



A meta-synthesis of emotional intelligence research: methodological rigour and linguistic hype in systematic reviews and meta-analyses

Sarah K. Davis¹ · Bérénice Mahoney¹

Received: 18 February 2026 / Accepted: 10 June 2026
© The Author(s) 2026

Abstract

Across 3 decades of emotional intelligence (EI) research, the number of EI-focused systematic reviews and meta-analyses has grown. Regarded as rigorous evidence syntheses, these reviews can be used by academics and non-academics to inform interventions, yet their quality has not been critically examined. We attempt such evaluation using a mixed-method meta-synthesis methodology. We first apply the SIGN quality appraisal tool to assess the methodological rigour of 73 reviews published up to April 2024. This is followed by a deductively oriented thematic analysis of each *Discussion* section, where we examine use of linguistic hype—a form of misrepresentation in scientific reporting in which language is incongruent with the underlying evidence. Findings show most reviews ($n=59$; 81%) are of ‘low’ or ‘unacceptable’ quality yet highly cited. Thematic analysis developed three themes that describe linguistic hype across the full sample: 97% of *Discussion* sections contained claims positioning EI as *prestigious*, asserting causal reach and applied necessity beyond what the evidence supports; 95% used language that *exaggerates* the evidentiary basis, typically by stating caveats then overridden by confident claims; and 88% framed methodological weaknesses as *redeemable* artefacts of researcher decisions rather than as signals about the construct itself. Linguistic hype was present across all SIGN quality levels, though its frequency and form were attenuated in higher quality reviews. We conclude that linguistic hype in EI review writing may obscure unresolved conceptual and empirical concerns, posing a barrier to the field’s critical self-evaluation.

Keywords Emotional intelligence · Meta-synthesis · Systematic review · Meta-analysis · Linguistic hype

✉ Sarah K. Davis
sarah.davis@worc.ac.uk
Bérénice Mahoney
b.mahoney@worc.ac.uk

¹ School of Health and Wellbeing, University of Worcester, Worcester, UK

1 Introduction

Over 30 years ago, Salovey and Mayer (1990) launched their academic model of *emotional intelligence* (EI), describing this as a distinct form of intelligence for reasoning about emotions, arising from an interacting set of cognitive emotion information processing abilities. Later that same decade, the concept was popularised through Daniel Goleman's (1995) bestselling book: *Emotional intelligence: Why it can matter more than IQ*. Here, Goleman claimed his version of EI which included, "*self-control, zeal and persistence, and the ability to motivate oneself*", was more important than general intelligence (IQ) in predicting success in life, adding, "*these skills...can be taught to children, giving them a better chance to use whatever intellectual potential the genetic lottery may have given them*". (p. 8). Goleman's message fit well with egalitarian values, a growing rejection of IQ research and testing and the 'psychological fundamentalism' characterising Western socio-political culture. Emotional determinism was increasingly used to position the subjective self, coupled with the idea that emotional deficiencies were drivers of societal problems and solutions that could be 'fixed' (Ecclestone and Hayes 2009; Furedi 2004). As a quantifiable construct that scholars claim could be trained and taught, EI was an attractive idea. Claims that its absence explained, and its presence could solve a diversity of human problems across the lifespan, from mental ill-health and problematic interpersonal relationships to stalled careers and educational success (Matthews et al. 2004; Mayer et al. 2016; Petrides et al. 2016), catalysed the commercialisation of EI to those who saw the value of an 'emotionally smart' workforce. Numerous tests purporting to measure EI were developed and marketed to industry (Conte and Dean 2006). Simultaneously, EI was incorporated into universal school interventions to 'teach' EI to young people. For example, in the UK, the government invested £14 million in 2006–2008 to support a national roll-out of 'Social and Emotional Aspects of Learning' in primary schools (Hansard Written Answers 2007)—a now defunct programme based on Goleman's approach (Humphrey et al. 2013).

Not all scholars welcomed EI. Notable early critiques queried both the fundamentals of the concept, and its measurement (e.g., Landy 2005; Locke 2005). Others identified EI as a 'fad' owing to its rapid emergence and grand, 'cure-all' claims, against a backdrop of research marked by controversy and some resistance to disconfirmatory evidence (Antonakis 2015; Murphy and Sideman 2006). Recent critical reviews suggest somewhat limited progress has been made towards addressing these issues, citing unresolved difficulties with measurement stemming from inadequate theory (e.g., poor face validity and tautological findings, with EI tools measuring *outcomes* not components or drivers of EI) (see Dasborough et al. 2022 for an adversarial exchange rehearsing these issues). Despite these conceptual and empirical challenges, and critics urging EI researchers to "*...abandon their construct or go back to the drawing board*" (Antonakis et al. 2009, p. 255), the concept of EI continues to receive research attention. This includes the publication of academic articles seeking to advise non-experts on the use of theoretically contested EI tools (e.g., O'Connor et al. 2019).

As researchers with a history of examining EI, we attempt to respond to critics' calls for greater introspection and a "self-critical approach" in EI research (Dasborough et al. 2022, p. 14) by exploring why academic interest in EI has endured in this seemingly uncertain empirical landscape. We attempt this by considering whether the field of EI and its applications could be at risk of 'linguistic hype' (LH). This form of scientific reporting is a type of

misrepresentation characterised by incongruity between writers' language and the research described (Millar et al. 2019). This incongruity, comprised of hyperbolic language use that glamorises, promotes and exaggerates aspects of the research (Millar et al. 2019), could facilitate 'spin', an umbrella term for "*reporting that fails to faithfully reflect the nature and range of findings and that could affect the impression that the results produce in readers*" (Boutron and Ravaut 2018, p. 2613). We focus our critical examination on identifying LH in EI-focused systematic reviews (and by extension meta-analyses) since these are seen as 'gold standard' evidence syntheses that command a wide readership among researchers, practitioners, and the media. We scrutinise LH through two analytical methods. First, given the sustained growth in EI research and applications, we assess the methodological rigour of EI-focused systematic reviews and meta-analyses to determine whether this growth is underpinned by high-quality research evidence, or, whether lower quality reviews are indicative of a potential research climate characterised by misrepresentation. Second, we conduct a qualitative analysis of the *Discussion* sections of these reviews for evidence of LH, incorporating the assessed methodological rigour of these reviews. We do not to re-review or revisit issues associated with primary literature on the topic of EI (e.g., construct coherence), since others (as previously cited) have done so comprehensively, including those writing about the application of EI within applied areas; see, e.g., organisation and management: Harms and Credé (2010b, a); Education: Zeidner et al. (2002); Healthcare: Cherry et al. (2014). Instead, we attempt to critically assess the presence of LH, as a potential facilitator of 'spin' in secondary analyses of EI research. We do not claim that LH is unique to EI research. However, we argue that there are good reasons to suspect EI could be particularly vulnerable to this type of misrepresentation in scientific reporting, and we set out our case for this in the paragraphs that follow.

1.1 Emotional intelligence: Criticisms, responses, progress?

After more than a decade of EI research activity, Matthews et al. (2004) critically reviewed whether academic progress had been made to address growing concerns about EI's construct validity. They queried whether conceptual and theoretical consensus on EI existed, and whether measures of EI were valid, could predict important outcomes, and how useful EI was in practice. Their review concluded uncertainty in researchers' conceptual and theoretical notions of EI, as either 'trait' (a set of non-cognitive competencies and characteristics, akin to Goleman's approach) or 'ability' (a set of cognitive abilities specialised for emotion processing, aligned with Mayer and Salovey's model). They identified the adverse impact of this uncertainty on the operationalisation of EI, including the growing number of EI measures with questionable psychometric rigour. Debate at that time also centred on whether EI was representative of a 'jingle' or 'jangle' fallacy (Gignac 2009)—whether EI was a novel predictor of important outcomes and not merely an existing construct re-labelled (e.g., personality or cognitive intelligence). At that time, Matthews et al. (2004) concluded that continued methodologism would resolve some of these issues: "*In the case of EI, there are some grounds for optimism that ability-based tests will in time prove to be reliable and valid, supporting real-world applications*" (p. 192) and "*following historical precedent in ability testing, development of good operational measures and a psychometrically sound multi-stratum model may lead to theoretical progress*" (p. 193).

Fast-forward to the present era, and critics continue to argue that EI is a somewhat fuzzy and variously defined umbrella concept, which, as a consequence, remains inadequately measured (see e.g., Dasborough et al. 2022; Maul 2012). Furthermore, EI draws on a now-contested version of emotion correctness theorem whose implicit epistemology and ontology assumes universal truths about how we create our feelings, what ‘good’ and ‘bad’ feelings are, and that these emotional ‘do’s’ and ‘don’ts’ can be measured reliably in a social vacuum, assessed against pre-determined, external criteria (Deonna and Teroni 2024). However, advocates continue to suggest that concerns will be addressed and the field clarified by re-labelling different ‘types’ of EI measures to mark their differentiation from the original Mayer–Salovey ability EI model. Instead of ‘trait’ and ‘ability’ EI, organisation and leadership researchers prefer to identify ‘streams’ of EI; *stream 1* for measures of *maximal performance* based on the ability EI model, *stream 2*, for any *self-report* measures based on this model, and *stream 3*, for *‘mixed’ self-report* measures that tap competencies beyond those described in the ability EI model (Ashkanasy and Daus 2005). Mayer and Caruso (2025) go further to suggest that the self-report EI label (whether ‘trait EI’, ‘mixed EI’, etc.) should be dropped in favour of ‘socioemotional attributes’, and only ability-based research should be identified as ‘EI’. However, such academic debate is at odds with public and practitioner conceptions about EI. For example, unlike the narrow definition of the EI ability model, lay definitions are typically more inclusive or superordinate in scope and culturally influenced, referencing practical and social elements useful for good communication (Kittisiam et al. 2025; Lindebaum 2015). Indeed, we tend to agree with O’Connor et al. (2019) who in their review of measurement tools concede, “we are not proposing that these terms are ideal or even useful when classifying EI...” (p. 3), qualifying that, “many readers will simply be looking for an overall measure of emotional functioning that can predict personal and professional effectiveness.” (p. 5). Given non-academic consumers of EI research appear unlikely to nuance their conceptualisation of EI, we do not focus here on identifying trends in quality according to measurement model type. Rather, we are concerned with critically reviewing the quality of EI research reviews *at concept level*, given the high-status afforded to the conclusions of systematic reviews and meta-analyses, including by advocates of the construct (see Dasborough et al. 2022).

The growth of published EI outputs suggests researchers followed Matthews et al.’s (2004) call for continued methodologism, with outputs increasing noticeably from 2019 (see Fig. 1). This followed the release of the United Nation’s 2030 Agenda for Sustainable Development in 2015 and claims by UNESCO that EI, under the banner of Social and Emotional Learning, is “necessary” to attaining Sustainable Development Goals (Bryan 2024; Singh and Duraiappah 2020, p. 2) assertions that are repeated in industry, e.g., by test publishers acting as ‘UN action partners of the SDGs’: “With emotional intelligence, people, all of us, have the capacity to focus on what’s most important. Emotions can contribute to challenges... but with EQ, they become resources to empower change, fuel collaboration, and spark the innovation needed to make the SDGs real.” (Six Seconds, n.d.). In academia, 2016 marked the release of an updated ability EI model (Mayer et al. 2016) including new emotion skills (e.g., recognising cultural differences in emotion), and some 40 tools are also now available to measure EI (Bru-Luna et al. 2021). The number of published EI- focused systematic reviews (SRs) and meta-analyses (MAs) also increased during this period (see Fig. 1). Using standardised, comprehensive and transparent methods, SRs and MAs move beyond the limits of single primary research studies to summarise the overall strength and

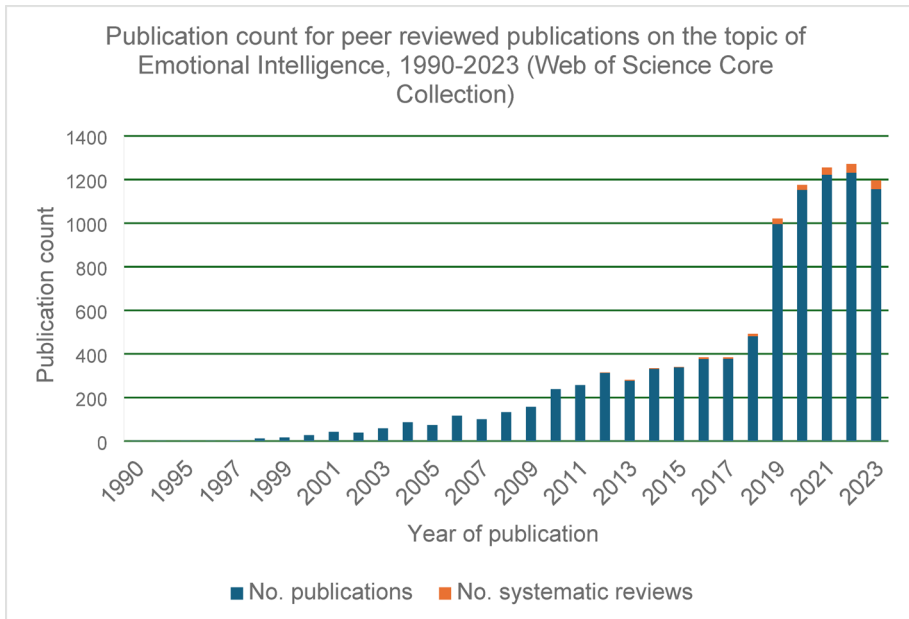


Fig. 1 9623 peer reviewed publications on the topic of EI in the period 1990–2023 (1990 marks the inception of the construct according to Mayer and Salovey’s original EI model). There is a significant increase in research output from 2019 to 2023, with increasing numbers of peer-reviewed, systematic reviews across this period

quality of a body of findings, identifying where further investigation and methodological improvements are needed (Page et al. 2021). Thus, these are regarded as high-quality evidence syntheses and are typically highly cited. Consequently, SRs and MAs are of interest to a wide audience and trusted to provide an unbiased orientation to a subject area and possess high quality standards of methodological rigour necessary to inform high-stakes evidence-based decision making about real-world applications including interventions (South and Lorenc 2020). In the case of EI, most of these reviews focus on assessing the predictive validity of EI applied in areas where a critical mass of research has been assumed to have reached sufficiency and/or stability, such as health. Consequently, EI-focused SRs and MAs are perceived as the ‘best’ source of evidence for academic and non-academic readers to understand whether claims of the constructs value and application are valid.

1.2 The case for assessing methodological quality and linguistic hype in EI SRs and MAs

As modern academia has become more competitive and marketised, restoring transparency, integrity, and truth in scientific reporting is regarded as a priority to combat questionable incentives that undermine research quality (Edwards and Roy 2017). SRs and MAs are important metascientific tools for accomplishing this restoration given they embody Open Science principles that aim to enhance “*openness, transparency, rigor, reproducibility, replicability, and accumulation of knowledge*” in research practices (Crüwell et al. 2019, p. 237). Open Science efforts have focussed mostly on encouraging good reporting practices in

primary research to mitigate various forms of misrepresentation in scientific writing including *misreporting of methods* (e.g., changing objectives/post-hoc analyses), *misreporting of results* (e.g., selective reporting of favourable outcomes), and *misinterpretation* (e.g., opinion vs. evidence; inappropriate use of causal statements; justifying after the results are known; misleading extrapolation of findings to wider settings/groups) (Boutron and Ravaud 2018). SRs and MAs are viewed as a ‘safety net’ that should mitigate such misrepresentation of primary research when conducted and reported in line with best-practice guidelines, e.g., PRISMA (Page et al. 2021), which include systematic quality appraisal of included studies. However, recent analyses suggest that Psychology reviews often contain methodological inconsistencies and reporting gaps, prompting scholars to argue for quality benchmarking of SRs/MAs themselves (see e.g., Hohn et al. 2020; Shercliffe et al. 2009). One recent analysis of 150 MAs published in *Psychological Bulletin* found that authors commonly omitted reporting search results, screening, and extraction procedures from individual studies, essential elements for transparency and reproducibility (Polanin et al. 2020). This parallels issues found in biomedical sciences where, “*instead of clarifying the evidence, too many of these reviews often suffer from extensive redundancy, little value, misleading claims, and/or vested interests*” (Ioannidis 2016, p. 487). Yavchitz et al (2016) identified 13 types of misrepresentation within SRs of therapeutic interventions, classifying these as (1) *misleading reporting* (e.g., selective reporting favouring a beneficial effect), (2) *misleading interpretation* (ignoring or failing to conduct risk of bias assessments of primary studies, the heterogeneity of results, and/or publication bias), and (3) *inappropriate extrapolation* (i.e., to different population, context, or intervention). Notably, unfounded or hyped recommendations for clinical practice in *Discussion* sections of these SRs (i.e., LH) was considered the most severe type of misrepresentation given they are most likely to distort reader interpretation.

Scholars debate the scale of language units at which LH operates (Ädel 2023; Hyland 2017; Millar et al. 2019), from individual words and phrases, to broader textual patterns (see Ädel 2023). In our assessment of EI SRs/MAs, we adopt a definition that encompasses words, phrases and broader textual patterns, such that LH is represented by *instances of incongruence between the nature of the underpinning research (e.g., methodological rigour/quality, statistical findings such as effect size, author acknowledged limitations of their research) and the language (words and phrases) used to make claims about what the research shows*. Importantly, specific words and phrases *on their own* cannot be considered linguistic hype, nor do they signal less or more hype. Rather, language becomes LH in meaning within the context of this incongruence, where its presence has the potential to impress “*judgments on readers that might undermine objective and disinterested evaluation of new knowledge*” (Millar et al. 2019, p. 149). Since linguistically, LH claims are exaggerated assertions, they threaten scientific credibility, even when the underpinning methods and analysis of the research are sound (Hoekstra and Vazire 2021). Thus, LH is distinguished from fraud (e.g., fabricating data), which violates scientific integrity directly—it can occur without intent to deceive and even with technically accurate statements, but it misleads through selective emphasis, or omission, to frame evidence misleadingly and is thus a communication error (Intemann 2022), exaggerating “*the state of scientific progress, the degree of certainty in models or bench results, or the potential applications of research*” (Caulfield et al. 2016, p. 776).

Evidence suggests that the prevalence of LH is growing across scientific fields (Hyland and Jiang 2021; Millar et al. 2023; Vinkers et al. 2015) and most often found in *Discussion* sections of primary empirical research papers. Hyland and Jiang (2021) found the incidence of two forms of linguistic hype—expressions of heightened conviction (e.g., ‘demonstrate’, ‘clearly’) and positive evaluative language (e.g., ‘promisingly’, ‘interestingly’)—increased by 93.8 across 360 multidisciplinary papers published between 1965 and 2015. Their analysis suggests authors use this language to project values of *certainty*, *contribution*, *novelty*, and *potential* about their research (Hyland and Jiang 2021), and, consequently, their expertise and preferred research methodology (Millar et al. 2019). This matters since LH claims cascade into mass media via university press releases (Sumner et al. 2014) and can facilitate spin, shaping public understanding and behaviour in ways not supported by the underpinning evidence (Matthews et al. 2016). Misplaced expectations and false beliefs arising from LH claims, if exposed, can lead to the erosion of public trust in science, as well as wasteful resource allocation, inaccurate decision-making, and “*a failure to pursue other solutions that might better address the problems they intend to solve*” (Intemann 2022, p. 2). With claims of continued significant investment from industry in EI training (e.g., TalentSmartEQ), supported by academic advocates (e.g., Chamorro-Premuzic and Sanger 2017; Crummenerl et al. 2019) it is pertinent to investigate whether the findings of SRs and MAs on the topic contain evidence of LH.

