alleged SRs reported to have identified perpetrators based on various non-facial sources of information, such as clothing (Towers, 2020). Analysis of such *non*-facial cues, such as behavior is easily trainable and in some security agencies represents a facet of police officiers' vocational training (Bättig, 2018). Since training effects for face cognition are at best moderate (Towler et al., 2014, 2019), it makes sense that law enforcement would seek to identify superior skills that are effectively not trainable, i.e. look for SRs among their employees. Nonetheless, it is crucial to understand the varied needs of practitioners (Ramon et al., 2019a) to ensure they are met through appropriate tools. In some cases, when existing tools fall short of meeting practical needs, bespoke solutions may even need to be developed.

A prime example is BeSuRe™ — a multi-level tool developed in a longstanding and ongoing collaboration with the Berlin Police in order to identify individuals with specific face processing skills among their >25K employees (Ramon and Rjosk, in press; Ramon, 2020a,b). Its development – combining scientific and applied expertise – was necessary for two reasons. First, it was crucial to ensure assessment of skills that directly reflect clearly defined intended areas of deployment. This required designing specific experiments, which most importantly, involve the use of authentic police material. Second, it was essential to ensure that the procedures were empirically validated under controlled conditions, in a representative cohort, i.e. among police employees. Preliminary findings from the BeSuRe™ pilot completed in December 2020 are promising in confirming that the test modules involve increasing levels of difficulty as intended, and yield consistent performance across controlled and uncontrolled test settings (Ramon, 2020b). The final version of BeSuRe™ is currently being rolled out among the Berlin Police. To date a number of state polices have communicated their interest in utilizing BeSuRe™, given our expressed intention to make it available to other law enforcement agencies interested identifying SRs for free (Ramon and Rjosk, in press; Ramon, 2020a,b).

The bespoke solution created to meet the Berlin Police's specific needs does not (have to) align with the diagnostic framework outlined here. However, all employees who volunteer to participate BeSuRe™ will in a first step complete the three tests outlined above. The resulting data obtained from such a large heterogeneous cohort will provide normative data of unprecedented scope and allow to formally address the currently open question of how lab-assessed face cognition relates to professionally critical skills *and* on-the-job performance.

performance profiles among SR-individuals.

Different players with unique challenges and responsibilities. SR work involves three main interest groups: researchers, potential SR individuals, and government practitioners from international law enforcement agencies. Motivated by exchanges with these groups, I provide practical recommendations (Box 4) and guidelines that characterize my interdisciplinary work with these groups (Box 5).

With respect to law enforcement, governments' decisions to prioritize identification and deployment of SRs may come at the expense of other areas that are also critical to public safety and security. Given the potentially wide-reaching implications of SR research, and to effectuate progress in the field, researchers should commit to the use of (i) clear and consistently applied concepts and definitions, (ii) systematic methodological approaches, and (iii) transparent communication of procedures and findings regardless of their outcome. Moreover, researchers must continuously critically (re)evaluate their procedures, and adapt their practices in order to meet interests outside of our confined field of expertise. This may involve shifting attention away from commonplace research methods suitable for specific questions that are relevant to a select group of peers. With decision-makers and practitioners seeking solutions for pressing real-world problems, we as researchers have to accept a change in pace, more pragmatically oriented approaches, and the need to transfer knowledge into other fields. Box 6 describes my experiences in a longstanding collaboration with the Berlin State Police. Such challenging collaborations create win-win opportunities that can generate findings required for evidence-informed practices *and* advance fundamental science in the future.

6. Conclusion: Ongoing work and future outlook

The 70 SR cases identified using the novel diagnostic framework proposed here will be expanded to include individuals identified in the context of large-scale government collaborations such as BeSuRe™ (Box 6). The present diagnostic framework thus lays the necessary foundation for future SR research seeking to advance our understanding of brain-behavior relationships.

