Health Psychology Review

ISSN: 1743-7199 (Print) 1743-7202 (Online) Journal homepage: www.tandfonline.com/journals/rhpr20

Adoption of the Transparency and Openness Promotion (TOP) guidelines within health psychology and behavioural medicine journal policies: a cross-sectional study

Elaine Toomey, Rory Coyne, Christina Derksen, Sean P. Grant, Christopher M. Jones, Marta Kijowska, Ilona McNeill, Felix Naughton, Aoife O'Mahony & Emma Norris

To cite this article: Elaine Toomey, Rory Coyne, Christina Derksen, Sean P. Grant, Christopher M. Jones, Marta Kijowska, Ilona McNeill, Felix Naughton, Aoife O'Mahony & Emma Norris (13 Jun 2025): Adoption of the Transparency and Openness Promotion (TOP) guidelines within health psychology and behavioural medicine journal policies: a cross-sectional study, Health Psychology Review, DOI: 10.1080/17437199.2025.2516010

To link to this article: https://doi.org/10.1080/17437199.2025.2516010

HEALTH PSYCHOLOGY REVIEW
https://doi.org/10.1080/17437199.2025.2516010

Routledge
Taylor & Francis Group

REVIEW ARTICLE

OPEN ACCESS Check for updates

Adoption of the Transparency and Openness Promotion (TOP) guidelines within health psychology and behavioural medicine journal policies: a cross-sectional study

Elaine Toomey[a,b], Rory Coyne[c], Christina Derksen[d], Sean P. Grant[e], Christopher M. Jones[f], Marta Kijowska[g], Ilona McNeill[h], Felix Naughton[i], Aoife O'Mahony[j] and Emma Norris[k]

[a]Centre for Health Research Methodology, School of Nursing and Midwifery, Áras Moyola, University of Galway, Galway, Ireland; [b]Institute for Clinical Trials, College of Medicine, Nursing and Health Sciences, University of Galway, Galway, Ireland; [c]School of Psychology, University of Galway, Galway, Ireland; [d]Centre for Cancer Screening, Prevention, and Early Diagnosis, Wolfson Institute of Population Health, Queen Mary University of London, London, UK; [e]HEDCO Institute for Evidence-Based Educational Practice, College of Education, University of Oregon, Eugene, OR, USA; [f]Center for Preventive Medicine and Digital Health, Medical Faculty Mannheim, Heidelberg University, Heidelberg, Germany; [g]Institute of Psychology, Health & Coping Research Group, SWPS University, Warsaw, Poland; [h]School of Psychological Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia; [i]Addiction Research Group, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK; [j]School of Public Health, University College Cork, Cork, Ireland; [k]Department of Health Sciences, Brunel University London, Uxbridge, England

ABSTRACT

Scientific journals play a crucial role in promoting open science. The Transparency and Openness Promotion (TOP) guidelines identify a range of standards that journals can adopt to promote the verifiability of the research they publish. We evaluated the adoption of TOP standards within health psychology and behavioural medicine journal policies, as this had not yet been systematically assessed. In a cross-sectional study on 19 health psychology and behavioural medicine journals, eight raters evaluated TOP standard adoption by these journals using the TRUST journal policy evaluation tool. Out of a total possible score of 29, journal scores ranged from 1 to 13 (median = 6). Standards related to use of reporting guidelines and data transparency were adopted the most, whereas standards related to pre-registration of study analysis plans and citation of code were adopted the least. TOP guidelines have to-date been poorly adopted within health psychology and behavioural medicine journal policies. There are several relatively straightforward opportunities for improvement, such as expanding policies around research data to also consider code and materials, and reducing ambiguity of wording. However, other improvements may require a collaborative approach involving all research stakeholders.