1.3 Purpose of the current work: Are EI-focused SRs and MAs characterised by poor methodological quality and linguistic hype?

With an increasing interest in meta scientific reporting across Psychology (Guilera et al. 2013), and the number of EI-focussed SRs and MAs growing, evaluating their methodological rigour, and extent of LH in their *Discussion* sections, seems timely. Again, we are not suggesting that LH is uniquely associated with EI research, however, we argue there may be certain qualities of the concept that make it more vulnerable to hype. Identifying the presence of LH (or lack thereof) within the *Discussion* sections of EI-focused SRs and MAs is an approach that can be used to scrutinise critically, whether the quality of academic evidence justifies the claim-making about EI’s psychological and social cure-all status in the public sphere.

We propose that assessing the methodological rigour of EI-focused SRs and MAs and identifying the presence of LH within their *Discussion* sections is important because the nature of the construct appears to fit Oyserman’s (2023) notion of a ‘culturally fluent’ theory. These feel intuitively sensible, palatable, and easy to grasp but are likely to be accepted without the level of critical appraisal applied to counter-intuitive ideas or findings given “*when reviewers share the cultural assumption, gaps [in evidence] may be less apparent and/or seem more tolerable*” (p. 227). For instance, labelling EI as a (different form of) *intelligence* suggests it is something prized, useful, and potentially achievable by all through learning/training, making it attractive to researchers who wish to address societal issues. Indeed, a rapid review of evidence concerning the potential drawbacks of EI, found that it was common for counter-intuitive findings to be buried within results sections with inadequate interpretation (Davis and Nichols 2016). We contend that, within a hypercompetitive and marketized ‘publish or perish’ academic environment invested in finding ‘quick fixes’ for psychological harm (Frawley et al. 2024), EI’s cultural fluency as an ‘everyday’ psycho-

logical phenomenon (Uher 2020) may act as ‘push’ factors undermining the methodological rigour of EI-focused SRs/MAs, and increase the likelihood of authors employing LH to ‘sell’ findings within *Discussion* sections of these reviews.

In line with Oyserman’s (2023) suggestion, we take a culturally disfluent lens in this current work, to assess the quality of EI SRs/MAs and the presence of LH in their *Discussion* sections, and address the following research questions:

1. Methodological rigour: following the work of Yavchitz et al. (2016) is there evidence of poor methodological quality in EI SRs and MAs? The growth of EI research and applications should be underpinned by rigorous evidence, and we focus on assessing whether SRs and MAs as ‘gold-standard’, secondary evidence syntheses (which include assessment of the methodological quality of primary EI studies), meet these high-quality standards, yielding trustworthy conclusions for readers. High quality EI reviews would be free from misleading reporting, misleading interpretation, or inappropriate extrapolation to recommend applications of EI beyond those supported by the findings.
2. LH claim-making: following the work of Hyland and Jiang (2021), is there evidence of linguistic hype—defined as language materially incongruent with the methodological quality and stated limitations of the review—in the *Discussion* sections of EI-focused SRs and MAs?

Our goal here is not to quantitatively sum effect sizes from across reviews in the manner of a meta-meta-analysis or umbrella review, but rather to present a new perspective on EI research by identifying themes in the claims made by review authors and interpreting their validity in relation to the *quality* of these reviews.

2 Methodology

The methodology was informed by Strobel and van Barneveld’s (2009) approach used to evaluate meta-syntheses, and is comprised of five steps: (1) search for articles, (2) selection and exclusion of articles (outlined in the *Method*), (3) appraisal of articles, (4) analysis of studies against one research question and aims (described in the *Findings*), (5) synthesis of analysis (outlined in the *Discussion*). This approach allows for the integration of qualitative and quantitative studies (as data, or units of analysis), and analysis of these units to understand themes within a research literature (Blair 1999), particularly when a field of research has grown but continues to experience conceptual and empirical uncertainty (Strobel and van Barneveld 2009) Here, we focus on EI SRs/MAs as units of data, appraising their methodological quality and understanding how claims about EI are represented and constructed, qualitatively, by review authors to examine the presence of LH.

2.1 Method

To understand the nature of the claims made about EI in research reviews, and whether these are underpinned by good quality research, we sought to retrieve published, systematic reviews that were specifically focussed on evaluating associations between EI and life outcomes (i.e., reviews focussed on summarising descriptive, correlational research). As the

purpose of our work centred on critically analysing published reviews rather than primary studies/unpublished grey literature, this suited a streamlined systematic search akin to a rapid review (King et al. 2022). Hence, we followed PRISMA guidelines (Page et al. 2021) to conduct a tailored search of two, topic appropriate databases, with a keyword search designed to retrieve reviews that placed a primary focus on ‘emotional intelligence’ rather than adjacent concepts or peripheral terms. We searched PsycINFO and Web of Science databases in April 2024 to retrieve published, peer reviewed, meta-analyses and systematic reviews using the terms “emotional intelligence” AND “systematic review” OR “emotional intelligence” AND “meta-analysis”. To be eligible for inclusion papers had to meet the following criteria:

1. Papers had to report a peer-reviewed, systematic review of primary research data, rather than narrative review, and be published in English, since 1990 (publication of the original Mayer and Salovey model of EI).
2. EI had to be the focal variable under investigation and not merely included in subsidiary analyses as part of broader investigations (e.g., emergent outcome from wider evaluation of training interventions, or EI as a moderator of a main effect).
3. Papers had to refer to ‘Emotional Intelligence’ explicitly as a construct, and include a measure of EI, rather than components of this, or adjacent concepts (e.g., emotion knowledge, emotion identification, empathy, social cognition, etc.)

Our initial keyword search for peer reviewed studies published in English since 1990, yielded 166 papers through PsycINFO, and 135 from Web of Science. 98 duplicates were removed prior to screening (see Fig. 2: PRISMA flow diagram). The remaining 203 papers were screened based on abstracts against the inclusion criteria. 97 papers were discarded

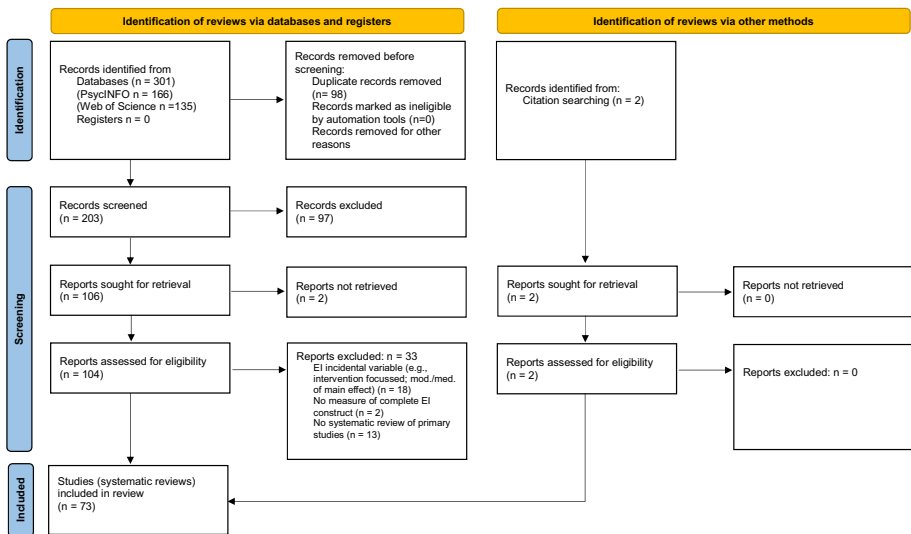


Fig. 2 PRISMA 2020 flow diagram mapping search for published systematic reviews using databases and manual search. *From:* Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. <https://doi.org/10.1136/bmj.n71>. For more information, visit: <http://www.prisma-statement.org/>

at this stage, leaving 106 papers eligible for full-text screening. 2 reviews could not be retrieved. The remaining 104 reviews were screened, with 33 subsequently excluded at this final stage for not meeting inclusion criteria. Examples of exclusions included, reviews that were intervention focussed, or included EI as an incidental variable only (i.e., as a mediator or moderator of effects), did not include any analysis or interpretation of EI study findings despite their inclusion in the review, or reported a narrative review, a scoping review, second order meta-analysis or cross-temporal meta-analysis. Please see [S1 supplementary materials](#) for full list of studies excluded at this stage along with reasoning for exclusion. A further 2 reviews were identified from a manual search of reference lists, producing a final set of 73 systematic reviews.

Studies were retrieved, screened for eligibility, and coded by the first author for methodological details and key findings. This included publication year, country of origin (first author affiliation), type of review (SR or MA), and citation count. Where meta-analyses were conducted, effect sizes relating to main effects were recorded. Any recommendations for research and/or application of EI in practice that were stated by authors in discussion and conclusion, were identified and coded categorically using an iterative process. This allowed us to examine whether there was evidence of misrepresentation as *inappropriate extrapolation* (Yavchitz et al. 2016) on the basis of the quality of the review. Categories included: *train EI through interventions; use EI in screening and/or selection processes; conduct further EI research (e.g., explore incremental validity, longitudinal designs; explore EI sub-facet analysis); use advanced statistical techniques; use new technology for data collection; extend study to include different population groups; refine EI measurement; search for further moderators of effects*). The second author independently checked screening and data extraction. To evaluate *misleading reporting and interpretation*, both authors appraised the quality of each review using the Scottish Intercollegiate Guidelines Network checklist for systematic reviews and meta-analyses (SIGN 2012), as outlined in Sect. 3.1. Findings are in Sect. 3.

2.2 Coding reliability thematic analysis of SR/MA Discussions

The study used coding reliability thematic analysis (CRTA; Braun et al. 2021), a type of codebook thematic analysis that employs two or more ‘coders’ to analyse data using a given set of codes and themes (codebook), but with the addition of inter-rater reliability to indicate coding quality (Braun and Clarke 2021). CRTA is appropriate for our study aims given it mixes quantitative with qualitative data, and because it enhances the transparency of our definition and assessment of LH, and the empirical analysis path from SR/MA Discussion text to codes and themes (Roberts et al. 2019). The codebook was developed by first categorising the review articles by area of focus, and these fell into three domains: links between EI and other individual difference variables (personality, cognition, emotion), health outcomes, or performance outcomes (Table 1). Following Strobel and van Barneveld (2009), we randomly sampled three papers from each of the three domains, given that data richness could be assumed equivalent across all peer-reviewed sources, to produce a sample of nine papers (Abdulla Alabbasi et al. 2021; Andrei et al. 2014; Arora et al. 2010; Gómez-Leal et al. 2022; Kopp and Jekauc 2018; Kun and Demetrovics 2010; Mancini et al. 2024; Miao et al. 2019a, b; Van Rooy et al. 2005). The nine *Discussion* sections were analysed to develop the codes, with both authors independently applying deductively oriented open coding to

Table 1 Scope of EI reviews

Domain	<i>N</i> reviews (<i>n</i> = meta-analyses)	Example outcomes studied
Personality, cognition and emotion	21 (15)	Dark triad; mindfulness; divergent thinking or creativity; personality; general cognitive ability; giftedness; emotion regulation; emotion complexity; manipulation; stress reactivity; attachment; relationship satisfaction
Health	21 (10)	Physical health conditions; mental health conditions (within general population and specific contexts e.g., work related burnout)
Performance	31 (17)	Training; academic success; leadership; patient interactions/caring; sports performance

generate initial codes and themes (Byrne 2022). Codes were developed openly, without closed a priori definitions, but guided by a broad deductive orientation toward the nature of knowledge claims about EI *conceptualisation* and *methodological status*, its *application*, and our working definition of LH.

As noted earlier, LH only arises when claims are misleading; specific words and phrases on their own cannot be considered linguistic hype, nor do they signal less or more hype. To be consistent with our working definition of LH as a form of ‘overclaiming’ denoted by *incongruence* between the nature of the underpinning research and the language used to make claims about what the research shows, passages were coded only where language was materially incongruent with what the evidence actually supports (i.e., context in the form of claims evaluated in relation to the reported effect size in meta-analyses, SIGN methodological quality rating for each review, and stated limitations in the *Discussion*). Thus, the unit of analysis was the relationship between the language and the evidentiary context. This analytic process—where small units of meaning (words or phrases) and context, were identified and grouped into themes based on shared patterns—allowed capture of LH in flexible, unspecified forms while remaining grounded in the data and avoiding premature restriction of how LH might be expressed (Braun et al. 2022; Byrne 2022). Initial codes and themes were discussed collaboratively, followed by two further rounds of analysis on additionally sampled reviews to refine the coding framework and confirm no major new thematic categories were emerging. The resulting codebook was applied deductively to the remaining 64 reviews, yielding a full-sample analysis (N = 73). The rationale for extending to the full sample was to provide evidence proportionate to the scope of the claims, and to allow frequency and quality-distribution analyses across all SIGN ratings.

To evaluate the consistency and transparency with which the final codebook could be applied, inter-coder reliability was assessed on a 10% subsample of the deductive dataset (110 clearly delineated paragraphs drawn from the discussion sections of seven academic articles). The two authors independently coded each paragraph for the binary presence or absence of the three non-mutually exclusive themes:

1. *Prestigious*: Claims of EI's causal reach, importance, or application, are asserted beyond reported effects or study quality
2. *Exaggerated*: Hedging language precedes or follows a strong claim. Coded only when a hedge does not meaningfully qualify the claim (i.e. the caveat is overridden)
3. *Redeemable*: Weaknesses framed as methodological artefacts protecting the conceptual status of EI

As the themes could co-occur within the same paragraph, coding was performed separately for each theme. For each theme, a 2×2 contingency table was constructed pooling decisions across all 110 paragraphs, and Cohen's Kappa (κ) was calculated to quantify agreement beyond chance (McHugh 2012). Discrepancies were discussed collaboratively, and the codebook definitions, inclusion/exclusion criteria, and illustrative examples were refined iteratively as needed. The resulting reliability coefficients support the trustworthiness of the deductive coding process while acknowledging the interpretive nature of qualitative analysis; according to guidelines proposed by Landis and Koch (1977), inter-coder reliability was 'substantial' for all three themes (Theme 1 Prestigious: $\kappa=0.62$, $SE=0.08$, 95% CI [0.45, 0.78]; Theme 2 Exaggerated: $\kappa=0.62$, $SE=0.09$, 95% CI [0.44, 0.80]; Theme 3 Redeemable: $\kappa=0.64$, $SE=0.15$, 95% CI [0.35, 0.93]), together indicating good consistency in the application of the coding scheme.

3 Findings

The earliest systematic review in our sample ($N=73$) was published in 2004, and most reviews sought to include primary research using both 'trait' (self-reported EI) and 'ability' EI (maximal performance measures)—11 focussed exclusively on trait EI, and 2 on ability EI only (see [S2 supplementary materials for a complete listing of included studies/characteristics](#)). The largest proportion of reviews originated from first-author affiliations with institutions in Spain (27%), the USA (25%), and Australia (11%), see Fig. 3, which also shows the in-country distribution of SRs according to subject domain.

Of the 73 systematic reviews, 42 reported meta-analytical findings, providing a statistical summary of the strength of effects linking EI to outcomes. To characterise the sample, we recorded effects for 'overall EI' and by 'type' of EI where reported. Methodological variability in the classification of 'types' of EI across meta-analyses can be seen in the *notes* column of Table S3 (see [Supplementary file](#)).

In some cases, 'self-report EI' aligns with the 'trait EI' approach—a methodological distinction where all self-report measures are classified as distinct from tests of performance, referred to as 'ability EI' (e.g., Martins et al. 2010). Elsewhere, this 'trait EI' approach is split into different 'streams'—either classified as 'self-report EI/stream 2' or 'mixed EI/stream 3'. The former includes trait measures that are purportedly based on the Mayer–Salovey ability EI model, whilst 'mixed EI' are self-report/trait measures that index a range of self-perceptions beyond the Mayer–Salovey model, including emotional self-efficacy, personality traits and motivation (see e.g., work of Miao and colleagues). We note that it was often unclear how authors classified individual measures/studies, and/or variation in approach when this was made patent to the reader, for instance, the Trait Meta Mood Scale has been classified as a stream 2 measure, e.g., Sánchez-Álvarez et al. (2020), or stream 3, e.g., Kopp and Jekauc (2018). Moreover, these labels are inconsistently applied (e.g., Ngoc et al. 2020 used 'trait EI self-report' vs. 'self-report ability based'; Van Rooy et al. 2005 used 'mixed EI' for *all* self-report/trait EI measures). Across our sample of meta-analyses, we note that 17 authors adopted the EI 'stream' labelling approach, 14 classified EI outcomes according to instrument methodology, whilst 2 used an ad-hoc approach not fitting with either of the former (e.g., Schlaerth et al. 2013). One meta-analysis used both approaches, covering all bases (Kopp and Jekauc 2018). Some MAs focussed on only one approach (i.e.,

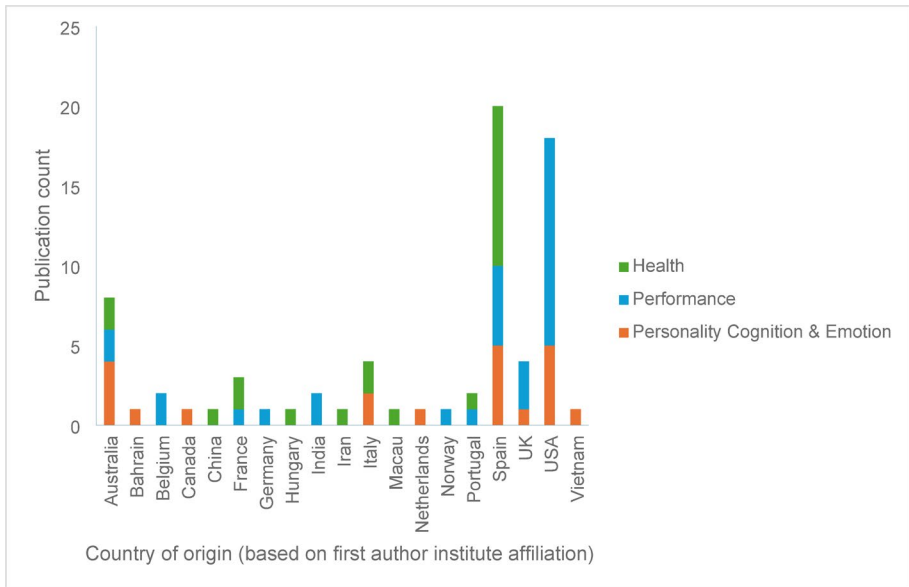


Fig. 3 Subject focus for systematic reviews by country of origin

either ‘trait’ or ‘ability’), yet used the aggregate term ‘EI’ rather than demarcating further (see e.g., Malouff et al. 2014), whilst others chose not to apply any classification and simply combined all measures of EI together (e.g., Pirsoul et al. 2023; Quílez-Robres et al. 2023). Clearly, these approaches could lead to confusion for readers who may not be aware of methodological and theoretical distinctions which carry implications for any recommended interventions or applications.