Considering the limited data available in the new field of SR research (Fig. 1a), various questions remain unanswered – and the cognitive basis of SRs' abilities is entirely unclear. Research is needed to determine whether SRs' abilities extend to other domains and modalities, and/or come at a potential cognitive cost. Relatedly, we sought to determine whether SRs present with abnormal salience for faces as measured by reliable fixation biases (de Haas et al., 2019). Our recent work extends previous findings (Bobak et al., 2017) describing individual differences in visual sampling in greater detail (Linka et al., 2021). We have investigated the nature of perceptual processing differences (i.e., qualitative vs. quantitative) using psychophysical paradigms varying information content during face matching (Nador et al., 2021), and the relationship between proficiency in face cognition and stimulus memorability (Bainbridge et al., 2018; Nador et al., submitted). Ultimately, our goal is to characterize objective neural correlates of individual differences in face cognition, including those specific to SRs (Nador & Ramon, 2021), for which evidence is lacking (cf. Bellanova et al., 2018; Faghel-Soubeyrand et al., 2020).

Adopting the formal conservative diagnostic framework presented here will increase the informative value of future work towards a deeper understanding of the underlying mechanisms of superior face processing. Moreover, reproducible, rigorous and transparent research is crucial to answer societally relevant questions regarding the utility of SR deployment for law enforcement (Ramon, 2019), and how SRs can inform the development and deployment of AI and automatic face processing to enhance public safety (Ramon and Rjosk, in press). Combining complementary methods (Nickels et al., 2011) and inter-disciplinary approaches (Boxes 5 & 6), my lab's goal is to advance research on SRs and special populations, and more generally to address emerging societal challenges through sustainable science with real-world applications and ethical research practices.

Author contributions

MR wrote the manuscript.

Significance statement

What are Super-Recognizers (SRs)? Individuals with exceptional *person* recognition ability, or people who possess *face processing* skills? The term SR sparks interest from groups seeking to enhance scientific knowledge, public safety, or their monetary gain. Despite this interest, a clear definition of what actually constitutes a SR is lacking, and there is no accepted framework for diagnosis. This leads to varied findings and facilitates the spread of unsupported claims in the media. This is unfortunate for scientists interested in SRs to advance our understanding of human brain functioning, as well as international law enforcement interested in deploying SRs to enhance public safety. This paper provides a framework for SR diagnosis, introduces 70 SR cases, offers recommendations for practitioners, and highlights current caveats and future challenges.

Declaration of competing interest

The author declares no competing interests.

Acknowledgements

I thank Marlene Behrmann, Johannes B Bergmann, Anna K Bobak, Benjamin de Haas, and Micah Murray for helpful discussions and/or comments on a previous draft of this article. MR is supported by a Swiss National Science Foundation PRIMA (Promoting Women in Academia) grant (PR00P1_179872).

References

Bate, S., Frowd, C., Bennetts, R., Hasshim, N., Murray, E., Bobak, A.K., Wills, H., Richards, S., 2018. Applied screening tests for the detection of superior face recognition. Cognitive research: principles and implications 3, 22. https://doi.org/10.1186/s41235-018-0116-5.

Bate, S., Portch, E., Mestry, N., Bennetts, R.J., 2019. Redefining super recognition in the real world: skilled face or person identity recognizers? London, England : 1953 Br. J. Psychol. 110 (3), 480–482. https://doi.org/10.1111/bjop.12392.

Bättig, F., 2018. ASPECT® – Verhaltenserkennung. Analysing suspicious persons and cognitive training, 2018 Format 8, 20–25. Retrieved from. https://www.institut-police.ch/06-wissen/06-04-format-magazine/2018-08/inhalt-2018-08/20-Bttig.pdf.

Belanova, E., Davis, J.P., Thompson, T., 2018. Cognitive and neural markers of super-recognisers' face processing superiority and enhanced cross-age effect. Nov Cortex 108, 92–111. https://doi.org/10.1016/j.cortex.2018.07.008. Epub 2018 Jul 31. PMID: 30149236.