ARTICLE HISTORY
Received 24 October 2024
Accepted 15 April 2025

KEYWORDS
Open science; publishing guidelines; journal policies; TOP factor; transparency; verifiability

1. Introduction

In recent years, there have been substantial efforts to move science towards a more transparent and inclusive approach, aiming to facilitate equitable access to resources, increase trustworthiness,

CONTACT Elaine Toomey Elaine.toomey@universityofgalway.ie
Supplemental data for this article can be accessed online at https://doi.org/10.1080/17437199.2025.2516010.

enhance reproducibility and thus improve scientific rigour (Munafò et al., 2022). This open science approach is especially important for applied sciences such as health psychology and behavioural medicine (Segerstrom et al., 2023). For one, research in this area is often publicly funded through federal and state agencies or charity organisations. Furthermore, applied health research often involves other stakeholders such as healthcare professionals with the overall aim of improving the efficiency, safety and effectiveness of daily practice and healthcare services. Other stakeholders such as patients, the public, policymakers and governments also stand to benefit from increased accessibility and availability of research findings that can and should be translated into policy and real-life impact. Thus, there is an increased demand for accountability and an ethical obligation for reliable systems and practices to ensure best practice research (Taylor & Gorman, 2022).

Open science requires an accordingly open mindset from researchers (Hagger, 2022), institutional training and supervision (Cole et al., 2023), and reinforcement from funding bodies (Kwasnicka et al., 2021). In addition, scientific journals play a crucial role in promoting open science. As the main gatekeepers of traditional academic research dissemination, journals are highly relevant for facilitating open science practices such as implementing registered reports to mitigate against publication bias, encouraging pre-registration of study protocols and analysis plans, use of repositories for housing open data, protocols, materials and code, publishing preprints and open access articles and employing open peer review (Castro et al., 2017; Kwasnicka et al., 2021; Wolfram et al., 2020). A previous example of the positive potential impact of journal policies on research practices is the endorsement of clinical trials registration by the International Committee of Medical Journal Editors, which is reported to have led to an increase in clinical trial registration (De Angelis et al., 2004).

To support journals in taking a rigorous approach in promoting transparency and openness in the research they publish, the Transparency and Openness Promotion (TOP) Guidelines were developed by journals, funders and societies in 2015 (Nosek et al., 2015). There are eight TOP standards described within the guidelines; Data Citation, Data Transparency, Analytic Methods (Code) Transparency, Research Materials Transparency, Reporting Guidelines (i.e., design and analysis reporting transparency), Pre-registration of Studies, Pre-registration of Analysis Plans and Replication. Each standard can be implemented at increasing levels of stringency, ranging from requiring authors to disclose the use of an open science practice (Level 1) to requiring authors to use the practice (Level 2), to requiring an independent verification of those practices before publication (Level 3). In 2020 the Centre for Open Science (COS) developed the TOP Factor (TOP Guidelines (cos.io)) metric to quantify the extent to which TOP standards are being implemented within journals (Table 1). The metric also considers two additional standards – one relating to publication bias of original studies and the use of registered reports, and another relating to the use of open science badges. The metric was developed to provide an alternative approach to measuring journal quality beyond the traditional journal impact factor.

Researchers have previously used the TOP Factor to audit journals' adoption of TOP standards into their policies in areas such as social intervention, pain, sports medicine, and medical and health sciences (Cashin et al., 2021; Gardener et al., 2022; Grant et al., 2023; Hansford et al., 2022; Lee et al., 2018; Patarčić & Stojanovski, 2022). Many of the rated journals can be found in the TOP Factor Database which provides a searchable repository of journals. Receiving a rating and inclusion in the TOP Factor Database may provide several tangible benefits to scientific journals, such as increased trustworthiness and reputation for rigour among the scientific community. Furthermore, journals that have higher levels of TOP implementation may attract high-quality submissions from researchers who are committed to open science. Implementing the practices outlined at higher levels of the TOP Factor provides opportunities for enhancing the quality of published work in a scientific journal overall. However, although some health psychology and behavioural medicine journals have previously been rated and included in the TOP Factor Database, the majority have not. Additionally, there has not been any systematic study of journals from health psychology and behavioural medicine to enable comparison and establish progress within this field. Moreover, until recently there has not been any established instruments or tools for calculating the TOP Factor

Table 1. TOP Factor (table reproduced with permission from Mayo-Wilson et al., 2021).