We note that beyond classification labels, there is also wide variation within the sampling domains of different ‘trait’ self-report measures (e.g., EQi, TEIQue, SSEI, TMMS), and increasing variation in measures classified as ‘ability EI’ (e.g., in addition to the MSCEIT and its predecessor MEIS, the STEM, STEU, TIEFBA, etc.), with Van der Linden et al. (2017) identifying several ‘custom made’ ability EI measures. On this point, Kopp and Jekauc (2018) note the wide variation of measures included across studies within their review (14 different measures per 21 studies), adding that 7 could not be classified as belonging to a ‘stream’, with Jardine et al. (2022) being unable to apply any classification due to the range of measures. This issue will be compounded as the field becomes increasingly saturated with new measures (see, Bru-Luna et al. 2021). Overall, classification and labelling issues makes comparison of summary effect sizes and key findings across meta-analytic reviews potentially misleading and unhelpful. For readers interested in comparing the outcomes of reviews, we would therefore only recommend comparing summary statistics for discrete measurement tools and interpreting accordingly.

3.1 Methodological rigour: quality assessment of systematic reviews and meta-analyses using SIGN

The methodological quality of each review was systematically appraised by both authors using the [Scottish Intercollegiate Guidelines Network checklist for systematic reviews](#) and meta-analyses (SIGN 2012) which was selected based on its suitability for evaluating the quality of non-clinical reviews (Baker et al. 2010), in line with the nature of research reviews in the EI field. Across 12 domains, SIGN requires appraisers to evaluate sources of bias and confounding that can influence the internal validity of a review in terms of the design, execution, and rigour of the systematic search and its reporting, and where applicable, the methods used to statistically combine findings (for meta-analyses). Explicit guidelines for the assessment of each SIGN domain are provided for raters; reproduced for readers in column two, Table 2. It should be noted that SIGN *domain 1* (specification of research question) and *domain 2* (comprehensive literature search) are viewed as critical indicators of review quality, with a failure to meet either of these criteria automatically yielding a 'reject' quality rating for a review owing to the unacceptable risk of bias in the methodology employed. Failure to meet *domain 7* (documenting the characteristics of primary studies included within the review) or *domain 8* (assessment and reporting of the quality of primary studies included in the review) would automatically result in an overall rating of 'low' quality where there is a high risk of methodological bias in the review. Relatedly, if *domain 8* is absent, *domain 9* cannot be met (the scientific quality of included studies used appropriately in analysis/conclusions). Both authors independently evaluated the reviews applying SIGN guidelines. There was no disagreement in domain ratings. Once the review has been evaluated according to each of the 12 domains, an overall rating of methodological quality is derived. SIGN provides guidance for four ratings: *High quality* (++) : The majority of criteria met (i.e., little or no risk of bias); *Acceptable quality* (+) : Most criteria met with some flaws in the study with an associated risk of bias; *Low quality* (-) : Either most criteria not met, or significant flaws relating to key aspects of study design, and *Reject* (0) : Poor quality study with significant flaws. Since SIGN domain guidelines are clear in identifying studies that should be graded as *reject* or *low quality*, the authors agreed through discussion that *high quality* studies should meet all criteria but there could be some exception for domain 3 and 4 (i.e., some allowable/small amount of potential bias in the reporting of study selection and data extraction processes). Table 2 provides aggregated quality appraisal data per SIGN domain; see [S2 supplementary materials for individual SIGN domain ratings for each review](#).

A high proportion of reviews showed weaknesses in study selection (64%), extraction processes (51%), the inclusion of 'grey' unpublished literature (56%), reporting of excluded studies (93%), and considering potential conflicts of interest (53%). Most critically, 78% of works failed to appropriately assess the quality of primary studies comprising the review. Only 10% of studies providing a quality analysis of included literature, used this information effectively when drawing conclusions (i.e., by excluding poor quality studies, recommending caution in interpreting findings, or explicitly discussing lack of methodological rigor in connection with recommendations). This was a significant factor in why 88% of meta-analytical reviews did not meet required standards for combining individual study results (SIGN domain question 10). Ratings for each of the 12 SIGN domains along with the overall quality judgement for each review are displayed in Fig. 4 as a risk-of-bias 'traffic

light' plot (McGuinness and Higgins 2021). Only 5 (7%) reviews were high quality, showing little risk of bias. These studies typically followed guidelines for reporting systematic reviews (e.g., PRISMA), were published between 2010 and 2020, with a citation range of 36–526 (Google Scholar, April 2024). 9 (12%) were of acceptable quality, published between 2021 and 2024, and cited 4–131 times. However, 36 (49%) were deemed low quality with significant weaknesses (publication date 2010–2023, citation range 4–1589), and 23 (32%) were of an unacceptable standard for publication (publication date 2004–2021, citation range 5–1848). Rejectable quality reviews commonly lacked a systematic and comprehensive search (no keywords, insufficient databases, no documenting of search process, no supplementary search) and 3 did not have clear research questions. None of the rejectable reviews performed any quality checks on the primary studies included in the review, nor did they comment on methodological quality in their discussion sections. It is notable that the poorer quality studies have reached a wider readership.

Frequency analysis of the recommendations made by authors in the discussion and concluding sections of the reviews are displayed in Fig. 5. Review authors were most likely to call for further research ($n=68$), implying that EI is worthy of further study. However, the second most popular recommendations were calls to train or teach EI in participants to enhance performance ($n=50$) and to extend the study of EI to other populations ($n=50$). Whilst poor quality reviews were more likely to make a greater number of recommendations per se, the 'higher stakes' claims (i.e., calls for training interventions, to use EI as a screening tool) were made most often in lower quality reviews, and appeared at the highest rate in rejectable quality reviews.

In our sample of meta-analytical reviews, where statistical summaries were attempted for EI types (i.e., 'trait' vs. 'ability'), we note that most effect sizes are reported in reviews with rejectable or low methodological quality, with relatively few coming from acceptable or high-quality studies (see [S4 online table for data extraction](#)). This reinforces the finding that methodological rigour is a significant concern across the sampled meta-analyses. Most high-quality meta-analyses report small effects according to both Cohen (1988) and Gignac and Szodorai's (2016) classification criteria.

3.2 LH claim-making

The CRTA developed three themes, each comprised of two to four sub-themes (see Table 3 which includes frequency counts and SIGN quality distributions; the full analysis and coded transcripts can be found in [S5 supplemental materials](#)).

LH was reflected sample-wide, with all three themes evident in 59 of the 73 coded reviews (81%). All 23 Reject-rated reviews and 35 of 36 Low-rated reviews contained coded instances across at least two themes; Miao et al. (2018b) was the exception in this quality band, with coded instances for Theme 3 only. All nine Acceptable-rated reviews contained coded instances across at least two themes, and all five High-rated reviews contained coded instances in at least one theme, though with lower density and in more constrained forms. Nine reviews contained no coded instances for Theme 3 (Kotsou et al. 2019; Peterson et al. 2011; Pirsoul et al. 2023; Ranjbar and Bakhshi 2018; Vega et al. 2021; Laborde et al. 2016; Malouff et al. 2014; Encarnação et al. 2018; Domínguez-García and Fernández-Berrocal 2018); and four contained no coded instances for Theme 2 (Lea et al. 2019; Miao et al. 2018b; Toriello et al. 2022; Walker et al. 2022)—all others activated all three themes to

Table 2 Summary of review quality ratings for each SIGN checklist domain (figures in bold show indicate where > 50% of reviews failed to meet criterion)

SIGN checklist domain	Guidance notes (reproduced from SIGN checklist)	<i>n</i> reviews meeting criteria "Yes"	<i>n</i> reviews not meeting criteria "No"
(1) The research question is clearly defined and the inclusion/exclusion criteria must be listed in the paper	The PICO must be clear in the paper even if not directly referred to. The research question and inclusion criteria should be established before the review is conducted. If 'no' reject (unacceptable quality)	69 (95%)	4 (5%)
(2) A comprehensive literature search is carried out	At least two relevant electronic sources must be searched. The report must list the databases used (a grey literature search counts as supplementary; PubMed and MEDLINE count as one database) Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. Dates for the search should be provided The paragraph above is the minimum requirement. If 'no' reject (unacceptable quality)	55 (75%)	18 (25%)
(3) At least two people should have selected studies	All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or/and experts in the field of study, and by reviewing the references in the studies found. The paragraph above is a quality criterion which affects the overall rating of the review	37 (51%)	36 (49%)
(4) At least two people should have extracted data	At least two people should select papers. There should be a consensus process to resolve any differences	26 (36%)	47 (64%)
(5) Grey/unpublished data was sought	At least two people should extract data and should report that a consensus was agreed. One person checking the others data extraction is accurate is acceptable	4 (9%)	37 (51%)
(6) The excluded studies are listed	The authors should state that they searched for reports regardless of their publication status. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status If review indicates that there was a search for "grey literature" or "unpublished literature," indicate "yes." SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains both grey and non-grey, must specify that they were searching for grey/unpublished lit Limiting the excluded studies to references is acceptable	32 (44%)	41 (56%)
		5 (7%)	68 (93%)

Table 2 (continued)

SIGN checklist domain	Guidance notes (reproduced from SIGN checklist)	<i>n</i> reviews meeting criteria "Yes"	<i>n</i> reviews not meeting criteria "No"
(7) The relevant characteristics of the included studies are provided	In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the included studies e.g., age, race, sex, relevant socioeconomic data should be reported. (Note that a format other than a table is acceptable, as long as the information noted here is provided). Absence of this will make it impossible to form guideline recommendations. Mark as (- low quality) original papers would need to be examined	48 (66%)	25 (34%)
(8) The scientific quality of the included studies was assessed and reported	It can include use of a quality scoring tool or checklist, e.g. risk of bias assessment, or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable. Absence of this will make it impossible to form guideline recommendations. Mark as (- low quality)	16 (22%)	57 (78%)
(9) Was the scientific quality of the included studies used appropriately?	Examples include sensitivity analysis based on study quality, exclusion of poor quality studies, and statements such as 'the results should be interpreted with caution due to poor quality of included studies'. The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations. Cannot score "yes" for this question if scored "no" for question 8	7 (10%)	66 (90%)
(10) Appropriate methods are used to combine the individual study findings	Studies that are very clinically heterogeneous should not be combined in a meta-analysis. Look at the forest plot—do the results look similar across the studies? For the pooled result a test should be done to assess statistical heterogeneity i.e. Chi-squared (χ^2) test for homogeneity and/or I^2 test for inconsistency. If significant heterogeneity is apparent the authors should have explored possible explanations using methods such as sensitivity analysis or meta-regression. A random effects analysis may be used to take account of between-study variation but is not a 'fix' for heterogeneity. Planned subgroup analyses should be pre-specified and limited in number because conducting many subgroup analyses increases the probability of obtaining a statistically significant result by chance. Conclusions based on post-hoc subgroup analyses must be interpreted with caution. Cannot score "yes" for this question if scored "no" for question 8	5 (12%)	37 (88%) (N/A= 31 SR only)
(11) The likelihood of publication bias was assessed appropriately	The possibility of publication bias should be assessed where possible, commonly done by visual inspection of a funnel plot together with a statistical test for asymmetry (e.g., Egger regression test) although other statistical and modelling approaches may be reported. Absence of a funnel plot doesn't mean the likelihood of publication bias was not assessed appropriately (there are other methods); 10 studies is just a ball-park minimum number for a funnel plot and a plot is of little use when there are few studies	29 (69%)	13 (31%) (N/A= 31 SR only)
(12) Conflicts of interest are considered	Potential sources of support should be clearly acknowledged in both the systematic review and the included studies. 'Yes' if a COI statement is included	34 (47%)	39 (53%)

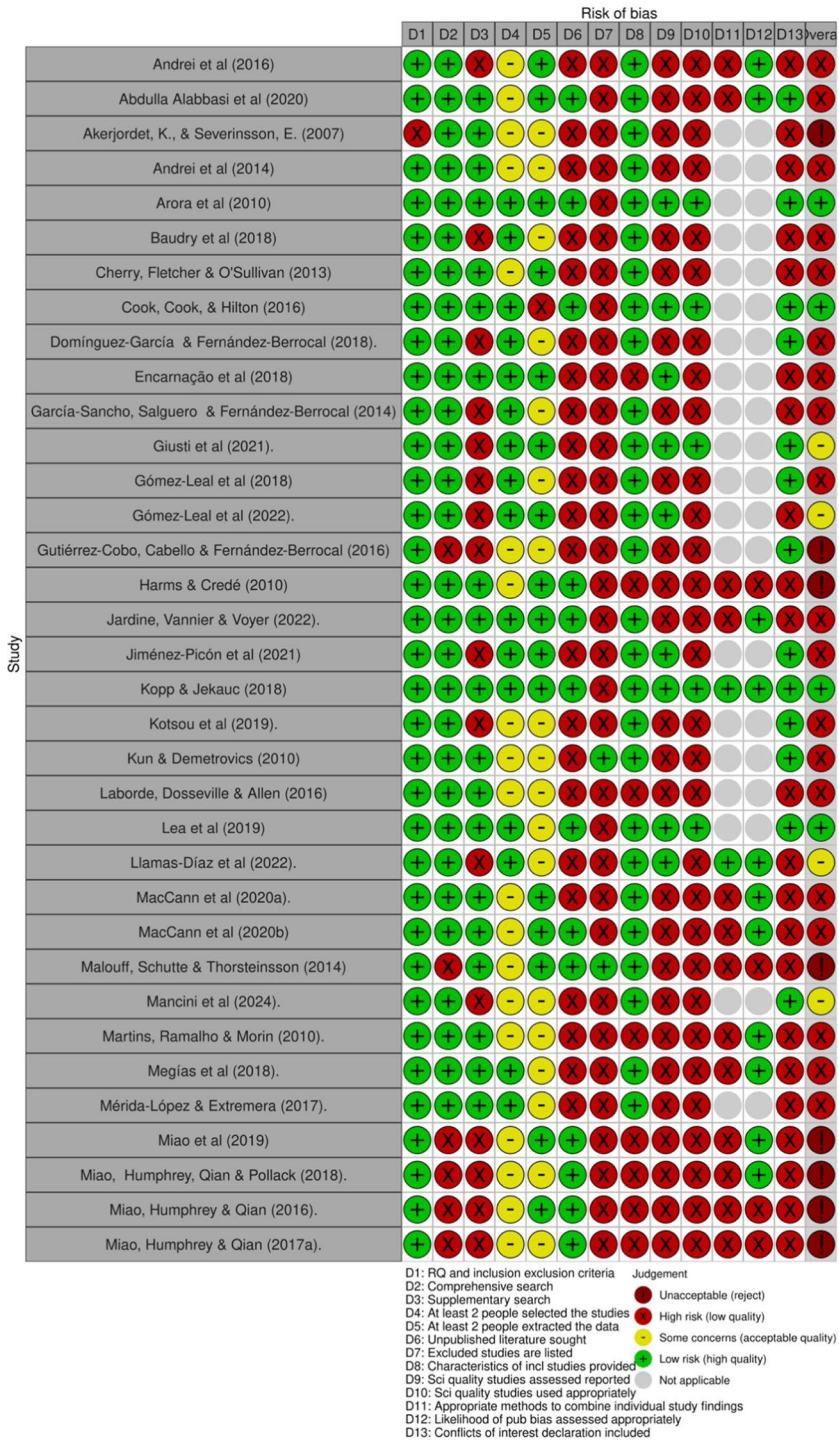


Fig. 4 Risk of bias ‘traffic light’ plot for SIGN domain judgements for each review (N=73)



Figure 4 (continued)

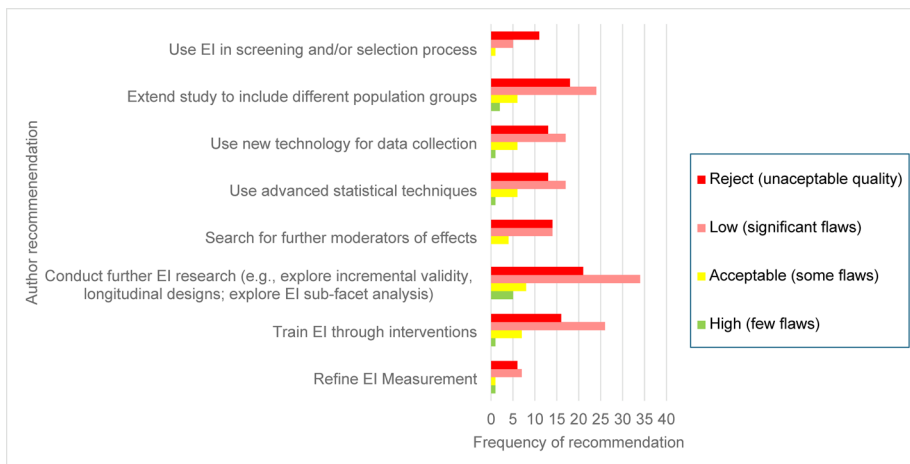


Fig. 5 Nature and frequency of author recommendations grouped by quality of review

some degree. Coding density (i.e., number of coded passages per *Discussion*) was highest in Reject-rated reviews and lowest in High-rated reviews, indicating that *quality appraisal affects the frequency and form of LH rather than its presence*. This is discussed further in Sect. 3.2.4.

The overarching thematic pattern in the reviews sampled is one in which *Discussion* sections collectively construct EI as a *prestigious* construct (Theme 1, $n=71$, 97%) whose applied value is asserted with language that *exaggerates* the evidentiary basis of the research reviewed (Theme 2, $n=69$, 95%). Where methodological weaknesses are acknowledged, they are rarely used to qualify EI's conceptual worth; instead, they are typically rendered *redeemable* as artefacts of researchers' decisions rather than as signals about the construct itself (Theme 3, $n=64$, 88%). We describe and illustrate each theme and its sub-themes below.

3.2.1 Theme 1: prestigious

Nearly all *Discussion* sections ($n=71$, 97%) contain claims asserting EI's importance, causal power, or applied necessity in ways incongruent with the quality or nature of the evidence reviewed. This theme comprises four sub-themes that reflect qualitatively distinct forms of prestige claim-making.