Bobak, A.K., Bennetts, R.J., Parris, B.A., Jansari, A., Bate, S., 2016a. An in-depth cognitive examination of individuals with superior face recognition skills. Cortex; a journal devoted to the study of the nervous system and behavior 82, 48–62. https://doi.org/10.1016/j.cortex.2016.05.003.

Bobak, A.K., Dowsett, A.J., Bate, S., 2016b. Solving the border control problem: evidence of enhanced face matching in individuals with extraordinary face recognition skills. PloS One 11 (2), e0148148. https://doi.org/10.1371/journal.pone.0148148.

Bobak, A.K., Hancock, P., Bate, S., 2016c. Super-recognisers in action: evidence from face-matching and face memory tasks. Appl. Cognit. Psychol. 30 (1), 81–91. https://doi.org/10.1002/acp.3170.

Bobak, A.K., Parris, B.A., Gregory, N.J., Bennetts, R.J., Bate, S., 2017. Eye-movement strategies in developmental prosopagnosia and "super" face recognition. Feb Q. J. Exp. Psychol. 70 (2), 201–217. https://doi.org/10.1080/17470218.2016.1161059. Epub 2016 Mar 31. PMID: 26933872.

Bruck, M., Cavanagh, P., Ceci, S.J., 1991. Fortysomething: recognizing faces at one's 25th reunion. Mem. Cognit. 19 (3), 221–228. https://doi.org/10.3758/bf03211146.

Burton, A.M., White, D., McNeill, A., 2010. The Glasgow face matching test. Behav. Res. Methods 42 (1), 286–291. https://doi.org/10.3758/BRM.42.1.286.

Cologne Police, 2016. POL-K: 160128-4-K/LEV Polizeipräsident verabschiedet "Super Recognizer. Retrieved from. https://www.presseportal.de/blaulicht/pm/12415/.

Davis, J.P., 2019. The worldwide public impact of identifying super-recognisers for police and business. The Cognitive Psychology Bulletin 4, 17–21.

de Haas, B., Iakovidis, A.L., Schwarzkopf, D.S., Gegenfurtner, K.R., 2019. Individual differences in visual salience vary along semantic dimensions. Jun 11 Proc. Natl. Acad. Sci. U. S. A. 116 (24), 11687–11692. https://doi.org/10.1073/pnas.1820553116. Epub 2019 May 28. PMID: 31138705; PMCID: PMC6576124.

Devue, C., 2019. Breaking face processing tasks apart to improve their predictive value in the real world: a comment on Ramon, Bobak, and White (2019). Br. J. Psychol. 110 (3), 483–485. https://doi.org/10.1111/bjop.12391.

Dresler, M., Shirer, W.R., Konrad, B.N., Müller, N., Wagner, I.C., Fernández, G., Czisch, M., Greicius, M.D., 2017. Mnemonic training reshapes brain networks to support superior memory. Neuron 93 (5), 1227–1235. https://doi.org/10.1016/j.neuron.2017.02.003 e6.

Dunn, J.D., Summersby, S., Towler, A., Davis, J.P., White, D., 2020. UNSW Face Test: A screening tool for super-recognizers. PLoS One 15 (11), e0241747.

European Union Agency for Fundamental Rights, 2019. Facial Recognition Technology: Fundamental Rights Considerations in the Context of Law Enforcement. Retrieved from. https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-facial-recognition-technology-focus-paper.pdf.

Faghel-Soubeyrand, S., Ramon, M., Bamps, E., Zoia, M., Woodhams, J., et al., 2020. Multivariate pattern analysis reveals domain-general enhancement of visual representations in individuals with "super-recognition" of faces. J. Vis. 20 (11), 502. https://doi.org/10.1167/jov.20.11.502.

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nat. Neurosci. 18 (11), 1664–1671. https://doi.org/10.1038/nn.4135.

Fysh, M.C., Stacchi, L., Ramon, M., 2020, 7. Royal Society Open Science, p. 200233.