Standard	Level 1	Level 2	Level 3
Data citation	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article requires appropriate citation for data used consistent with the journal's author guidelines.	Article is not published until providing appropriate citation for data following journal's author guidelines.
Data transparency	Articles must state whether or not data are available.	Articles must have publicly available data, or explain why ethical/legal constraints prevent it.	Articles must have publicly available data and must be used to computationally reproduce or confirm results prior to publication.
Analytical code transparency	Articles must state whether or not code is available.	Articles must have publicly available code, or explain why ethical/legal constraints prevent it.	Articles must have publicly available code and must be used to computationally reproduce or confirm results prior to publication.
Materials transparency	Articles must state whether or not materials are available.	Articles must have publicly available materials, or explain why ethical/legal constraints prevent it.	Articles must have publicly available materials and must be used to computationally reproduce or confirm results prior to publication.
Reporting guidelines	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Study pre-registration	Articles will state if work was pre-registered.	Article states whether work was pre-registered and, if so, journal verifies adherence to pre-registered plan.	Journal requires that confirmatory or inferential research must be pre-registered.
Analysis plan pre-registration	Articles will state if work was pre-registered with an analysis plan.	Article states whether work was pre-registered with an analysis plan and, if so, journal verifies adherence to pre-registered plan.	Journal requires that confirmatory or inferential research must be pre-registered with an analysis plan.
Replication	Journal encourages submission of replication studies.	Journal will review replication studies blinded to results.	Registered Reports for replications as a regular submission option.
Publication bias	Journal states that significance or novelty are not criteria for publication decisions.	Journal will review (novel) studies blinded to results.	Journal accepts Registered Reports for novel studies as a regular submission option.
Open science badges	Journal awards 1 or 2 open science badges	Journal awards all 3 open science badges	

and the majority of previous evaluations have used self-developed bespoke tools (Kianersi et al., 2023; Mayo-Wilson et al., 2021). In 2021, Mayo-Wilson and Grant established the Transparency of Research Underpinning Social Intervention Tiers (TRUST) initiative, part of which involved the rigorous development of a tool to rate journal policy adherence to TOP guidelines, the TRUST policy evaluation tool (Kianersi et al., 2023; Mayo-Wilson et al., 2021).

The primary aim of this study was to evaluate the adoption of TOP standards within health psychology and behavioural medicine journal policies. Our secondary aim was to calculate the interrater agreement and interrater reliability of the TRUST policy evaluation tool and the TOP Factor rating process used within this study.

2. Materials and methods

2.1. Study design

We conducted a cross-sectional study to examine the adoption of TOP standards in health psychology and behavioural medicine journal policies. Eight raters used the TRUST tool during an in-person 'hackathon' event at the European Health Psychology Society Conference (EHPS) 2024. We registered our study protocol on the Open Science Framework on August 29th 2024. The protocol is available at: https://doi.org/10.17605/OSF.IO/GNMEQ.

Table 2. Health psychology and behavioural medicine journals identified.

Journals not in the TOP Factor Database	Journals already rated in the TOP Factor Database (date rated)
1. Health Psychology Review	1. Health Psychology and Behavioral Medicine (rated in Jan 2024)
2. Psychology and Health	2. Health Psychology (Nov 2022)
3. Annals of Behavioral Medicine	3. British Journal of Health Psychology (Nov 2022)
4. Translational Behavioral Medicine	4. European Journal of Health Psychology (Nov 2022)
5. Journal of Behavioral Medicine	5. Journal of Health Psychology (Nov 2022)
6. Behavioral Medicine	6. International Journal of Clinical and Health Psychology (Nov 2022)
7. International Journal of Behavioral Medicine	7. Journal of Occupational Health Psychology (Nov 2022)
8. Health Psychology Open	8. Applied Psychology-Health and Wellbeing (July 2023)
9. Health Psychology Report	9. Psychology, Health and Medicine (Feb 2024)
10. Behavioral Sleep Medicine	10. *To be excluded*: Health Psychology Bulletin (Mar 2023)