The most pervasive form of Prestige was *cross-domain causal reach* (1a). This involves EI being positioned as a powerful and general determinant of life outcomes, typically without calibration to reported effect sizes, design limitations, or SIGN quality. In most reviews sampled, claims of this type span multiple domains in the same passage, projecting EI's influence across health, performance, clinical, and interpersonal outcomes simultaneously. For example, a Reject-rated meta-analysis states that: “*The literature clearly demonstrates that people with a higher level of EI have healthier, happier and more productive professional lives*” (Akerjordet and Severinsson 2007, p. 1410). That equivalent cross-domain summaries appear in Low-rated reviews (“*emotionally savvy individuals are not only high-performing... but are also more satisfied with their jobs*”; Miao et al. 2017b, p. 194) and, in

Table 3 Thematic summary of linguistic hype across discussion sections of N=73 emotional intelligence systematic reviews and meta-analyses

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distri- bution
			Theme 1: Prestigious.EI presented as important, causally powerful, and uniquely valuable, in ways incongruent with the quality or nature of the evidence reviewed	n = 70/73 (96%)	Reject: n=23 Low: n=34 Accept- able: n=9 High: n=4
	(1a) Cross-domain causal reach.EI positioned as a powerful determinant of life domains without calibration to effect sizes, design limitations, or SIGN quality rating	Bold T1	<p>Akejordet and Severinsson (2007), p. 1410: <i>Reject</i></p> <p>“The literature clearly demonstrates that people with a higher level of EI have healthier, happier and more productive professional lives”</p> <p>[Review of 16 unevaluated studies; no quality assessment; ‘clearly demonstrates’ presents as settled fact what cross-sectional correlational evidence cannot establish]</p> <p>Baudry et al. (2018), p. 217: <i>Low</i></p> <p>“[EI sub-dimensions] are associated with better mental, physical, and general health in both clinical and healthy populations”</p> <p>[42 primary studies, none quality-appraised; 39/42 cross-sectional; the claim spans three health domains from self-report correlational data only]</p> <p>Miao et al. (2017b), p. 194: <i>Low</i></p> <p>“Emotionally savvy individuals are not only high-performing... but are also more satisfied with their jobs”</p> <p>[Cumulative portrait combining claims from two separate meta-analyses (O’Boyle et al. 2011; this study), neither of which quality-appraised primary studies; substantial unexplained heterogeneity across EI types]</p>		

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distrib- ution
				n = 67/73 (92%)	Reject: n=23 Low: n=32 Accept- able: n=9 High: n=3
	(1b) Imperative application. EI presented as necessary for professional practice, clinical intervention, training, or policy; intervention and screening recommendations made on the basis of correlational or cross-sectional evidence	Bold T1	Baudry et al. (2018), p. 219: <i>Low</i> “ IEI is crucial for clinical interventions ” [No intervention studies included in the review; ‘crucial’ is a prescriptive imperative applied to a finding, from cross-sectional self-report correlations only] Kopp and Jekauc (2018), p. 18: <i>High</i> “ Practitioners should promote the implementation of EI screening and EI-development programs as an integral part of the training process ” [High quality review; however, screening and training recommendation derives from a small pooled effect ($r=0.16$, ~2.6% variance explained) in a predominantly cross-sectional evidence base; no training studies included] Llamas-Diaz et al. (2022), p. 935: <i>Acceptable</i> “ Based on the evidence, we recommend fostering children’s social and emotional skills within educational settings from early childhood ” [Quality assessment conducted but results not used to moderate claims; recommendation moves from correlational evidence to early childhood intervention prescription without causal evidence]	n = 26/73 (36%)	Reject: n=14 Low: n=10 Accept- able: n=2 High: n=0
	(1c) Ontological elevation. EI framed as a fundamental human capacity, core to professional identity, personal integrity, or existential well-being; the construct elevated beyond the scope of what any empirical review could establish	Bold T1	Akerjordet and Severinsson (2007), p. 1410: <i>Reject</i> “ Emotional intelligence is an essential phenomenon and one that enhances the quality of care from a humanistic existentialist perspective, which involves the ontology of human nature and agency ” [16 unevaluated studies, no quality assessment; grounds EI’s importance in the ontology of human nature. This is a philosophical prestige claim that no empirical review can warrant, and incongruent with the authors’ own acknowledgement that EI’s conceptual status remains unresolved] Encarnação et al. (2018), p. 7: <i>Low</i> “ For an effective and successful nursing, the nurse must be able to perceive, understand, regulate and enjoy their emotions ” [$r=0.04$ correlation between EI and job satisfaction in this review; mandatory professional imperative (‘must be able to’) applied from a near-zero correlation in 9 heterogeneous, unevaluated studies] Ogurlu (2021), p. 9: <i>Reject</i> “ Since success in life requires both head strengths and heart strengths ” [Small pooled effect ($g=0.12$) from 17 unevaluated studies with undated search; life-wisdom framing (‘success in life’) elevates findings to an existential claim incongruent with the evidence base]		

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(1d) Consensus assertion. The state of EI evidence characterised as settled, cumulative, or beyond dispute; contested findings and ongoing debates suppressed in fa- vour of a unified positive narrative	Bold T1	Miao et al. (2019a, 2019b), p. 194: <i>Reject</i> “Given the clear research findings, developing, promoting, and enhancing EI is a goal across multiple streams of research” [Opening statement used to contextualise Dark Triad findings; ‘clear research findings’ characterises the broader EI literature as settled, relying substantially on the authors’ own prior REJECT-rated meta-analyses; severe publication bias confirmed for the narcissism distribution in this study] Martins et al. (2010), p. 561: <i>Low</i> “No additional studies are needed to establish the existence of the phenomenon” [Forecloses future research on whether EI is a valid construct; from a Low quality review where primary study quality was not assessed and effect sizes were not calibrated to design limitations] Quiñez-Robres et al. (2023), p. 11: <i>Low</i> “EI must be understood as a predictor of socio-emotional wellbeing” [Prescriptive framing (‘must be understood’) applied to correlational findings; from a review where quality assessment results were not used to moderate conclusions]	n = 39/73 (53%)	Reject: n = 17 Low: n = 17 Accept- able: n = 5 High: n = 0
			Theme 2: Exaggerated. Acknowledged limitations and caveats overridden within the same passage or across the Discussion by confident, often applied claims; null and weak findings absorbed rather than treated as informative		

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(2a) Hedging fallacy. A stated caveat (limitations, small effect size, causal indeterminacy) is immediately followed or preceded by a confident claim that functionally overrides it; the hedge is structurally present but rhetorically ineffective	Underline T2	<p>Andrei et al. (2016), p. 272: <i>Low</i> “Although small, the overall effect size confirms the distinctiveness and theoretical importance of trait EI” <i>[‘Although small’ precedes a claim that $\Delta R^2=0.06$ ‘confirms’ theoretical importance; confirmed publication bias (Egger’s $z=4.78$, $p<0.001$) and 94% criterion overlap mean the pooled effect likely overestimates true incremental validity; the hedge does not constrain the confirmatory claim]</i> Akerjordet and Severinsson (2007), p. 1408: <i>Reject</i> “In general EI research exhibits promising results, regardless of critical comments” <i>[‘Regardless of critical comments’ explicitly discounts the critical literature as a class; 16 unevaluated studies with no quality assessment cannot sustain a positive general characterisation of the EI evidence base as ‘promising’]</i> Cherry et al. (2013), p. 186: <i>Low</i> “Despite the included studies’ methodological limitations, the findings lend support for the value of applying... EI to the study of patient-centred care” <i>[‘Despite’ nominally acknowledges limitations but the application claim that follows is not qualified in kind; null finding on EI–patient satisfaction was the primary result of this review]</i></p>	n = 69/73 (95%)	Reject: n = 23 Low: n = 33 Accept- able: n = 9 High: n = 4

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
(2b) Null findings as anomalous. Absence of effect characterised as surprising, puzzling, or a temporary artefact relative to an advocacy literature, rather than treated as informative evidence about EI's predictive limits		Underline T2	<p>Cherry et al. (2013), p. 183: <i>Low</i> “A surprising finding given the growth in review pieces and editorials advocating the importance of EI for developing the patient–doctor relationship” <i>[Null empirical result characterised as ‘surprising’ relative to advocacy literature; this implicitly treats advocacy-based claims as a credible baseline against which an empirical null is anomalous rather than informative]</i></p> <p>Harms and Credé (2010a, 2010b), p. 12: <i>Reject</i> “Although there have been a number of studies conducted assessing the role of EI in transformational leadership, very few have actually been conducted using each of the different measures of EI... As newer EI assessment tools are developed and older tests are refined with criteria prediction in mind, it could be expected that validities will improve” <i>[Converts weak validity estimates into a field-maturity problem rather than acknowledging they may reflect genuine construct limitations ‘it could be expected’ frames future positive results as theoretically predicted]</i></p> <p>Gutiérrez-Cobo et al. (2016), p. 10: <i>Reject</i> “[Cool task null results] may reflect the situation that EI favours cognitive performance only when it has emotional information” <i>[Null finding on cool cognitive tasks which could indicate EI is not a general cognitive ability is reframed as evidence of construct specificity; converts disconfirmatory result into a theoretical asset without evidential basis in a Reject-rated review with no quality appraisal]</i></p>	<p>n = 41/73 (56%)</p>	<p>Reject: n = 16 Low: n = 20 Acceptable: n = 4 High: n = 1</p>

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(2c) Causal escalation. Correlational findings restated as directional, causal, or predictive without acknowledged transition; the move from association to mechanism treated as unremarkable and unqualified	Underline T2	<p>Miao et al. (2016), p. 21: <i>Reject</i> “Emotionally intelligent leaders will produce satisfied followers” <i>[‘Will produce’ applies causal deterministic language to a correlational meta-analysis ($\rho = 0.31$); $k = 20$ samples with no primary study quality appraisal; 10% of variance explained by statistical artefacts indicates substantial unmeasured confounding]</i></p> <p>Megias et al. (2018), p. 202: <i>Low</i> “If we help these individuals to develop their EI, we will stimulate the quality of their interpersonal relationships, facilitating their adaptation to society” <i>[Causal intervention chain (‘if we... we will stimulate’) applied from cross-sectional correlational data; no intervention or longitudinal studies cited to support the directional claim]</i></p> <p>Jiménez-Picón et al. (2021), p. 10: <i>Low</i> “EI in general was shown to increase significantly after practicing mindfulness” <i>[Causal directionality (‘increase after’) applied to a correlation between mindfulness practice and EI scores; no quality appraisal of included studies; ‘was shown to increase’ implies experimental evidence for a finding from observational data]</i></p>	n = 54/73 (74%)	<p>Reject: n = 21</p> <p>Low: n = 27</p> <p>Acceptable: n = 5</p> <p>High: n = 1</p>

Theme 3: Redeemable. Evidential weaknesses, be that methodological, conceptual, or scope-based, are attributed to remediable artefacts of researchers’ decisions rather than to the construct itself, protecting EI’s theoretical and applied status

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(3a) Measurement methodologism. <i>Validity threats, null results, or inconsistent findings attributed to inadequate measurement instruments; construct validity protected by attributing failures to the tools rather than to the underlying construct</i>	<i>Italic</i> T3	<p>Andrei et al. (2016), p. 274: <i>Low</i> “The confident conclusion is that trait EI reliably accounts for substantial variation in a wide range of criteria that is not accounted for by other constructs” <i>[Follows acknowledged publication bias (Egger’s $z=4.78$, $p<0.001$), common-method variance, and 94% criterion overlap; these are measurement concerns that protect EI’s construct status rather than being used to constrain conclusions; $\Delta R^2=0.06$ is described as ‘substantial variation’]</i> Llamas-Diaz et al. (2022), p. 934: <i>Acceptable</i> “To overcome these limitations, we must develop more instruments with execution capacity” <i>[Measurement heterogeneity attributed to inadequate tools rather than construct ambiguity; redirects a potential validity problem into a technical development agenda, preserving EI’s status as a coherent target for better measurement]</i> Arora et al. (2010), p. 761: <i>High</i> “Without agreed definitions and measures, the medical education community will find it hard to link individuals’ characteristics with competency development” <i>[Nine different EI tools used across 14 studies in a High quality review; failure to map EI onto theoretical framework (ACGME competencies) attributed to measurement inconsistency rather than construct incoherence; reframes a potential construct validity problem as a standardisation task]</i></p>	n = 61/73 (84%)	Reject: n = 20 Low: n = 29 Acceptable: n = 8 High: n = 4

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(3b) Design methodologism. Cross- sectional dominance, shared-method variance, or small-sample limita- tions acknowledged but immediately redirected to future research agendas rather than used to con- strain current claims	<i>Italic</i> T3	Baudry et al. (2018), p. 219: <i>Low</i> “A meta-analysis will be necessary in order to provide a definitive consensus across the literature about these effects” <i>[Meta-analysis framed as a step toward confirming already-asserted effects rather than as a method for testing whether effects exist; three of 42 studies were longitudinal; cross-sectional limitation does not prevent applied intervention recommendations appearing throughout]</i> MacCann et al. (2020b), p. 173: <i>Low</i> “EI training programs are likely to increase academic performance” <i>[Immediately follows the authors' own acknowledgement that no causal evidence exists in this review; 'likely to increase' projects a causal outcome from correlational data, using probabilistic language to advance a training claim without evidential warrant]</i> Baudry et al. (2018), p. 219: <i>Low</i> “Future studies should employ longitudinal designs to support the predictive impact of EI” <i>[To support the predictive impact' frames longitudinal research as a confirmatory step for an established effect rather than as investigation of whether a predictive relationship exists; presupposes the answer before the test is conducted]</i>	n = 64/73 (88%)	Reject: n = 21 Low: n = 29 Accept- able: n = 9 High: n = 5

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(3c) Speculative future adequacy. Null or weak results attributed to untested moderators, unexplored populations, or unexamined sub-facets; EI's relevance preserved by projecting support into hypothetical future contexts not yet studied	<i>Italic</i> T3	<p>Andrei et al. (2016), p. 274: <i>Low</i> “Even though sociability and emotionality did not emerge as strong predictors in our meta-analysis, it is reasonable to expect that they will assume salience when examined in relation to criteria that are more social in nature” <i>[Two non-predicting factors redeemed by asserting salience in a hypothetical future context (‘criteria that are more social in nature’) not yet studied; converts a disconfirmatory finding into a deferred prediction; confirmed publication bias and Low SIGN quality mean the null results are themselves from an imperfect evidence base]</i></p> <p>Gutiérrez-Cobo et al. (2016), p. 10: <i>Reject</i> “[Cool cognitive task null results] may reflect the situation that EI favours cognitive performance only when it has emotional information to go on” <i>[Null result on cool tasks—9 studies—attributed to a domain-specificity hypothesis that is untested in the review; ‘may reflect’ preserves EI’s cognitive relevance by deferring null interpretation to a future context with emotionally loaded stimuli; only nine cool-task studies against 35 hot-task instances, making this asymmetry a permanent feature of the literature rather than a temporary data gap]</i></p>	n = 57/73 (78%)	<p>Reject: n = 19 Low: n = 28 Accept- able: n = 8 High: n = 2</p>

Table 3 (continued)

Theme	Sub-theme	Code	Illustrative example quotes with evidential annotations	Frequency n reviews (N=73)	SIGN quality distribu- tion
	(3d) Construct redefinition. <i>EI</i> relabelled, <i>narrowed</i> , or <i>re-theorised</i> when evidence is weak or construct validity is questioned; the construct's scope adjusted <i>post-hoc</i> to preserve its relevance when challenged by inconsistent findings	<i>Italic</i> T3	Cherry et al. (2013), p. 186: <i>Low</i> “ <i>If EI can be reframed as the ‘sensitive and intelligent problem-solving activities emerging from deliberate, structured group learning’, then it is not difficult to see the relevance of EI to... medical education</i> ” [<i>EI's poor discriminant validity from personality and its measurement heterogeneity are acknowledged, then the construct is redefined using external literature (Lewis et al.) to preserve its applicability in medical education; reframing a validity problem as a matter of conceptual emphasis rather than a reason for caution</i>] van der Linden et al. (2017), p. 44: <i>Reject</i> “ <i>trait EI strongly overlaps and is perhaps substantively synonymous with the GFP [general factor of personality]</i> ” [<i>Near-total overlap with the general factor of personality reframed as theoretical integration rather than discriminant validity failure; ‘substantively synonymous’ converts a construct validity threat into a theoretical contribution, preserving EI’s conceptual territory by folding its redundancy with personality into a broader definitional claim</i>] Van Rooy et al. (2005), p. 458: <i>Reject</i> “ <i>the two models of EI are not mutually exclusive</i> ” [<i>Follows finding of $p=0.14$ (80% CI including zero) between mixed and ability EI models; ‘not mutually exclusive’ minimises near-orthogonality, reframing what could constitute a jingle fallacy (two different constructs sharing a label) as partial theoretical overlap warranting continued unified EI research</i>]	n=31/73 (42%)	Reject: n=13 Low: n=14 Accept- able: n=4 High: n=0

Note: Findings are based on thematic analysis of Discussion sections across the full $N=73$ corpus of SIGN-rated systematic reviews and meta-analyses on emotional intelligence. SIGN quality distribution: High (++), $n=5$; Acceptable (+), $n=9$; Low (-), $n=36$; Reject (0), $n=23$. Frequency counts reflect the number of coded reviews in which each sub-theme was identified (one or more coded passages); they do not reflect passage counts. Illustrative quotes are drawn from Discussion or Conclusions sections of the source review. Square-bracketed annotations following each citation identify the specific evidential context that establishes coding incongruence, including SIGN quality anchors, effect size calibration, acknowledged limitations, and structural rhetorical features. Co-occurring codes are reported within the sub-theme row where the dominant rhetorical function is located. T1 = Prestigious; T2 = Exaggerated; T3 = Redeemable. Two reviews showed no coded instances for Theme 1 (Miao et al. 2018b; Sarrionandia and Mikolajczak 2020). Four reviews showed no coded instances for Theme 2 (Lea et al. 2019; Miao et al. 2018b; Toriello et al. 2022; Walker et al. 2022). Nine reviews showed no coded instances for Theme 3 (Kotsou et al. 2019; Peterson et al. 2011; Ramjbar and Bakhshi 2018; Pirsoul et al. 2023; Vega et al. 2021; Malouff et al. 2014; Encarnação et al. 2018; Domínguez-García and Fernández-Berrocal 2018; Laborde et al. 2016). All remaining reviews activated at least one sub-theme across each of the three themes

attenuated form, in several Acceptable and High-rated reviews, suggests how common the *Prestige* frame is across the corpus.