Fysh, M.C., Ramon, M., 2021. It's about time. Testing tests of normal face matching with prosopagnosia.

Gabrieli, J., Ghosh, S.S., Whitfield-Gabrieli, S., 2015. Prediction as a humanitarian and pragmatic contribution from human cognitive neuroscience. Neuron 85 (1), 11–26. https://doi.org/10.1016/j.neuron.2014.10.047.

Jenkins, R., White, D., Van Montfort, X., Mike Burton, A., 2011. Variability in photos of the same face. Cognition 121 (3), 313–323. https://doi.org/10.1016/j.cognition.2011.08.001.

Linka, M., Alsheimer, T., Broda, M.D., de Haas, B.*, Ramon, M.*, 2021. Atypical visual salience in super-recognizers. Tagung Experimentell Arbeitender Psychologen (Conference of Experimental Psychologists).

Medina, J., Fischer-Baum, S., 2017. Single-case cognitive neuropsychology in the age of big data. Cogn. Neuropsychol. 34 (7–8), 440–448. https://doi.org/10.1080/02643294.2017.1321537.

Moreton, R., Pike, G., Havard, C., 2019. A task- and role-based perspective on super-recognizers: commentary on 'Superrecognizers: from the lab to the world and back again. Br. J. Psychol. 110 (3), 486–488. https://doi.org/10.1111/bjop.12394.

Nador, J.D., Zoia, M., Pachai, M.V., Ramon, M., 2021. Super-Recognizers: more consistent or qualitatively different psychophysical profiles? http://doi.org/10.31234/osf.io/2deap, 2021, March 5.

Nador, J.D., Alsheimer, T.A., Gay, A., Ramon, M., 2021, March 9. Image or identity? Only Super-Recognizers' (memor)ability is consistently viewpoint-invariant. https://doi.org/10.31234/osf.io/jgvuk.

https://onlinelibrary.wiley.com/doi/full/10.1111/ejn.15115, 2021.

Nickels, L., Howard, D., Best, W., 2011. On the use of different methodologies in cognitive neuropsychology: drink deep and from several sources. Cogn. Neuropsychol. 28 (7), 475–520. https://doi.org/10.1080/02643294.2012.672406.

Noyes, Eilidh, Phillips, P. Jonathon, O'Toole, Alice, 2017. What is a Super-Recogniser?.

Phillips, P.J., Yates, A.N., Hu, Y., Hahn, C.A., Noyes, E., Jackson, K., Cavazos, J.G., Jeckeln, G., Ranjan, R., Sankaranarayanan, S., Chen, J.C., Castillo, C.D., Chellappa, R., White, D., O'Toole, A.J., 2018. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. Proc. Natl. Acad. Sci. U. S. A 115 (24), 6171–6176. https://doi.org/10.1073/pnas.1721355115.

Ramon, M., 2019. Super-Recognizers in criminal investigation – hype or hope? J. Vis. 19, 137a.

Ramon, M., 2020a. Super-Recognizers: what We Know, Don't Know, and Need to Know. Keynote Talk Held at the 2nd Consortium Workshop of the EU-Project SafeCi – Safer Space for Safer Cities. EU Commission Berlin, Germany.

Ramon, M., . The Berlin Model for Super-recognizer Identification – from Development to Implementation. Keynote Talk Held at the 3rd Workshop of the EU-Project SafeCi – Safer Space for Safer Cities #policemeetsscience. Berlin, Germany. https://www.youtube.com/watch?v=KgeFMFTyaCU.

Ramon, M., Busigny, T., Gosselin, F., Rossion, B., 2016. All new kids on the block? Impaired holistic processing of personally familiar faces in a kindergarten teacher with acquired prosopagnosia. Vis. Cognit. 24, 321–355.

Ramon, M., Bobak, A.K., White, D., 2019a. Super-recognizers: from the lab to the world and back again. Br. J. Psychol. 110 (3), 461–479. https://doi.org/10.1111/bjop.12368.