2.2. Journal selection process

A search for eligible journals was conducted on Clarivate/Web of Science on August 16th, 2024 using the keywords 'health psychology' and 'behavioral/behavioural medicine'. Journals were deemed eligible for inclusion in the study if they had an explicit focus on health psychology or behavioural medicine. This was screened for by selecting journals with any combination of the words 'health' and 'psychology', or 'behaviour/behavior' in their title. This search yielded 20 eligible journals, of which ten had previously received ratings on www.topfactor.org. We included the ten previously-rated journals in the study, as the majority of these ($n = 8$) had been rated prior to 2024, and journal policies may have been updated in the intervening time since they were initially rated. The other ten journals identified had not been previously rated; of these, one journal (*Health Psychology Bulletin*) had ceased production as of June 2023, so it was not included in the study. This resulted in a total overall sample of 19 journals (see Table 2).

2.3. Rater training and preparation

Ahead of the hackathon, all raters completed standardised training which included reviewing the original TOP guidelines paper by Nosek et al. (2015) and the TOP rubric to familiarise themselves with the transparency standards in more detail. Raters also reviewed the TRUST policy evaluation tool and watched a 90-minute training video on rating journals using the TOP guidelines that was uploaded by the Center for Open Science to the Open Science Framework (available at: https://osf.io/tf4yn). On the morning of the hackathon, raters participated in a refresher training session delivered by ET which consisted of a presentation on TOP standards and a demonstration of the TRUST tool used on one example journal from a different discipline to health psychology and behavioural medicine. Raters then practiced rating another sample journal using the TRUST tool as a group.

2.4. Data collection and rating processes

In the week prior to the hackathon, relevant policy documents for each journal were identified and uploaded on the Open Science Framework by ET, EN, RC and CD. Where relevant information was absent from a journal's policy documents but provided in the publisher policy, the publisher policy documents were also uploaded. Where relevant information was absent from a journal's policy documents but provided in the publisher policy, the publisher policy documents were also uploaded. In most instances journal policies provided sufficient information to make a judgement, however in seven instances publisher policy information was used where journal policies did not address the standard. In one instance, the journal policy information was deemed to provide insufficient information so publisher policy information was also used.

The hackathon took place during the 38th EHPS Conference in Cascais, Portugal, on Tuesday, September 3rd, 2024. Eight raters participated in the hackathon. Of these, two of the raters evaluated four journals each, and the other six raters evaluated five journals each. The raters were randomly allocated using the online random number generator 'Pickerwheel' (https://pickerwheel.com/tools/random-

number-generator/). Additionally, raters were allocated such that they could not be assigned to rate journals with which they were affiliated or had a conflict of interest (e.g., member of the editorial board).

Raters used the TRUST Journal Policy Evaluation Tool (Mayo-Wilson et al., 2021) to rate journals with respect to each of the ten TOP factor domains. This rating instrument was designed to minimise ambiguity when rating journal policies in terms of the TOP factor domains. The tool comprises a series of signalling questions which for the most part have 'yes/no' binary response options. These questions were set up within a Google Forms survey (Supplementary File 1) with display logic, such that answers would inform the appearance of subsequent questions within each TOP domain. For the Data Citation standard, the TRUST tool splits this into (a) data and (b) code citations to be assessed separately, and the final score for the Data Citation standard is the higher of the two ratings. Once the rating for each journal had been completed, the raters then used an algorithm (Supplementary File 2) developed by the TRUST team to calculate the TOP score for each domain and the journal's overall TOP score using the TOP factor rubric. The algorithm assigns a score of 0–3 according to the level achieved. For the Study Pre-registration standard there was ambiguity in the TRUST tool algorithm; as such we used the TOP factor rubric to calculate the score for this standard. Next, scores for each standard were entered into a data extraction spreadsheet (Supplementary File 3) along with the following metadata: the name of the journal, the journal's ISSN, the URL to the journal's webpage, the URL to the journal's author guidelines, and the name of the publisher. For each domain, all raters included a justification for the score that they provided. In most cases, this consisted of a quoted extract from the author guidelines supporting the score that was given. In the event that policy made no mention of the relevant domain (e.g., data citation), the justification was entered as 'Not mentioned' to indicate this was absent from the policy.