Imperative application claims (1b) were present in 67 reviews (92%), and translate EI's asserted causal reach directly into professional, clinical, or policy prescriptions. These recommendations (e.g., for training programmes, screening processes, or curricular integration) are routinely made based on correlational or cross-sectional evidence, and their linguistic form (use of “*must*,” “*should*,” “*crucial*,” and “*need*”) conveys a sense of professional obligation rather than conditional inference. Crucially, this pattern was not confined to the lowest-quality reviews: a High-rated meta-analysis recommended that “*practitioners should promote the implementation of EI screening and EI-development programs as an integral part of the training process*” (Kopp and Jekauc 2018, p. 18) from a corpus where the effect size was labelled a “weak determinant” of sports performance in the same *Discussion*. An Acceptable-rated review concluded that “*based on the evidence, we recommend fostering children's social and emotional skills within educational settings from early childhood*” (Llamas-Díaz et al. 2022, p. 935), a recommendation extending beyond both the reviewed age range (adolescents, 10–19) and the correlational scope of the evidence. Given that unfounded recommendations for clinical practice represent the most severe form of misrepresentation (Yavchitz et al. 2016), it is notable that our analysis found these claims present in nine in ten of our coded *Discussions* (67/73, 92%), across all SIGN quality levels—albeit less frequently in High-rated reviews (3/5, 60%) than in reviews of lower methodological quality.

Ontological elevation (1c) identifies the framing of EI as a fundamental attribute. This was less prevalent than the preceding sub-themes ($n=26$, 36%) and was concentrated in Reject- and Low-rated reviews. The most pronounced instances appeared in reviews drawing explicitly on humanistic or professional identity frameworks: “*Emotional intelligence is an essential phenomenon and one that enhances the quality of care from a humanistic existentialist perspective, which involves the ontology of human nature and agency*” (Akerjordet and Severinsson 2007, p. 1410, SIGN Reject); “*for an effective and successful nursing, the nurse must be able to perceive, understand, regulate and enjoy their emotions*” (Encarnação et al. 2018, p. 7, SIGN Low). Notably, no High-rated review in our sample contained an instance of ontological elevation.

Consensus assertion (1d), present in 39 reviews (53%), involves characterising the state of EI research as settled, cumulative, or beyond dispute, in ways that pre-empt critical scrutiny. In our corpus, its most striking instances appear in reviews by authors who cite their own prior meta-analyses (included in the $N=73$ corpus) as the evidential basis for EI's established importance: “*given the clear research findings, developing, promoting, and enhancing EI is a goal across multiple streams of research*” (Miao et al. 2019a, b, p. 194, SIGN Reject). Choosing to use “clear” in a sentence that opens the *Discussion* of a new meta-analysis, implies a settled consensus where heterogeneous and contested evidence exists. Similarly, Martins et al. (2010, p. 561, SIGN Low) concluded that “*no additional studies are needed to establish the existence of the phenomenon*”—a directive-framing using meta-analytic stability indicators as substantive closure of an empirical question. A self-referential scaffolding pattern, in which prior low-quality reviews are used to anchor a current review's claims, was particularly pronounced in clustered research programmes in our sample (e.g., the Miao et al. series, the Fernández-Berrocal group papers), and repre-

sents a form of citation-based prestige amplification that operates independently of primary evidence quality.

3.2.2 Theme 2: exaggerated

All but four coded *Discussions* ($n=69$, 95%) contained passages in which acknowledged limitations, caveats, or weak findings, were incongruous with confident claims made in the same passage, or across the *Discussion* as a whole. This theme is characterised by the structural incongruence between *stated epistemic constraint* and *claimed epistemic authority* and can be further elaborated through three sub-themes.

The *hedging fallacy* (2a) is the single most pervasive form of LH in the corpus, present in 69 reviews (95%) and the sub-theme with the most frequent presence in High-rated reviews (4 of 5, 80%). Its defining structure involves a stated caveat (e.g., acknowledging small effect sizes, causal uncertainty, cross-sectional design) that is immediately overridden by a strong inferential or applied claim. In other words, the hedge is structurally present but rhetorically inert. For instance, a Low-rated meta-analysis reads: “*although small, the overall effect size confirms the distinctiveness and theoretical importance of trait EI*” (Andrei et al. 2016, p. 272), where the concessive “although small” functions as a pro-forma acknowledgement before the evaluative claim proceeds unqualified. The concessive hedge is structurally present as a subordinate clause, whose function is to *inoculate* the main claim against the very caveat being named. A Reject-rated review provided one of the clearest examples: “*in general EI research exhibits promising results, regardless of critical comments*” (Akerjordet and Severinsson 2007, p. 1408). This single sentence simultaneously registers the critical literature and dismisses its force. In High-rated reviews, the hedging fallacy took more constrained forms, typically emerging in Conclusion sections after appropriately hedged main discussions, illustrating that quality appraisal reduces but does not eliminate the pattern. For instance, having documented mixed and null findings with appropriate circumspection throughout their Discussion, Lea et al. (2019, SIGN High) concluded that participants with higher EI “*do seem to recover more quickly from the ordeal, regardless of how EI or stress is measured*” (p.16), which indicates a mild ‘universalisation’ that sets aside both the measurement heterogeneity and the small study base ($k=6$) that had been previously acknowledged. Across quality tiers, these hedges are therefore ‘freeing’ rather than constraining of central claims: by appearing to register doubt, they make the confident conclusion that follows seem more, not less, trustworthy.

The *null findings as anomalous* sub-theme (2b) was present in 41 reviews (56%). This involves the treatment of absent or inconsistent effects as puzzling exceptions, rather than as informative findings about EI’s predictive limits. Cherry et al. (2013, p. 185, SIGN Low) characterised a null finding for EI and patient-centred care as “*a surprising finding given the growth in review pieces and editorials advocating the importance of EI for developing the patient–doctor relationship*” (p. 185). This structure suggests the advocacy literature as the epistemic baseline against which empirical data are measured, inverting the normal direction of evidence-to-claim inference. In a Reject-rated review, null results for self-report EI and cognitive tasks were absorbed by claiming that performance-based EI results had emerged “*as expected,*” selectively foregrounding the confirming stream while treating the disconfirming stream as anomalous (Gutiérrez-Cobo et al. 2016, p. 9). This sub-theme was

less prevalent in High-rated reviews, where null or partial findings were more commonly reported without evaluative framing.

Causal escalation (2c), present in 54 reviews (74%), involves the re-statement of correlational or associational findings in explicitly directional or causal language, without any acknowledged transition and without the caveat-then-override structure of the hedging fallacy. Here, causation is simply assumed. Reject-rated reviews produced the most direct instances: “*emotionally intelligent leaders will produce satisfied followers*” (Miao et al. 2016, p. 21) converts a cross-sectional correlation with no quality-appraised primary studies directly into a deterministic personnel outcome. Low-rated reviews provided clinical equivalents: “*if we help these individuals to develop their EI, we will stimulate the quality of their interpersonal relationships, facilitating their adaptation to society*” (Megías et al. 2018, p. 202). Notably, causal escalation also appeared in Acceptable-rated reviews using the pre-test/post-test intervention literature as a bridging mechanism: “*EI in general was shown to increase significantly after practicing mindfulness*” (Jiménez-Picón et al. 2021, p. 10), thus treating a heterogeneous set of predominantly uncontrolled pre-post designs as establishing causal efficacy. This form of escalation, in which correlational evidence and adjacent intervention evidence are implicitly merged to generate causal claims, was a recurring structural feature of the corpus across all SIGN quality levels.

3.2.3 Theme 3: redeemable

Theme 3 was present in 64 reviews in our sample (88%). Where methodological weaknesses were acknowledged, the dominant pattern was one of *rhetorical redirection*, rather than to qualify EI's conceptual status or to constrain applied claims: limitations were acknowledged then immediately recast as fixable problems awaiting methodological solutions, preserving EI's standing as a construct whose importance is established in principle and whose fuller validation is contingent only on researcher effort. This theme comprises four sub-themes.

Measurement methodologism (3a) was present in 61 reviews in our sample (84%). This involves attributing null results, inconsistent findings, or construct validity concerns to the inadequacy of measurement instruments, rather than to questions about the underlying construct. In many cases, fundamental conceptual problems (e.g., whether different EI operationalisations measure the same or related things, or whether self-report and ability measures tap genuine ‘emotional competence’), are treated as instrumentation gaps requiring technical solutions. A Low-rated meta-analysis exemplifies this: having acknowledged publication bias (Egger's $z=4.78$, $p<.001$), common-method variance, and 94% criterion overlap, the same paper concluded that “*the confident conclusion is that trait EI reliably accounts for substantial variation in a wide range of criteria*” (Andrei et al. 2016, p. 274). This structural sequence in claim-making—to acknowledge a limitation, then set aside to restore confidence—was replicated across the corpus at all quality levels, including in High-rated reviews where standardisation of measurement was recommended as the pathway to resolving inconsistencies (Arora et al. 2010), without engaging with whether inconsistency might reflect construct-level rather than measurement-level problems.

Design methodologism (3b) was the most prevalent sub-theme within Theme 3 ($n=64$, 88%), reflecting the frequent presence of cross-sectional and shared-method-variance limitations in the reviewed literature. Its defining pattern involves the acknowledgement of a design limitation (typically that cross-sectional data cannot establish causation, or that

common method variance inflates correlations), followed by its redirection into a future research agenda, such that the limitation constrains future methodological choices but not present-tense claims. For example, Baudry et al. (2018, p. 219, SIGN Low) recommended that “*a meta-analysis will be necessary in order to provide a definitive consensus across the literature about these effects*”. This formulation frames meta-analysis as a device to confirm effects already treated as established, rather than to test whether they exist. MacCann et al. (2020a, b, p. 173, SIGN Low) concluded that “*EI training programs are likely to increase academic performance*” immediately following an explicit acknowledgement that only three longitudinal studies existed in the meta-analytic database and that causal inference was not possible. This is a particularly clear example of the design-methodologism pattern in the reviews analysed because the causal limitation and the causal conclusion appear within a few sentences of each other. This sub-theme was present in all SIGN quality bands, including all 5 High-rated reviews, demonstrating that even well-conducted reviews routinely redirect design limitations rather than treating them as constraints on interpretive confidence.

Speculative future adequacy (3c) was present in 57 reviews (78%) and involves the attribution of null or weak results to untested moderators, unexplored populations, or unexamined sub-facets, with the implication that EI's relevance will emerge once the relevant conditions are examined. In our sample, this protective manoeuvre converts disconfirmatory evidence into an invitation for further research, rather than treating it as informative about the construct's limits. Andrei et al. (2016, p. 274, SIGN Low) addressed two non-predicting meta-analytic factors by claiming that “*it is reasonable to expect that they will assume salience when examined in relation to criteria that are more social in nature*”, asserting, post-hoc, that the factors will matter in a different, as-yet-untested domain. Gutiérrez-Cobo et al. (2016, p. 10, SIGN Reject) interpreted null results for cool cognitive tasks by proposing that EI “*favors cognitive performance only when it has emotional information to go on,*” converting a null finding into a theoretical specificity claim that simultaneously narrowed EI's scope and protected it from disconfirmation.

Construct redefinition (3d) was present in 31 reviews (42%) and more so in lower quality reviews sampled. This involves the post-hoc relabelling or re-theorising of EI when evidence is weak or construct validity is challenged. Unlike measurement methodologism, where the problem is located in tools, construct redefinition adjusts the conceptual boundaries of EI itself. Cherry et al. (2013, p. 186, SIGN Low) responded to evidence of EI–personality overlap by proposing that “*if EI can be reframed as the 'sensitive and intelligent problem-solving activities emerging from deliberate, structured group learning', (Lewis et al. [75], p. 351), then it is not difficult to see the relevance of EI to both undergraduate and postgraduate medical education*”. This is a redefinition that moves EI so far from its original operationalisation as to render the review's own data tangential to the new concept. After reporting a mixed-ability EI correlation of only $\rho=0.14$ (which indicates near-orthogonality of the two main EI models), Van Rooy et al. (2005, p. 458, SIGN Reject) reframed the finding by asserting that mixed and ability EI “*are not mutually exclusive*” as if partial overlap constituted theoretical coherence. Van der Linden et al. (2017, p. 44, SIGN Reject) took construct redefinition to its furthest point in the corpus, characterising the very high GFP-trait EI correlation as evidence that EI “*is perhaps substantively synonymous with the general factor of personality*”, thus converting a discriminant validity failure into a theoretical unification claim.

Table 4 Prevalence of linguistic hype sub-themes by sign methodological quality band across N = 73 emotional intelligence systematic reviews and meta-analyses

Sub-theme	Total n (%)	Reject (n = 23)	Low (n = 36)	Acceptable (n = 9)	High (n = 5)
Theme 1: Prestigious					
<i>1a</i> Cross-domain causal reach	70 (96%)	100%	94%	100%	80%
<i>1b</i> Imperative application	67 (92%)	100%	89%	100%	60%
<i>1c</i> Ontological elevation	26 (36%)	61%	28%	22%	0%
<i>1d</i> Consensus assertion	39 (53%)	74%	47%	56%	0%
Theme 2: Exaggerated					
<i>2a</i> Hedging fallacy	69 (95%)	100%	92%	100%	80%
<i>2b</i> Null findings as anomalous	41 (56%)	70%	56%	44%	20%
<i>2c</i> Causal escalation	54 (74%)	91%	75%	56%	20%
Theme 3: Redeemable					
<i>3a</i> Measurement methodologism	61 (84%)	87%	81%	89%	80%
<i>3b</i> Design methodologism	64 (88%)	91%	81%	100%	100%
<i>3c</i> Speculative future adequacy	57 (78%)	83%	78%	89%	40%
<i>3d</i> Construct redefinition	31 (42%)	57%	39%	44%	0%

Note: Cells show the percentage of reviews within each SIGN quality band in which the sub-theme was identified (one or more coded passages). Shading encodes prevalence within band: white=0%; very light grey=1–25%; light grey=26–50%; medium grey=51–75%; dark grey=76–100%. SIGN quality band sizes: Reject (0) n=23; Low (-) n=36; Acceptable (+) n=9; High (++) n=5. Total column reports number of reviews (N=73) and corpus-level percentage. Sub-theme codes are italicised in the first column

3.2.4 Thematic patterns in relation to SIGN quality

Across the corpus, we observed a consistent *gradient* in the number of coded passages per *Discussion* (coding density) according to SIGN quality levels (see Table 4 for heat map). Reject-rated reviews exhibited the highest density and High-rated reviews the lowest. However, we also noted three cross-cutting patterns.

First, all three themes, and most sub-themes, were present across all SIGN quality levels. Four of the five high-rated reviews contained coded instances for Theme 1, and all five for Theme 2, demonstrating that LH is not simply a function of methodological weakness. Rather, it appears to be a feature of the genre of EI review writing more broadly in the corpus, with quality determining its *form* and *frequency* rather than its *presence*. In High-rated reviews, LH tended to emerge at specific structural junctures; predominantly in *Conclusion* sections and applied *Recommendations* paragraphs instead of being distributed throughout the *Discussion*. This pattern suggests that even reviews that handle evidence responsibly in

their main analytical sections may revert to prestige-claiming language when summarising implications.

Second, the dominant coding pattern across the corpus was the co-occurrence of Themes 1 and 2 (T1+T2), in which prestige claims and hedging fallacies appear in the same passage. This was most concentrated in applied and clinical *Recommendations* sections. This suggests that linguistic hype in this corpus most commonly operates through a *rhetorical sequence* rather than through isolated prestige claims or isolated hedging failures. The typical structure is: (a) acknowledge a limitation or small effect; (b) assert that EI's importance is confirmed despite this; (c) recommend training, clinical application, or policy implementation. This sequence was reproduced with near-formulaic consistency across reviews spanning four decades, multiple domains, and the full range of SIGN quality ratings.

Third, theme 3 (Redeemable) was widespread in review *Discussions*, appearing in all but nine of the sampled reviews (Kotsou et al. 2019; Peterson et al. 2011; Pirsoul et al. 2023; Ranjbar and Bakhshi 2018; Vega et al. 2021; Laborde et al. 2016; Malouff et al. 2014; Encarnação et al. 2018; Domínguez-García and Fernández-Berrocal 2018). This indicates that the *redirection of limitations as future research opportunities* is a near-universal feature of EI review writing in our sample, functioning less as a response to specific evidentiary challenges and more as a structural convention. The nine reviews that did not employ redeemable framing were notably characterised by more extensive methodological self-critique and for naming construct-level rather than only measurement-level problems, suggesting that more analytically rigorous interpretation of findings is detectable within the field, even if rarely deployed.

4 Discussion

Using a culturally disfluent, critical lens (Oyserman 2023), we examined the nature of claim-making about EI, by exploring whether poor methodological quality and LH could be detected in a sample of 73 systematic reviews (SRs) and meta-analyses (MAs) on the topic. We used a meta-synthesis methodology to first assess the methodological rigour of the corpus of reviews published between 2004 and April 2024, and second to assess the use of LH in claim-making in these reviews. In line with Yavchitz et al. (2016), our assessment of methodological quality suggests the presence of *misleading interpretation* (e.g., formulating recommendations for practice not supported by the findings; inadequate consideration of the risk of bias of primary studies included in results interpretation) and *inappropriate extrapolation* (i.e., inadequate extrapolation of the results to a larger population, a larger setting or a wider set of interventions; conclusion extrapolating the review's findings to a different population or setting). Further, we show how LH in the *Discussion* sections of the reviews produces a consistent pattern of claim-making about EI's value that is materially incongruent with the quality of the reviews and evidence they summarise. Together, these analyses show methodological issues and misrepresentation in the interpretation and reporting of findings in the reviews sampled.

4.1 Methodological rigour in EI reviews

Of the 73 SRs included in our sample, 42 provided statistical summaries of EI effects drawn from primary research. Meta analytical techniques are seen by some as a solution for correcting EI measurement issues, who cite these reviews as evidence of field-wide advancement in volume, rigour and ‘practically meaningful’ EI effect sizes (e.g., Dasborough et al. 2022). However, our analysis illustrates how variation in EI measures *and* their treatment within meta-analyses, renders theoretically meaningful interpretations challenging. This suggests that, potentially, EI meta-analyses may distract audiences from unresolved fundamental issues relating to the lack of coherency in EI theory, and that a singular, unified “EI” construct does not exist. Thus, discussing “EI” in general terms as a superordinate entity is misleading, and produces an evidentiary landscape that potentially facilitates ‘spin’. It is notable that in our quality analysis, only 5 of 73 reviews analysed were deemed ‘good’, with a low risk of bias (Arora et al. 2010; Cook et al. 2016; Kopp and Jekauc 2018; Lea et al. 2019; Sarrionandia and Mikolajczak 2020), and these higher quality reviews discuss heterogeneity in defining and measuring EI as a barrier to qualifying trends. 81% (59/73) of the systematic reviews in our sample were classified as ‘poor’ or ‘rejectable’ quality, with evidence of misrepresentation—that is, *misleading interpretation* and *inappropriate extrapolation*—common in these reviews. These forms of misrepresentation appear to be pervasive and sustained, rather than ‘historical’ artefacts located within older reviews and research practices, given poor and rejectable quality reviews were published very recently in our sample (up to 2023). Consistent with other analyses from psychology and health (e.g., Hohn et al. 2020; Shercliffe et al. 2009), we find EI review authors within our sample often fail to report key aspects of the systematic search process (e.g., selection, extraction, exclusions), and neglect to include unpublished literature, or to report and interpret the quality appraisals of primary studies. Hohn et al. (2020) suggests that some of these omissions may reflect a failure in reporting, rather than doing, e.g., authors may conduct quality checks of primary literature but may be unaware of the need to explicitly report this, given they report other aspects relating to methodology (e.g., study design elements, exclusion criteria)—albeit one would expect the journal peer review process to mitigate such reporting oversights. In our sample of SRs, this appears an unlikely explanation as 90% of the reviews did not consider methodological quality when formulating conclusions and recommendations (i.e., recommending caution, excluding poor quality studies). Moreover, there appeared an incongruence between evidence and claims; poor quality reviews (i.e., with multiple sources of bias) were more likely to make a greater number of recommendations for advancing understanding and application of EI per se, and particularly ‘higher stakes’ claims (i.e., calls for training interventions; to use EI as a screening tool). Nevertheless, these reviews reached a wide audience (based on citation count). Overall, our quality analysis shows that the reproducibility, transparency and validity, of the majority of EI SRs in our review sample are questionable.