Ramon, M., Docherty, N., 2021. Online behavioral experimentation - how not to mess it up.

M. Ramon, S. Rjosk(in press) Super-Recognizers in Policing – Best Practices Established during Development of the Berlin Model for SR-Identification. Best Practice Handbook of the EU-Project SafeCi – Safer Space for Safer Cities.

Robertson, D.J., Bindemann, M., 2019. Consolidation, wider reflection, and policy: response to 'Super-recognisers: from the lab to the world and back again. Br. J. Psychol. 110 (3), 489–491. https://doi.org/10.1111/bjop.12393.

Rosenbaum, R.S., Gilboa, A., Moscovitch, M., 2014. Case studies continue to illuminate the cognitive neuroscience of memory. Ann. N. Y. Acad. Sci. 1316, 105–133. https://doi.org/10.1111/nyas.12467.

Rossion, B., 2014. Understanding face perception by means of prosopagnosia and neuroimaging. Front. Biosci. (Online) 6, 258–307.

Rossion, B., Caldara, R., Seghier, M., Schuller, A.M., Lazeyras, F., Mayer, E., 2003. A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. Brain: J. Neurol. 126 (Pt 11), 2381–2395. https://doi.org/10.1093/brain/awg241.

Russell, R., Duchaine, B., Nakayama, K., 2009. Super-recognizers: people with extraordinary face recognition ability. Psychon. Bull. Rev. 16 (2), 252–257. https://doi.org/10.3758/PBR.16.2.252.

Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Krapohl, E., Simpson, M.A., Reichenberg, A., Cederlöf, M., Larsson, H., Lichtenstein, P., Plomin, R., 2015.

Thinking positively: the genetics of high intelligence. Intelligence 48, 123–132. https://doi.org/10.1016/j.intell.2014.11.005.

Stacchi, L., Huguenin-Elie, E., Caldara, R., Ramon, M., 2020. Normative data for two challenging tests of face matching under ecological conditions. Published 2020 Feb 19 Cogn Res Princ Implic 5 (1), 8. https://doi.org/10.1186/s41235-019-0205-0.

Sui, J., Jiang, R., Bustillo, J., Calhoun, V., 2020. Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. Advance online publication Biol. Psychiatr. S0006–3223 (20), 30111–30116. https://doi.org/10.1016/j.biopsych.2020.02.016.

Towers, T., 2020. Madeleine McCann Super Recognisers 'to Track Down Potential New Suspect in Holiday Pics'. Retrieved from. https://www.dailystar.co.uk/news/latest-news/madeleine-mccann-super-recognisers-to-21804254.

Towler, A., White, D., Kemp, R.I., 2014. Evaluating training methods for facial image comparison: the face shape strategy does not work. Perception 43 (2–3), 214–218. https://doi.org/10.1068/p7676.

Towler, A., Kemp, R.I., Burton, A.M., Dunn, J.D., Wayne, T., Moreton, R., White, D., 2019. Do professional facial image comparison training courses work? PloS One 14 (2), e0211037. https://doi.org/10.1371/journal.pone.0211037.

White, D., Phillips, P.J., Hahn, C.A., Hill, M., O'Toole, A.J., 2015. Perceptual expertise in forensic facial image comparison. Proc. Biol. Sci. 282, 1292. https://doi.org/10.1098/rspb.2015.1292.

White, D., Rivolta, D., Burton, A.M., Al-Janabi, S., Palermo, R., 2017. Face matching impairment in developmental prosopagnosia. *2006* Q. J. Exp. Psychol. 70 (2), 287–297. https://doi.org/10.1080/17470218.2016.1173076.

We need to talk about super-recognizers Invited commentary on Young, A.W., Noyes, E., Ramon, M., Bobak, A.K., White, D., 2019. Super-recognizers: from the lab to the world and back again, Br J Psychol Br. J. Psychol. 110 (3), 492–494. https://doi.org/10.1111/bjop.12395.