All journals were evaluated by two independent raters. Raters completed their scoring of journal policies independently, and then met in pairs to discuss their ratings and reach consensus on a final agreed rating. Where consensus could not be reached, these were resolved through consultation with a third rater. Since the hackathon was an in-person event, it facilitated raters to complete their independent evaluations simultaneously. Once the scores for each journal had been finalised, the data was exported in.xlsx format for analysis.

2.5. Statistical analysis

We used descriptive statistics to report the TOP standard ratings within and across all journals and the median and interquartile range (IQR) using JASP (Version 0.17.3). Additionally, we calculated inter-rater reliability (IRR) and inter-rater agreement (IRA) across all (a) health psychology journals and (b) TOP standards in R 4.4.1 using RStudio to assess alignment and consistency between raters. As the level of adoption of each TOP standard is on an ordinal scale (range = 0–3), we calculated intraclass correlation coefficients (ICC) as indicators of reliability. Specifically, we used the two-way random-effects ICC model with the 'consistency' and 'single rater' definitions in our analysis (Koo & Li, 2016). We interpreted the ICC values based on the lower bound of the 95% CI: < 0.50 = poor; 0.50–0.75 = moderate; 0.75–0.90 = good; > 0.90 = excellent (Koo & Li, 2016). We calculated the overall percentage agreement for each two raters per journal as an indicator of IRA. Here, overall agreement was defined as the number of cases in which raters agreed exactly relative to the total number of their ratings.

All data pre-processing, management, visualisation, and analysis steps were conducted in R 4.4.1 using RStudio (R Development Core Team, 2023). We used the tidyverse (Mayo-Wilson et al., 2021; Wickham et al., 2019) for all data wrangling parts and the irr package for IRR analysis (Gamer, 2019). To facilitate the reproducibility of our results and descriptives, we share annotated code as RMarkdown files on the OSF (https://osf.io/yxbtv/).

2.6. Deviations from protocol

In our study pre-registration and protocol (https://osf.io/yxbtv/) we had intended to report ICCs as indicators of IRR and IRA. However, given the very low adoption rates of TOP standards across

journals and the resulting lack of variability in the ratings, we decided to complement the ICC with the overall percentage agreement as an indicator of IRA (see Figure 1 for a visual explanation). Reporting the agreement percentage helps to more clearly appraise two specific cases. First, very low ICC values that result from, for example, raters agreeing to mostly award '0' but differing for one particular rating. Here, the ICC values appear very low due to floor effects in the data while the agreement is still high (and above chance). Second, cases of lacking ICC values as these could not be estimated due to low levels of variability while the overall agreement is still high