4.2 LH in EI reviews

Irrespective of the methodological quality, our CRTA of the reviews sampled found that authors use of LH in their *Discussions* is pervasive and structurally consistent, through prestige claiming (Theme 1: Prestigious), hedging fallacies (Theme 2: Exaggerated), and

the reframing of construct-level limitations as methodological artefacts (Theme 3: Redeemable). All three themes—*Prestigious*, *Exaggerated*, and *Redeemable*—were identified in 59 of 73 coded reviews (81%), and the co-occurrence of prestige claims with hedging fallacies (T1 + T2) was the dominant coding pattern across the corpus, concentrated most heavily in applied and clinical recommendations sections. This claim-making collectively positions EI as a prestigious and powerful predictive trait. Across academia, authors are increasingly using rhetorical resources such as ‘boosters’, and positive attitude markers, to highlight contribution, novelty, and project disciplinary authority, ostensibly in response to competitive pressures and self-serving needs to ‘sell’ research for visibility, citations, funding, and institutional success (Hyland and Jiang 2021). However, our themes suggest that prestige claims in EI reviews are markedly *construct* protective and promotional. EI is positioned as an entity of established scientific standing, independent of the quality of the research presented to the reader.

Three patterns in the distribution of LH across quality levels are notable. First, Rejected reviews exhibited the highest coding density and High-rated reviews the lowest, a *coding gradient* that confirms methodological rigour can attenuate LH in frequency and form. However, the presence of T1 (80%) and T2 (100%) instances in the five High-rated reviews demonstrates that LH is not simply a consequence of poor methodology: it seems to be a genre-level feature of EI review writing across the corpus. Even in reviews where heterogeneity and effect size limitations are carefully discussed, prestige-claiming and imperative application language tended to re-emerge in *Conclusions* and *Recommendations* paragraphs. This is consistent with Hyland and Jiang (2024) who observe this same pattern and attribute this to the growing pressure on researchers to generate ‘impactful’ findings, by demonstrating ‘value for money’ and translational benefits to society, to meet the goals of research funders and research assessment exercises. Indeed, authors from our cohort outlined how government policy and commitments to develop EI in educational settings (as socio-emotional learning) had driven research in the field (Llomas-Díaz et al. 2022, p. 934), creating additional external incentive for prestige-claiming regardless of review quality.

Second, the dominant coding pattern across the corpus was the co-occurrence of Themes 1 and 2 (T1 + T2): a sequence with rhetorical potential, in which a limitation or small effect is acknowledged, then functionally overridden by a confident claim that EI's importance is confirmed nonetheless, before a recommendation for training, screening, or application is made. In our sample, this sequence appeared with near-formulaic consistency across 4 decades of reviews, multiple outcome domains, and all SIGN quality levels, suggesting it is a discursive convention embedded in this literature rather than an idiosyncratic choice. Hyland (1998) describes hedges (signalling uncertainty and caution) and boosters (asserting certainty) as typically operating as opposing forces that calibrate an author's epistemic commitment to reflect the evidence at hand. The T1 + T2 pattern in our data, however, shows an apparent ‘*hedge-override*’ structure, in which the hedge appears not as a genuine epistemic constraint, but as a pro-forma acknowledgement whose function is to inoculate the confident claim that follows, making the claim that succeeds it seem more, not less, warranted. The rhetorical effect is that expressed doubt lends credibility to the confident claim that follows, rather than constraining it, contrary to the conventional function of hedging (Hyland 1998).

Third, the reframing of limitations as methodologically redeemable (theme 3), was common across the corpus (64/73, 88%) absent only in those reviews that engaged in extensive construct-level self-critique. This pattern means that the field's conceptual vulnerabilities

(including construct heterogeneity, measurement tautology, absence of causal designs), are routinely preserved, rather than interrogated, as authors redirect attention towards methodological refinement. This is consistent with Millar et al. (2019) who identify the reframing of study limitations as opportunities for future research as a characteristic promotional move in scientific *Discussion* sections, that serves to protect the credibility of the research programme rather than to interrogate its foundations. In the EI reviews, this operates at the construct level: by redirecting limitations into a methodological research agenda (better measures, larger samples, more diverse populations), authors preserve the conceptual prestige of EI while deferring the harder epistemological question of whether the field's foundational assumptions are warranted. In highlighting these caveats and then minimising their seriousness through boosted claims, authors maintain the prestige of EI.

From our analysis of the sampled reviews, we suggest these three themes may be collectively underpinned by a positivity bias that serves to maintain the prestige of EI and communicate its value to readers, even where the reviewed evidence does not warrant this. Their seemingly emphatic claim-making constructs EI as powerful, dominant, and a causal determinant of human behaviour, which thereby *always* warrants application in some form (i.e., to new settings, new groups of people, new measurement approaches). By explaining incongruous or unclear findings as a symptom of method (e.g., heterogeneity in measures, inadequate designs, or sampling), readers of these reviews may be distracted from understanding the issues that question the fundamental nature of the construct, and remain unresolved since early critical commentaries (Matthews et al. 2004). This is another example of how discussing “EI” in general terms as a superordinate entity is misleading and provides an evidentiary landscape that potentially facilitates ‘spin’ among consumers of these reviews. We note that this issue is not unique to the field of EI research—a preference for establishing “synthetic certainties” over real (paradigmatic) theory development was noted as an issue applicable to many ‘soft’ psychological fields by Meehl (1978) and Proulx and Morey (2021, p. 672). This can project the illusion of progress and that earlier conceptual issues have been settled. For instance, the rush to claim impactful implications in the absence of critical evaluation of the reviewed research, overlooks the theory–practice gap, and presupposes that scores on self-report measures (responses to abstract, hypothetical scenarios) represent emotionally intelligent behaviour in the real-world. This is compounded by heterogeneity in EI measures (and their content domains)—many authors within the sampled reviews point to this as an issue, yet still call for further research, inviting more new measures and greater application, even though definitions and descriptions of the construct remain unclear. In highlighting these caveats and then minimising their seriousness through boosted claims, authors maintain the prestige of EI.

Whilst our sample of EI SRs and MAs is substantial ($N=73$), it does not represent an exhaustive corpus of EI SRs and MAs. It will be important to update this line of analysis as newer reviews are published, and to diversify the search engines, search terms and inclusion criteria used to sample reviews. Nevertheless, tentative implications about the reporting of EI focused research, and scientific reporting of Psychology research more broadly, can be hypothesised from the sample corpus.

The thematic pattern identified across all 73 reviews, whereby methodological limitations are acknowledged but consistently overridden, and where construct-level concerns are redirected into methodological research agendas, is consistent with what Oyserman (2023) describes as a culturally fluent research climate, in which shared assumptions reduce the

threshold for scrutiny. We propose the term *epistemic sub-culture* to describe this pattern arising from our analysis. As a tentative interpretive implication beyond the document-level findings, we hypothesise that the socio-political climate, coupled with a shared understanding between authors and peer reviewers of the need to promote EI for the public good, may serve to maintain this *epistemic sub-culture* and the demonstrated incongruence between what the evidence supports and the language used to characterise it. Our thematic analysis shows that many of the original issues levelled at EI are still prevalent (e.g., coherency, measurement), yet appear to be minimised by authors, possibly in pursuit of ‘positive’ and ‘impactful’ outcomes. As has been found with related concepts such as mindfulness (Altgasen et al. 2024) and mindset theory, (Macnamara and Burgoyne 2023) positive, culturally fluent ideas, appear to require a higher threshold of disconfirmatory evidence in order to critique or reject (Oyserman 2023). It is therefore plausible that peer reviewers may be more likely to overlook reporting flaws or evidence gaps in EI-focused systematic reviews given EI’s cultural fluency—its status as an intuitively appealing, socially valued concept (Oyserman 2023)—which may reduce the perceived need for critical scrutiny.

4.3 Recommendations and directions for future research

Our findings and broader tentative implications suggest that a shift to prioritising critical thinking and theorising in EI research and psychology more broadly, is warranted, and this starts with training graduate students (Bosch 2018; Proulx and Morey 2021). This is consistent with calls to guard against poor research practices that erode scientific quality and public trust (Teixeira Da Silva 2023), and with our finding that 81% of EI reviews in this sample are of low or unacceptable methodological quality. The recourse, identified in our analysis, to ‘redeem’ the construct through continued refinement of methods, measures and analytical techniques (i.e., methodologism), is unlikely to address the pervasive conceptual concerns associated with the construct. For instance, we note that many review authors argue that EI is context dependent and further research should examine this (i.e., EI to be measured via multiple reports, tailored to context) but fail to acknowledge that their recommended approach (self-report) removes ‘real’ context entirely, and only serves to maintain the status quo of a fragmented psychology with conceptually redundant constructs (Anvari et al. 2025). Again, this issue is not unique to EI-modern psychology is characterised by a “silent move away” from ‘constructs’ and criticisms of their value, towards ‘observables’ (De Boeck et al. 2023, p. 240), and this, coupled with “*the modern reality of a duty to use big, ubiquitous, ‘organic’ data, integrated from multiple, diverse sources*” (Prosperi et al. 2019, p. 2), is further divorcing our understanding of complex human behaviour from the social context. Noting recent developments in affective computing which combines AI, computer science, and cognitive science to create and apply artificial emotional intelligence (see e.g., Shoumy et al. 2020), we hypothesise that foundational theoretical issues in EI research may well become further masked and forgotten.

Addressing poor research practices per se requires collective action from researchers, publishers, institutions, and funders who incentivise ‘impactful’ research. While some factors are beyond the control of individual researchers, we make an appeal for researchers to pursue *theory-led* effects-based research, and echo Hoekstra and Vazire’s (2021) call for humility regarding researchers’ perspectives and findings; biases should be examined rather than used to promote weak research. Without proper safeguards, research becomes advo-

cacy that may serve to maintain 'EI' for impact rather than truthfully evaluating findings to build a credible knowledge base. We argue for improved researcher and peer reviewer training to address bias in systematic reviews, along with publishers' endorsement of reporting standards. Recent developments such as methodological quality checklist tools (Yavchitz et al. 2016), Corneille et al.'s (2023) checklist of 'persuasive communication devices,' and guidelines for open reproducible systematic reviews in psychology (Carlsson et al. 2024) represent promising advances in enhancing transparency and reducing poor practices to benefit producers (apprentice and post-apprenticeship researchers), disseminators (journals and publishers) and research consumers and users (readers and interventionists).

It is our hope that the growing adoption of open science practices will, over time, strengthen both the conduct and reporting of EI reviews, though given the scale and diversity of EI's applied claims, spanning clinical, educational, organisational, and policy domains, the scope for construct-promotional LH to distort the evidence base seems considerable. Repeating and updating our metasynthesis will be useful for tracking changes in reporting. Finally, it remains an open empirical question whether the construct-focused, prestige-protective forms of LH documented here are unique to EI, or whether this is a feature of academic writing regarding other culturally fluent psychological concepts. For example, mindfulness (Altgassen et al. 2024) and mindset theory (Macnamara and Burgoyne 2023) share both EI's intuitive appeal and its contested evidence base; whether their review literatures exhibit comparable patterns of prestige-claiming and 'hedge-override' warrants systematic investigation.

4.4 Conclusion

This study presents evidence of poor methodological quality and pervasive LH in systematic reviews and meta-analyses on the topic of EI. Using the SIGN appraisal tool, we rated 81% of the 73 included reviews as low quality or rejectable, with only five reviews (7%) achieving a high-quality rating. Critically, poor quality was not confined to older reviews: low and rejectable ratings were present throughout the sample up to 2023, indicating that methodological weaknesses in the corpus are persistent rather than historical. Misleading interpretation and inappropriate extrapolation were common across the lower-quality majority of reviews.

Across all 73 coded reviews representing the full range of SIGN quality levels, LH manifested as prestige claiming (theme 1: *Prestigious*), hedging fallacies (theme 2: *Exaggerated*), and the reframing of construct-level limitations as methodological artefacts (theme 3: *Redeemable*). The dominant and most distinctive pattern was the co-occurrence of themes 1 and 2, presenting as a 'hedge-override' structure, where stated caveats functioned not to constrain confident claims, but to inoculate them, lending credibility to conclusions that exceed the evidence reviewed. The sequence of acknowledging a limitation, asserting EI's importance regardless, then recommending application, appeared consistently across the reviews we sampled, spanning multiple outcome domains and all quality tiers, suggesting it is a discursive convention of this literature. Redeemable framing was near-universal, with conceptual vulnerabilities routinely redirected into methodological research agendas, preserving EI's standing as a construct whose importance is established in principle. Together, these patterns construct EI as a powerful, causally valuable psychological entity whose application is always warranted, independent of the quality of review findings. Whilst the

pervasiveness of these patterns outside of the EI reviews sampled is yet to be established, they are notable here given the scale of EI's applied reach and the self-reinforcing nature of the claim-making documented in our analysis. Addressing them requires not only methodological reform, but a return to critical theorising about what EI is, what it measures, and whether the construct's foundational assumptions are warranted.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11135-026-02913-9>.

Author contributions All authors whose names appear on the submission (SD, BM) made substantial contributions to the conception or design of the work; the acquisition, analysis, and interpretation of data; drafted the work and revised it critically for important intellectual content; approved the version to be published; and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

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

Data availability All supplementary materials and project data files can be downloaded from an open access repository: https://osf.io/p3v9y/?view_only=7603f2a8011b4d49959f8621deeb10cf.

Declarations

Conflict of interest The authors declare no financial conflicts of interest. Non-financial interest: One study included in the meta-synthesis was co-authored by members of the research team; however, inclusion was determined using predefined eligibility criteria and review procedures.

Ethical approval.

Ethical approval was not required for this work (secondary data analysis).

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdulla Alabbasi, A.M., Ayoub, A.E., Ziegler, A.: Are gifted students more emotionally intelligent than their non-gifted peers? A meta-analysis. *High Abil. Stud.* **32**(2), 189–217 (2021). <https://doi.org/10.1080/13598139.2020.1770704>
- Ädel, A.: Adopting a 'move' rather than a 'marker' approach to metadiscourse: a taxonomy for spoken student presentations. *Engl. Specif. Purp.* **69**, 4–18 (2023). <https://doi.org/10.1016/j.esp.2022.09.001>
- Akerjordet, K., Severinsson, E.: Emotional intelligence: a review of the literature with specific focus on empirical and epistemological perspectives. *J. Clin. Nurs.* **16**(8), 1405–1416 (2007). <https://doi.org/10.1111/j.1365-2702.2006.01749.x>
- Altgassen, E., Geiger, M., Wilhelm, O.: Do you mind a closer look? A jingle-jangle fallacy perspective on mindfulness. *Eur. J. Pers.* **38**(2), 365–387 (2024). <https://doi.org/10.1177/08902070231174575>

- Andrei, F., Mancini, G., Baldaro, B., Trombini, E., Agnoli, S.: A systematic review on the predictive utility of the trait emotional intelligence questionnaire (TEIQue). *BPA Appl. Psychol. Bull.* **271**, 5–26 (2014)
- Andrei, F., Siegling, A.B., Aloe, A.M., Baldaro, B., Petrides, K.V.: The incremental validity of the trait emotional intelligence questionnaire (TEIQue): a systematic review and meta-analysis. *J. Pers. Assess.* **98**(3), 261–276 (2016). <https://doi.org/10.1080/00223891.2015.1084630>
- Antonakis, J.: Emotional intelligence: the hype, the hope, the evidence. In: *Emotion Researcher: ISRE's Sourcebook for Research on Emotion and Affect* (2015). <https://emotionresearcher.com/emotional-intelligence-the-hype-the-hope-the-evidence/>
- Antonakis, J., Ashkanasy, N.M., Dasborough, M.T.: Does leadership need emotional intelligence? *Leadersh. q.* **20**(2), 247–261 (2009). <https://doi.org/10.1016/j.leaqua.2009.01.006>
- Anvari, F., Alsalti, T., Oehler, L.A., Hussey, I., Elson, M., Arslan, R.C.: Defragmenting psychology. *Nat. Hum. Behav.* (2025). <https://doi.org/10.1038/s41562-025-02138-0>
- Arora, S., Ashrafian, H., Daus, C.S., Rathanasiou, T., Darzi, A., Sevdalis, N.: Emotional intelligence in medicine: a systematic review through the context of the ACGME competencies. *Med. Educ.* **44**(8), 749–764 (2010). <https://doi.org/10.1111/j.1365-2923.2010.03709.x>
- Ashkanasy, N.M., Daus, C.S.: Rumors of the death of emotional intelligence in organizational behavior are vastly exaggerated. *J. Organ. Behav.* **26**(4), 441–452 (2005). <https://doi.org/10.1002/job.320>
- Baker, A., Young, K., Potter, J., Madan, I.: A review of grading systems for evidence-based guidelines produced by medical specialties. *Clin. Med.* **10**(4), 358 (2010). <https://doi.org/10.7861/clinmedicine.10-4-358>
- Baudry, A.-S., Grynberg, D., Dassonneville, C., Lelorain, S., Christophe, V.: Sub-dimensions of trait emotional intelligence and health: a critical and systematic review of the literature. *Scand. J. Psychol.* **59**, 206–222 (2018). <https://doi.org/10.1111/sjop.12424>
- Blair, C.R.: Meta-synthesis. In: *Annual Meeting of the Association for the Study of Higher Education* (1999)
- Bosch, G.: Train PhD students to be thinkers not just specialists. *Nature* **554**(7692), 277–277 (2018). <https://doi.org/10.1038/d41586-018-01853-1>
- Boutron, I., Ravaud, P.: Misrepresentation and distortion of research in biomedical literature. *Proc. Natl. Acad. Sci.* **115**(11), 2613–2619 (2018). <https://doi.org/10.1073/pnas.1710755115>
- Braun, V., Clarke, V.: One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qual. Res. Psychol.* **18**(3), 328–352 (2021). <https://doi.org/10.1080/14780887.2020.1769238>
- Braun, V., Clarke, V., Hayfield, N., Davey, L., Jenkinson, E.: Doing reflexive thematic analysis. In: Bager-Charleson, S., McBeath, A. (eds.) *Supporting Research in Counselling and Psychotherapy: Qualitative, Quantitative, and Mixed Methods Research*, pp. 19–38. Springer (2022). https://doi.org/10.1007/978-3-031-13942-0_2
- Bru-Luna, L.M., Martí-Vilar, M., Merino-Soto, C., Cervera-Santiago, J.L.: Emotional intelligence measures: a systematic review. *Healthcare* **9**(12), 1696 (2021). <https://doi.org/10.3390/HEALTHCARE9121696>
- Bryan, A.: From ‘the conscience of humanity’ to the conscious human brain: UNESCO’s embrace of social-emotional learning as a flag of convenience. *Comp. J. Comp. Int. Educ.* **54**(5), 770–784 (2024). <https://doi.org/10.1080/03057925.2022.2129956>
- Byrne, D.: A worked example of Braun and Clarke’s approach to reflexive thematic analysis. *Qual. Quant.* **56**(3), 1391–1412 (2022). <https://doi.org/10.1007/s11135-021-01182-y>
- Carlsson, R., Batinović, L., Hyltse, N., Kalmendal, A., Nordström, T., Topor, M.: A beginner’s guide to open and reproducible systematic reviews in psychology. *Collab. Psychol.* **10**(1), 126218 (2024). <https://doi.org/10.1525/collabra.126218>
- Caulfield, T., Sipp, D., Murry, C.E., Daley, G.Q., Kimmelman, J.: Confronting stem cell hype. *Science* **352**(6287), 776–777 (2016). <https://doi.org/10.1126/science.aaf4620>
- Chamorro-Premuzic, T., Sanger, M.: How to boost your (and others’) emotional intelligence. *Harvard Business Review* (2017). <https://hbr.org/2017/01/how-to-boost-your-and-others-emotional-intelligence>
- Cherry, M.G., Fletcher, I., O’Sullivan, H.: The influence of medical students’ and doctors’ attachment style and emotional intelligence on their patient-provider communication. *Patient Educ. Couns.* **93**, 177–187 (2013). <https://doi.org/10.1016/j.pec.2013.05.010>
- Cherry, M.G., Fletcher, I., O’Sullivan, H., Dornan, T.: Emotional intelligence in medical education: a critical review. *Med. Educ.* **48**(5), 468–478 (2014). <https://doi.org/10.1111/medu.12406>
- Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*, 2nd edn. Lawrence Erlbaum (1988)
- Conte, J.M., Dean, M.A.: Can emotional intelligence be measured? In: Murphy, K.R. (ed.) *A Critique of Emotional Intelligence: What are the Problems and How Can They be Fixed?*, pp. 59–77. Routledge (2006)
- Cook, C.J., Cook, C.E., Hilton, T.N.: Does emotional intelligence influence success during medical school admissions and program matriculation?: A systematic review. *J. Educ. Eval. Health Prof.* **13**, 40 (2016). <https://doi.org/10.3352/jeehp.2016.13.40>

- Corneille, O., Havemann, J., Henderson, E.L., IJzerman, H., Hussey, I., Orban De Xivry, J.-J., Jussim, L., Holmes, N.P., Pilacinski, A., Boffara, B., Carroll, H., Outa, N.O., Lush, P., Lotter, L.D.: Beware 'persuasive communication devices' when writing and reading scientific articles. *Elife* **12**, e88654 (2023). <https://doi.org/10.7554/eLife.88654>
- Crummerl, C., Pendlebury-Green, A., Buvat, J., Khadikar, A., Sengupta, A., Shah, H.: Emotional Intelligence—The Essential Skillset for the Age of AI. Capgemini Research Institute (2019). <https://www.capgemini.com/dk-en/wp-content/uploads/sites/42/2019/10/Digital-Report-%E2%80%93Emotional-Intelligence.pdf>
- Crüwell, S., van Doorn, J., Etz, A., Makel, M.C., Moshontz, H., Niebaum, J.C., Orben, A., Parsons, S., Schulte-Mecklenbeck, M.: Seven easy steps to open science. *Z. Psychol.* **227**(4), 237–248 (2019). <https://doi.org/10.1027/2151-2604/a000387>
- Dasborough, M.T., Ashkanasy, N.M., Humphrey, R.H., Harms, P.D., Credé, M., Wood, D.: Does leadership still not need emotional intelligence? Continuing "The Great EI Debate." *Leadersh. q.* **33**(6), 101539 (2022). <https://doi.org/10.1016/j.leaqua.2021.101539>
- Davis, S.K., Nichols, R.: Does emotional intelligence have a 'dark' side? A review of the literature. *Front. Psychol.* **7**(1316), 213863 (2016). <https://doi.org/10.3389/fpsyg.2016.01316>
- De Boeck, P., Pek, J., Walton, K., Wegener, D.T., Turner, B.M., Andersen, B.L., Beauchaine, T.P., Lecavalier, L., Myung, J.I., Petty, R.E.: Questioning psychological constructs: current issues and proposed changes. *Psychol. Inq.* **34**(4), 239–257 (2023). <https://doi.org/10.1080/1047840X.2023.2274429>
- Deonna, J., Teroni, F.: Emotions and their correctness conditions: a defense of attitudinalism. *Erkenntnis* **89**(1), 45–64 (2024). <https://doi.org/10.1007/s10670-022-00522-0>
- Domínguez-García, E., Fernández-Berrocal, P.: The association between emotional intelligence and suicidal behavior: a systematic review. *Front. Psychol.* **9**, 2380 (2018). <https://doi.org/10.3389/fpsyg.2018.02380>
- Ecclestone, K., Hayes, D.: Changing the subject: the educational implications of developing emotional well-being. *Oxf. Rev. Educ.* **35**(3), 371–389 (2009). <https://doi.org/10.1080/03054980902934662>
- Edwards, M.A., Roy, S.: Academic research in the 21st century: maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environ. Eng. Sci.* **34**(1), 51–61 (2017). <https://doi.org/10.1089/ees.2016.0223>
- Encarnação, R.M.C., da Cunha Soares, E.M., de Carvalho, A.L.R.F.: Inteligência emocional: fatores influenciadores e impacto nos enfermeiros em cuidados intensivos. *Rev Rene* **19**, e33229 (2018). <https://doi.org/10.15253/2175-6783.20181933229>
- Frawley, A., Wakeham, C., McLaughlin, K., Ecclestone, K.: Constructing a crisis: mental health, higher education and policy entrepreneurs. *Sociol. Res.* **29**(2), 351–369 (2024). <https://doi.org/10.1177/13607804231215943>
- Furedi, F.: *Therapy Culture: Cultivating Vulnerability in an Uncertain Age*. Routledge (2004). <https://doi.org/10.4324/9780203098684>
- García-Sancho, E., Salguero, J.M., Fernández-Berrocal, P.: Relationship between emotional intelligence and aggression: a systematic review. *Aggress. Violent. Behav.* **19**, 584–591 (2014). <https://doi.org/10.1016/j.avb.2014.07.007>
- Gignac, G.E.: Psychometrics and the measurement of emotional intelligence. In: Stough, C., Saklofske, D.H., Parker, J.D.A. (eds.) *Assessing Emotional Intelligence: Theory, Research and Applications* (Issue the Springer Series on Human Exceptionality), pp. 9–40. Springer (2009)
- Gignac, G.E., Szodorai, E.T.: Effect size guidelines for individual differences researchers. *Pers. Individ. Differ.* **102**, 74–78 (2016). <https://doi.org/10.1016/j.paid.2016.06.069>
- Giusti, E.M., Manna, C., Scolari, A., Mestre, J.M., Prevedar, T., Castelnovo, G., Pietrabissa, G.: The relationship between emotional intelligence, obesity and eating disorder in children and adolescents: a systematic mapping review. *Int. J. Environ. Res. Public Health* **18**, 2054 (2021). <https://doi.org/10.3390/ijerph18042054>
- Goleman, D.: *Emotional Intelligence Why It Can Matter More Than IQ for Character, Health and Lifelong Achievement*. Bantam Books (1995)
- Gómez-Leal, R., Gutiérrez-Cobo, M.J., Cabello, R., Megías, A., Fernández-Berrocal, P.: The relationship between the three models of emotional intelligence and psychopathy: a systematic review. *Front. Psychiatry* **9**, 307 (2018). <https://doi.org/10.3389/fpsyg.2018.00307>
- Gómez-Leal, R., Holzer, A.A., Bradley, C., Fernández-Berrocal, P., Patti, J.: The relationship between emotional intelligence and leadership in school leaders: a systematic review. *Camb. J. Educ.* **52**(1), 1–21 (2021). <https://doi.org/10.1080/0305764X.2021.1927987>
- Guilera, G., Barrios, M., Gómez-Benito, J.: Meta-analysis in psychology: a bibliometric study. *Scientometrics* **94**(3), 943–954 (2013). <https://doi.org/10.1007/s11192-012-0761-2>
- Gutiérrez-Cobo, M.J., Cabello, R., Fernández-Berrocal, P.: The relationship between emotional intelligence and cool and hot cognitive processes: a systematic review. *Front. Behav. Neurosci.* **10**, 101 (2016). <https://doi.org/10.3389/fnbeh.2016.00101>

- Harms, P.D., Credé, M.: Remaining issues in emotional intelligence research: construct overlap, method artifacts, and lack of incremental validity. *Ind. Organ. Psychol.* **3**(2), 154–158 (2010a). <https://doi.org/10.1111/j.1754-9434.2010.01217.x>
- Harms, P.D., Credé, M.: Emotional intelligence and transformational and transactional leadership: a meta-analysis. *J. Leadersh. Organ. Stud.* **17**(1), 5–17 (2010b). <https://doi.org/10.1177/1548051809350894>
- Hoekstra, R., Vazire, S.: Aspiring to greater intellectual humility in science. *Nat. Hum. Behav.* **5**(12), 1602–1607 (2021). <https://doi.org/10.1038/s41562-021-01203-8>
- Hohn, R.E., Slaney, K.L., Tafreshi, D.: An empirical review of research and reporting practices in psychological meta-analyses. *Rev. Gen. Psychol.* **24**(3), 195–209 (2020). <https://doi.org/10.1177/1089268020918844>
- Humphrey, N., Lendrum, A., Wigelsworth, M.: Making the most out of school-based prevention: lessons from the social and emotional aspects of learning (SEAL) programme. *Emot. Behav. Diffic.* **18**(3), 248–260 (2013). <https://doi.org/10.1080/13632752.2013.819251>
- Hyland, K.: Boosting, hedging and the negotiation of academic knowledge. *Text Talk* **18**(3), 349–382 (1998). <https://doi.org/10.1515/text.1.1998.18.3.349>
- Hyland, K.: Metadiscourse: What is it and where is it going? *J. Pragmat.* **113**, 16–29 (2017). <https://doi.org/10.1016/j.pragma.2017.03.007>
- Hyland, K., Jiang, F.: ‘Our striking results demonstrate ...’: Persuasion and the growth of academic hype. *J. Pragmat.* **182**, 189–202 (2021). <https://doi.org/10.1016/j.pragma.2021.06.018>
- Hyland, K., Jiang, F.: Hying the REF: promotional elements in impact submissions. *High. Educ.* **87**(3), 685–702 (2024). <https://doi.org/10.1007/s10734-023-01030-y>
- Intemam, K.: Understanding the Problem of? Hype?: Exaggeration, values, and trust in science. *Can. J. Philos.* **52**(3), 279–294 (2022). <https://doi.org/10.1017/can.2020.45>
- Ioannidis, J.P.: The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q.* **94**(3), 485–514 (2016). <https://doi.org/10.1111/1468-0009.12210>
- Jardine, B.B., Vannier, S., Voyer, D.: Emotional intelligence and romantic relationship satisfaction: a systematic review and meta-analysis. *Pers. Individ. Differ.* **196**, 111713 (2022). <https://doi.org/10.1016/j.paid.2022.111713>
- Jiménez-Picón, N., Romero-Martín, M., Ponce-Blandón, J.A., Ramirez-Baena, L., Palomo-Lara, J.C., Gómez-Salgado, J.: The relationship between mindfulness and emotional intelligence as a protective factor for healthcare professionals: systematic review. *Int. J. Environ. Res. Public Health* **18**(10), 5491 (2021). <https://doi.org/10.3390/ijerph18105491>
- King, V.J., Stevens, A., Nussbaumer-Streit, B., Kamel, C., Garritty, C.: Paper 2: performing rapid reviews. *Syst. Rev.* **11**(1), 151 (2022). <https://doi.org/10.1186/s13643-022-02011-5>
- Kittisiam, T., Ritter, C., Morabito, E., Stacey, A., Khosa, D., Jones, A.: A qualitative exploration of the emotional experiences and applications of emotional intelligence in early-career veterinarians. *Vet. Rec.* **197**(4), e5599 (2025). <https://doi.org/10.1002/vetr.5599>
- Kopp, A., Jekauc, D.: The influence of emotional intelligence on performance in competitive sports: a meta-analytical investigation. *Sports* **6**(4), 175 (2018). <https://doi.org/10.3390/sports6040175>
- Kotsou, I., Mikolajczak, M., Heeren, A., Grégoire, J., Leys, C.: Improving emotional intelligence: a systematic review of existing work and future challenges. *Emot. Rev.* **11**(2), 151–165 (2019). <https://doi.org/10.1177/1754073917735902>
- Kun, B., Demetrovics, Z.: Emotional intelligence and addictions: a systematic review. *Subst. Use Misuse* **45**(7–8), 1131–1160 (2010). <https://doi.org/10.3109/10826080903567855>
- Laborde, S., Dosseville, F., Allen, M.S.: Emotional intelligence in sport and exercise: a systematic review. *Scand. J. Med. Sci. Sports* **26**(4), 395–407 (2016). <https://doi.org/10.1111/sms.12510>
- Landis, J.R., Koch, G.G.: An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* **33**, 363–374 (1977). <https://doi.org/10.2307/2529786>
- Landy, F.J.: Some historical and scientific issues related to research on emotional intelligence. *J. Organ. Behav.* **26**(4), 411–424 (2005). <https://doi.org/10.1002/job.317>
- Lea, R.G., Davis, S.K., Mahoney, B.K., Qualter, P.: Does emotional intelligence buffer the effects of acute stress? A systematic review. *Front. Psychol.* **10**, 810 (2019). <https://doi.org/10.3389/FPSYG.2019.00810>
- Lindebaum, D.: A qualitative study of emotional intelligence and its underlying processes and outcomes in management studies. In: *New Ways of Studying Emotions in Organizations*, vol. 11. Emerald Group Publishing Limited (2015). <https://doi.org/10.1108/S1746-979120150000011006>
- Llamas-Díaz, D., Cabello, R., Megías-Robles, A., Fernández-Berrocal, P.: Systematic review and meta-analysis: the association between emotional intelligence and subjective well-being in adolescents. *J. Adolesc.* **94**(7), 925–938 (2022). <https://doi.org/10.1002/jad.12075>
- Locke, E.A.: Why emotional intelligence is an invalid concept. *J. Organ. Behav.* **26**(4), 425–431 (2005). <https://doi.org/10.1002/job.318>

- MacCann, C., Erbas, Y., Dejonckheere, E., Minbashian, A., Kuppens, P., Fayn, K.: Emotional intelligence relates to emotions, emotion dynamics, and emotion complexity: a meta-analysis and experience sampling study. *Eur. J. Psychol. Assess.* **36**(3), 460–470 (2020a). <https://doi.org/10.1027/1015-5759/a000588>
- MacCann, C., Jiang, Y., Brown, L.E.R., Double, K.S., Bucich, M., Minbashian, A.: Emotional intelligence predicts academic performance: a meta-analysis. *Psychol. Bull.* **146**(2), 150–186 (2020b). <https://doi.org/10.1037/bul0000219>
- Macnamara, B.N., Burgoyne, A.P.: Do growth mindset interventions impact students' academic achievement? A systematic review and meta-analysis with recommendations for best practices. *Psychol. Bull.* **149**(3–4), 133–173 (2023). <https://doi.org/10.1037/bul0000352>
- Malouff, J.M., Schutte, N.S., Thorsteinsson, E.B.: Trait emotional intelligence and romantic relationship satisfaction: a meta-analysis. *Am. J. Fam. Therapy* **42**(1), 53–66 (2014). <https://doi.org/10.1080/01926187.2012.748549>
- Mancini, G., Özal, Z., Biolcati, R., Trombini, E., Petrides, K.V.: Trait emotional intelligence and adolescent psychological well-being: a systematic review. *Int. J. Adolesc. Youth* **29**(1), 2292057 (2024). <https://doi.org/10.1080/02673843.2023.2292057>
- Martins, A., Ramalho, N., Morin, E.: A comprehensive meta-analysis of the relationship between emotional intelligence and health. *Pers. Individ. Differ.* **49**(6), 554–564 (2010). <https://doi.org/10.1016/j.paid.2010.05.029>
- Matthews, G., Roberts, R.D., Zeidner, M.: Seven myths about emotional intelligence. *Psychol. Inq.* **15**, 179–196 (2004). https://doi.org/10.1207/s15327965pli1503_01
- Matthews, A., Herrett, E., Gasparrini, A., Staa, T.V., Goldacre, B., Smeeth, L., Bhaskaran, K.: Impact of statin related media coverage on use of statins: interrupted time series analysis with UK primary care data. *BMJ* **353**, i3283 (2016). <https://doi.org/10.1136/bmj.i3283>
- Maul, A.: The validity of the Mayer–Salovey–Caruso emotional intelligence test (MSCEIT) as a measure of emotional intelligence. *Emot. Rev.* **4**(4), 394–402 (2012). <https://doi.org/10.1177/1754073912445811>
- Mayer, J.D., Caruso, D.R.: Emotional intelligence and socioemotional attributes: distinct constructs and their uses. *Consult. Psychol. J.* (2025). <https://doi.org/10.1037/cpb0000293>
- Mayer, J.D., Caruso, D.R., Salovey, P.: The ability model of emotional intelligence: principles and updates. *Emot. Rev.* **8**(4), 290–300 (2016). <https://doi.org/10.1177/1754073916639667>
- McGuinness, L.A., Higgins, J.P.T.: Risk-of-bias VISualization (robvis): An R package and Shiny web app for visualizing risk-of-bias assessments. *Res. Synth. Methods* **12**(1), 55–61 (2021). <https://doi.org/10.1002/jrsm.1411>
- McHugh, M.L.: Interrater reliability: the kappa statistic. *Biochem. Med.* **22**(3), 276–282 (2012). <https://doi.org/10.11613/BM.2012.031>
- Meehl, P.E.: Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *J. Consult. Clin. Psychol.* **46**(4), 806–834 (1978). <https://doi.org/10.1037/0022-006X.46.4.806>
- Megías, A., Gómez-Leal, R., Gutiérrez-Cobo, M.J., Cabello, R., Fernández-Berrocal, P.: The relationship between trait psychopathy and emotional intelligence: a meta-analytic review. *Neurosci. Biobehav. Rev.* **84**, 198–203 (2018). <https://doi.org/10.1016/j.neubiorev.2017.12.003>
- Mérida-López, S., Extremera, N.: Emotional intelligence and teacher burnout: a systematic review. *Int. J. Educ. Res.* **85**, 121–130 (2017). <https://doi.org/10.1016/j.ijer.2017.07.006>
- Miao, C., Humphrey, R.H., Qian, S.: Leader emotional intelligence and subordinate job satisfaction: a meta-analysis of main, mediator, and moderator effects. *Pers. Individ. Differ.* **102**, 13–24 (2016). <https://doi.org/10.1016/j.paid.2016.06.056>
- Miao, C., Humphrey, R.H., Qian, S.: Are the emotionally intelligent good citizens or counterproductive? A meta-analysis of emotional intelligence and its relationships with organizational citizenship behavior and counterproductive work behavior. *Pers. Individ. Differ.* **116**, 144–156 (2017a). <https://doi.org/10.1016/j.paid.2017.04.015>
- Miao, C., Humphrey, R.H., Qian, S.: A meta-analysis of emotional intelligence and work attitudes. *J. Occup. Organ. Psychol.* **90**(2), 177–202 (2017b). <https://doi.org/10.1111/joop.12167>
- Miao, C., Humphrey, R.H., Qian, S.: A meta-analysis of emotional intelligence effects on job satisfaction mediated by job resources, and a test of moderators. *Pers. Individ. Differ.* **116**, 281–288 (2017c). <https://doi.org/10.1016/j.paid.2017.04.031>
- Miao, C., Humphrey, R.H., Qian, S.: Emotional intelligence and authentic leadership: a meta-analysis. *Leadersh. Org. Dev. J.* **39**(5), 679–690 (2018a). <https://doi.org/10.1108/LODJ-02-2018-0066>
- Miao, C., Humphrey, R.H., Qian, S.: The relationship between emotional intelligence and trait mindfulness: a meta-analytic review. *Pers. Individ. Differ.* **135**, 101–107 (2018b). <https://doi.org/10.1016/j.paid.2018.06.051>