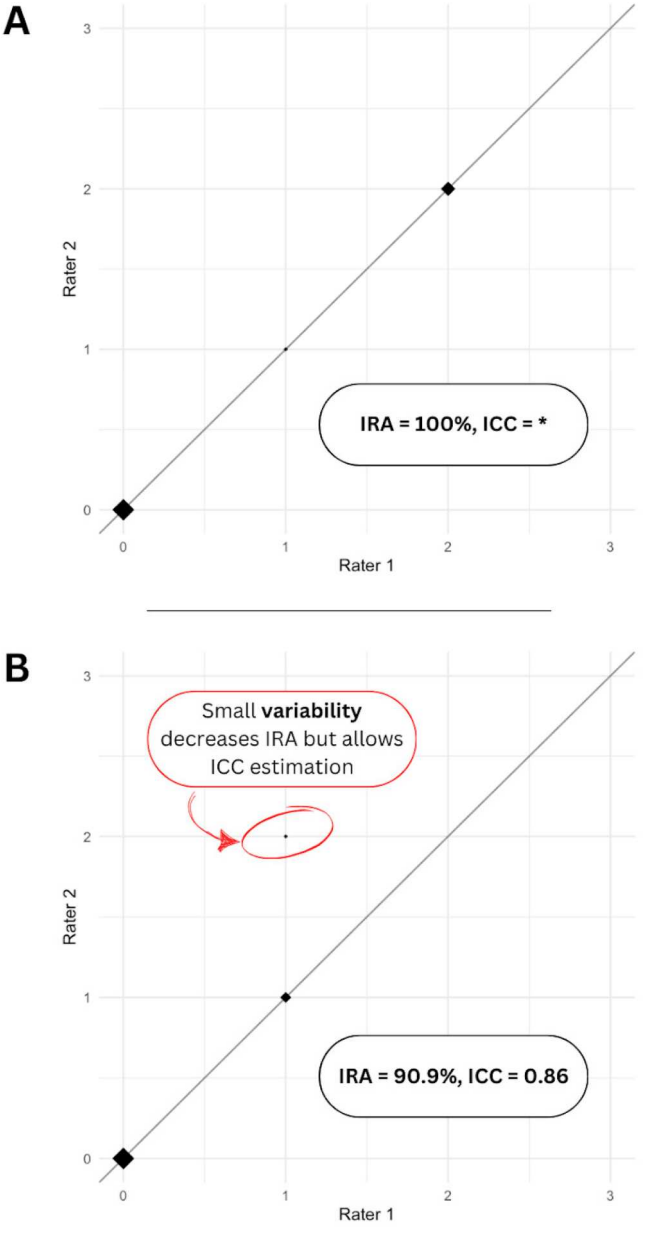

Figure 1. Differences in inter-rater agreement and inter-rater reliability scores in cases of low variability in ratings.

Note: The diagonal line indicates perfect agreement between both raters; the sizes of squares indicate the actual frequency of ratings between raters for the two examples (A) Health Psychology and Behavioral Medicine, and (B) Psychology, Health and Medicine.

(Figure 1). Thus, by combining two different indicators, we are able to more clearly explain the heterogeneity in the results. We also had initially randomised the journals across 9 raters, however due to unforeseen circumstances we had 8 raters present during the hackathon and re-randomised journals accordingly.

3. Results

3.1. Journal ratings

The scores for individual journals are provided in Table 3. As displayed in Figure 2, out of a total possible score of 29, final sum scores for TOP adoption across the 19 journals ranged from 1 (*Behavioral Medicine*) to 13 (*Translational Behavioral Medicine).* The median (IQR) score was 6.0 (6.0).

3.2. Adoption of standards across journals

3.2.1. Citation standards – data and code citation

Eleven of the 19 journals (57.9%) adopted the data citation standard to some extent. Of these, seven (36.8%) had adopted the standard to level 1, while the remaining four (21.1%) adopted it to level 2. However, only one journal (5.3%) had also adopted the code citation standard, which was adopted at level 2. Although the TRUST tool requires the evaluation of data and code citation standards separately, the final algorithm combines these to provide an overall Citation Standard score. As such, the final score for the overall Citation Standard was the same as the data citation, i.e., seven (36.8%) at level 1 and four (21.1%) at level 2. In total, therefore, a sum score of 15 was obtained for citation standards across the journals, as indicated in Figure 3. The median (IQR) for Citation Standard was 1.0 (1.0).

3.2.2. Transparency standards – data, analysis code and materials

The data transparency standard was adopted to some extent by 13 of the 19 journals (68.4%). Ten journals adopted this standard at level 1 (52.6%), while the remaining three adopted this at level 2 (15.8%). The sum of the scores provided an overall score of 16 across the journals for data transparency adopted, and a median (IQR) of 1.0 (1.0).