- Miao, C., Humphrey, R.H., Qian, S., Pollack, J.M.: Emotional intelligence and entrepreneurial intentions: an exploratory meta-analysis. *Career Dev. Int.* **23**(5), 497–512 (2018c). <https://doi.org/10.1108/CDI-01-2018-0019>
- Miao, C., Barone, M.J., Qian, S., Humphrey, R.H.: Emotional intelligence and service quality: a meta-analysis with initial evidence on cross-cultural factors and future research directions. *Mark. Lett.* **30**, 335–347 (2019a). <https://doi.org/10.1007/s11002-019-09495-7>
- Miao, C., Humphrey, R.H., Qian, S., Pollack, J.M.: The relationship between emotional intelligence and the dark triad personality traits: a meta-analytic review. *J. Res. Pers.* **78**, 189–197 (2019b). <https://doi.org/10.1016/j.jrp.2018.12.004>
- Miao, C., Humphrey, R.H., Qian, S.: Emotional intelligence and servant leadership: a meta-analytic review. *Bus. Ethics Environ. Responsib.* **30**(2), 231–243 (2021). <https://doi.org/10.1111/beer.12332>
- Millar, N., Salager-Meyer, F., Budgell, B.: “It is important to reinforce the importance of ...”: ‘Hype’ in reports of randomized controlled trials. *Engl. Specif. Purp.* **54**, 139–151 (2019). <https://doi.org/10.1016/j.esp.2019.02.004>
- Millar, N., Batalo, B., Budgell, B.: Promotional language (hype) in abstracts of publications of National Institutes of Health-Funded Research, 1985–2020. *JAMA Netw. Open* **6**(12), e2348706 (2023). <https://doi.org/10.1001/jamanetworkopen.2023.48706>
- Morales-Sánchez, L., Brandão, T., Guil, R.: Emotional intelligence and breast cancer: a systematic review. *Heliyon* **10**, e25061 (2024). <https://doi.org/10.1016/j.heliyon.2024.e25061>
- Murphy, K.R., Sideman, L.: The two Els. In: Murphy, K.R. (ed.) *A Critique of Emotional Intelligence: What are the Problems and How Can They be Fixed?*, pp. 37–58. Routledge (2006)
- Ngoc, N.N., Tuan, N.P., Takahashi, Y.: A meta-analytic investigation of the relationship between emotional intelligence and emotional manipulation. *SAGE Open* (2020). <https://doi.org/10.1177/2158244020971615>
- O’Boyle, E.H., Jr., Humphrey, R.H., Pollack, J.M., Hawver, T.H., Story, P.A.: The relation between emotional intelligence and job performance: a meta-analysis. *J. Organ. Behav.* **32**(5), 788–818 (2011). <https://doi.org/10.1002/job.714>
- O’Connor, P.J., Hill, A., Kaya, M., Martin, B.: The measurement of emotional intelligence: a critical review of the literature and recommendations for researchers and practitioners. *Front. Psychol.* **10**, 1116 (2019). <https://doi.org/10.3389/fpsyg.2019.01116>
- Ogurlu, U.: A meta-analytic review of emotional intelligence in gifted individuals: a multilevel analysis. *Pers. Individ. Differ.* **171**, 110503 (2021). <https://doi.org/10.1016/j.paid.2020.110503>
- Oussi, A., Hamid, K., Bouvet, C.: Managing emotions in panic disorder: a systematic review of studies related to emotional intelligence, alexithymia, emotion regulation, and coping. *J. Behav. Ther. Exp. Psychiatry* **79**, 101835 (2023). <https://doi.org/10.1016/j.jbtep.2023.101835>
- Oyserman, D.: Culturally fluent theories, metascience, and scientific progress: a case example: commentary on Macnamara and Burgoyne (2023) and Burnette et al. (2023). *Psychol. Bull.* **149**(3–4), 220–228 (2023). <https://doi.org/10.1037/bul0000376>
- Page, M.J., Moher, D., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McKenzie, J.E.: PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **372**, n160 (2021). <https://doi.org/10.1136/bmj.n160>
- Peña-Sarrionandia, A., Mikolajczak, M., Gross, J.J.: Integrating emotion regulation and emotional intelligence traditions: a meta-analysis. *Front. Psychol.* **6**, 160 (2015). <https://doi.org/10.3389/fpsyg.2015.00160>
- Perera, H.N., DiGiacomo, M.: The relationship of trait emotional intelligence with academic performance: a meta-analytic review. *Learn. Individ. Differ.* **28**, 20–33 (2013). <https://doi.org/10.1016/j.lindif.2013.08.002>
- Peterson, K., Malouff, J., Thorsteinsson, E.B.: A meta-analytic investigation of emotional intelligence and alcohol involvement. *Subst. Use Misuse* **46**(14), 1726–1733 (2011). <https://doi.org/10.3109/10826084.2011.618962>
- Petrides, K.V., Mikolajczak, M., Mavroveli, S., Sanchez-Ruiz, M.-J., Furnham, A., Pérez-González, J.-C.: Developments in trait emotional intelligence research. *Emot. Rev.* **8**(4), 335–341 (2016). <https://doi.org/10.1177/1754073916650493>
- Pirsoul, T., Parmentier, M., Sovet, L., Nils, F.: Emotional intelligence and career-related outcomes: a meta-analysis. *Hum. Resour. Manag. Rev.* **33**, 100967 (2023). <https://doi.org/10.1016/j.hrmr.2023.100967>
- Polanin, J.R., Hennessy, E.A., Tsuji, S.: Transparency and reproducibility of meta-analyses in psychology: a meta-review. *Perspect. Psychol. Sci.* **15**(4), 1026–1041 (2020). <https://doi.org/10.1177/1745691620906416>
- Prosperi, M., Bian, J., Buchan, I.E., Koopman, J.S., Sperrin, M., Wang, M.: Raiders of the lost HARK: a reproducible inference framework for big data science. *Palgrave Commun.* **5**(1), 1–12 (2019). <https://doi.org/10.1057/s41599-019-0340-8>

- Proulx, T., Morey, R.D.: Beyond statistical ritual: Theory in psychological science. *Perspect. Psychol. Sci.* **16**(4), 671–681 (2021)
- Quílez-Robres, A., Usán, P., Lozano-Blasco, R., Salavera, C.: Emotional intelligence and academic performance: a systematic review and meta-analysis. *Think. Skills Creat.* **49**, 101355 (2023). <https://doi.org/10.1016/j.tsc.2023.101355>
- Ranjbar, H., Bakhshi, M.: The association between internet addiction and emotional intelligence: a meta-analysis study. *Acta Fac. Med. Naissensis* **35**(1), 17–29 (2018). <https://doi.org/10.2478/afmna-2018-0002>
- Resurrección, D.M., Salguero, J.M., Ruiz-Aranda, D.: Emotional intelligence and psychological maladjustment in adolescence: a systematic review. *J. Adolesc.* **37**(4), 461–472 (2014). <https://doi.org/10.1016/j.adolescence.2014.03.012>
- Roberts, K., Dowell, A., Nie, J.-B.: Attempting rigour and replicability in thematic analysis of qualitative research data: a case study of codebook development. *BMC Med. Res. Methodol.* **19**(1), 66 (2019). <https://doi.org/10.1186/s12874-019-0707-y>
- Romero-Mesa, J., Peláez-Fernández, M.A., Extremera, N.: Emotional intelligence and eating disorders: a systematic review. *Eat. Weight Disord. Stud. Anorex. Bulim. Obes.* **26**, 1287–1301 (2021). <https://doi.org/10.1007/s40519-020-00968-7>
- Rueda, P., Pérez-Romero, N., Cerezo, M.V., Fernández-Berrocal, P.: The role of emotional intelligence in adolescent bullying: A systematic review. *Psicol. Educ.* **28**(1), 53–59 (2022). <https://doi.org/10.5093/psed2021a29>
- Saikia, M., George, L.S., Unnikrishnan, B., Nayak, B.S., Ravishankar, N.: Thirty years of emotional intelligence: a scoping review of emotional intelligence training programme among nurses. *Int. J. Ment. Health Nurs.* **33**, 37–51 (2024). <https://doi.org/10.1111/inm.13235>
- Salovey, P., Mayer, J.D.: Emotional Intelligence. *Imagin. Cogn. Pers.* **9**, 185–211 (1990). <https://doi.org/10.2190/DUGG-P24E-52WK-6CDG>
- Sánchez-Álvarez, N., Extremera, N., Fernández-Berrocal, P.: The relation between emotional intelligence and subjective well-being: a meta-analytic investigation. *J. Posit. Psychol.* **11**(3), 276–285 (2016). <https://doi.org/10.1080/17439760.2015.1058968>
- Sánchez-Álvarez, N., Berrios Martos, M.P., Extremera, N.: A meta-analysis of the relationship between emotional intelligence and academic performance in secondary education: a multi-stream comparison. *Front. Psychol.* **11**, 1517 (2020). <https://doi.org/10.3389/fpsyg.2020.01517>
- Sarrionandia, A., Mikolajczak, M.: A meta-analysis of the possible behavioural and biological variables linking trait emotional intelligence to health. *Health Psychol. Rev.* **14**(2), 220–244 (2020). <https://doi.org/10.1080/17437199.2019.1641423>
- Schlaerth, A., Ensari, N., Christian, J.: A meta-analytical review of the relationship between emotional intelligence and leaders' constructive conflict management. *Group Process. Intergroup Relat.* **16**(1), 126–136 (2013). <https://doi.org/10.1177/1368430212439907>
- Schutte, N.S., Malouff, J.M., Thorsteinsson, E.B., Bhullar, N., Rooke, S.E.: A meta-analytic investigation of the relationship between emotional intelligence and health. *Pers. Individ. Differ.* **42**(6), 921–933 (2007). <https://doi.org/10.1016/j.paid.2006.09.003>
- Shercliffe, R.J., Stahl, W., Tuttle, M.P.: The use of meta-analysis in psychology: A superior vintage or the casting of old wine in new bottles? *Theory Psychol.* **19**(3), 413–430 (2009). <https://doi.org/10.1177/0959354309104161>
- Shoumy, N.J., Ang, L.-M., Seng, K.P., Rahaman, D.M.M., Zia, T.: Multimodal big data affective analytics: a comprehensive survey using text, audio, visual and physiological signals. *J. Netw. Comput. Appl.* **149**, 102447 (2020). <https://doi.org/10.1016/j.jnca.2019.102447>
- SIGN: Methodology Checklist 1: Systematic Reviews and Meta-analyses. SIGN; Edinburgh (2012). <https://www.sign.ac.uk/what-we-do/methodology/checklists/>
- Singh, N., Duraipappah, A.: Rethinking Learning: A Review of Social and Emotional Learning Frameworks for Education Systems. UNESCO Mahatma Gandhi Institute of Education for Peace and Sustainable Development (2020).
- Singh, N., Kulkarni, S., Gupta, R.: Is emotional intelligence related to objective parameters of academic performance in medical, dental, and nursing students: a systematic review. *Educ. Health* **33**(1), 8–12 (2020). https://doi.org/10.4103/efh.Efh_208_17
- Six Seconds: EQ and the Sustainable Development Goals (n.d.). Retrieved 29 Oct 2025, from <https://www.6seconds.org/about/eq-and-the-sustainable-development-goals/>
- South, E., Lorenc, T.: Use and value of systematic reviews in English local authority public health: A qualitative study. *BMC Public Health* **20**(1), 1100 (2020). <https://doi.org/10.1186/s12889-020-09223-1>
- Strobel, J., van Barneveld, A.: When is PBL more effective? A meta-synthesis of meta-analyses comparing PBL to conventional classrooms. *Interdiscip. J. Probl. Based Learn.* **3**(1), 44–58 (2009). <https://doi.org/10.7771/1541-5015.1046>

- Sumner, P., Vivian-Griffiths, S., Boivin, J., Williams, A., Venetis, C.A., Davies, A., Ogden, J., Whelan, L., Hughes, B., Dalton, B., Boy, F., Chambers, C.D.: The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ* **349**(dec09 7), g7015–g7015 (2014). <https://doi.org/10.1136/bmj.g7015>
- Teixeira Da Silva, J.A.: Junk science, junk journals, and junk publishing management: risk to science's credibility. *Philosophia* **51**(3), 1701–1704 (2023). <https://doi.org/10.1007/s11406-022-00590-0>
- Thacoor, A., Smith, O., Nikkhah, D.: The role of emotional intelligence in predicting a successful career for plastic surgeons: a systematic review. *Plast. Reconstr. Surg. Glob. Open* **8**, e2699 (2020). <https://doi.org/10.1097/GOX.0000000000002699>
- Toriello, H.V., Van de Ridder, J.M.M., Brewer, P., Mavis, B., Allen, R., Arvidson, C., Kovar-Gough, I., Novak, E., O'Donnell, J., Osuch, J., Ulrich, B.: Emotional intelligence in undergraduate medical students: a scoping review. *Adv. Health Sci. Educ.* **27**, 167–187 (2022). <https://doi.org/10.1007/s10459-021-10079-2>
- Ubago-Jiménez, J.L., González-Valero, G., Puertas-Molero, P., García-Martínez, I.: Development of emotional intelligence through physical activity and sport practice: a systematic review. *Behav. Sci.* **9**(4), 44 (2019). <https://doi.org/10.3390/bs9040044>
- Uher, J.: Psychology's status as a science: peculiarities and intrinsic challenges. Moving beyond its current deadlock towards conceptual integration. *Integr. Psychol. Behav. Sci.* **55**(1), 212–224 (2020). <https://doi.org/10.1007/S12124-020-09545-0>
- van der Linden, D., Pekaar, K.A., Bakker, A.B., Schermer, J.A., Vernon, P.A., Dunkel, C.S., Petrides, K.V.: Overlap between the general factor of personality and emotional intelligence: a meta-analysis. *Psychol. Bull.* **143**(1), 36–52 (2017). <https://doi.org/10.1037/bul0000078>
- Van Rooy, D.L., Viswesvaran, C.: Emotional intelligence: a meta-analytic investigation of predictive validity and nomological net. *J. Vocat. Behav.* **65**, 71–95 (2004). [https://doi.org/10.1016/S0001-8791\(03\)00076-9](https://doi.org/10.1016/S0001-8791(03)00076-9)
- Van Rooy, D.L., Viswesvaran, C., Pluta, P.: An evaluation of construct validity: What is this thing called emotional intelligence? *Hum. Perform.* **18**(4), 445–462 (2005). https://doi.org/10.1207/s15327043hup1804_9
- Vega, A., Cabello, R., Megías-Robles, A., Gómez-Leal, R., Fernández-Berrocal, P.: Emotional intelligence and aggressive behaviors in adolescents: a systematic review and meta-analysis. *Trauma Violence Abuse* (2021). <https://doi.org/10.1177/1524838021991296>
- Vinkers, C.H., Tjebkink, J.K., Otte, W.M.: Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: Retrospective analysis. *BMJ* **351**, h6467 (2015). <https://doi.org/10.1136/bmj.h6467>
- Walker, S.A., Double, K.S., Birney, D.P.: The complicated relationship between the dark triad and emotional intelligence: a systematic review. *Emot. Rev.* **13**(3), 257–274 (2021). <https://doi.org/10.1177/17540739211014585>
- Walker, S.A., Double, K.S., Kunst, H., Zhang, M., MacCann, C.: Emotional intelligence and attachment in adulthood: a meta-analysis. *Pers. Individ. Differ.* **184**, 111174 (2022). <https://doi.org/10.1016/j.paid.2021.111174>
- Written answers: (2007). <https://hansard.parliament.uk/Commons/2007-01-29/debates/07012937000045/SEALProgramme>
- Xu, X., Pang, W., Xia, M.: Are emotionally intelligent people happier? A meta-analysis of the relationship between emotional intelligence and subjective well-being using Chinese samples. *Asian J. Soc. Psychol.* **24**(4), 477–498 (2021). <https://doi.org/10.1111/ajsp.12445>
- Yavchitz, A., Ravaud, P., Altman, D.G., Moher, D., Hrobjartsson, A., Lasserson, T., Boutron, I.: A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. *J. Clin. Epidemiol.* **75**, 56–65 (2016). <https://doi.org/10.1016/j.jclinepi.2016.01.020>
- Zeidner, M., Roberts, R.D., Matthews, G.: Can emotional intelligence be schooled? A critical review. *Educ. Psychol.* **37**(4), 215–231 (2002). https://doi.org/10.1207/S15326985EP3704_2
- Zhang, J., Wang, Y., Wu, C., He, J.: The relationship between emotional intelligence and eating disorders or disordered eating behaviors: a meta-analysis. *Pers. Individ. Differ.* **185**, 111239 (2022). <https://doi.org/10.1016/j.paid.2021.111239>