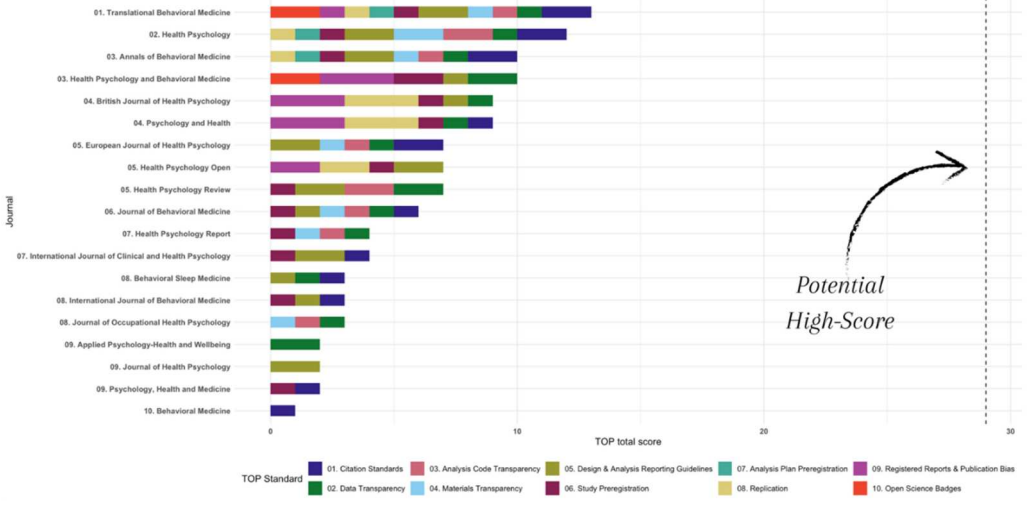

Figure 2. Ordered ranking of health psychology journals based on their respective adoption of TOP standards into their guidelines.
Note: Colours indicate different TOP standards while the length of each bar indicates the respective level of adoption within each standard.

Table 3. TOP Standard scores for individual journals.

	Citation			Transparency				Registration		Replication	Publication Bias	Badges	
Journal	1. Data Citation	1. Code Citation	Citation Standards – (higher of the two scores)	2. Data Transparency	3. Analysis Code Transparency	5. Materials Transparency	6. Design & Analysis Reporting Guidelines	7. Study Pre-registration	8. Analysis Plan Pre-registration	9. Replication	9. Registered Reports & Publication Bias	10. Open Science Badges	Total
Annals of Behavioral Medicine	2	0	2	1	1	1	2	1	1	1	0	0	**10**
Applied Psychology-Health and Wellbeing	0	0	0	2	0	0	0	0	0	0	0	0	**2**
Behavioral Medicine	1	0	1	0	0	0	0	0	0	0	0	0	**1**
Behavioral Sleep Medicine	1	0	1	1	0	0	1	0	0	0	0	0	**3**
British Journal of Health Psychology	0	0	0	1	0	0	1	1	0	3	3	0	**9**
European Journal of Health Psychology	2	2	2	1	1	1	2	0	0	0	0	0	**7**
Health Psychology	2	0	2	1	2	2	2	1	1	1	0	0	**12**
Health Psychology and Behavioral Medicine	0	0	0	2	0	0	1	2	0	0	3	2	**10**
Health Psychology Open	0	0	0	0	0	0	2	1	0	2	2	0	**7**
Health Psychology Report	0	0	0	1	1	1	0	1	0	0	0	0	**4**
Health Psychology Review	0	0	0	2	2	0	2	1	0	0	0	0	**7**

International Journal of Behavioral Medicine	1	0	1	0	0	0	1	1	0	0	0	0	**3**
International Journal of Clinical and Health Psychology	1	0	1	0	0	0	2	1	0	0	0	0	**4**
Journal of Behavioral Medicine	1	0	1	1	1	1	1	1	0	0	0	0	**6**
Journal of Health Psychology	0	0	0	0	0	0	2	0	0	0	0	0	**2**
Journal of Occupational Health Psychology	0	0	0	1	1	1	0	0	0	0	0	0	**3**
Psychology and Health	1	0	1	1	0	0	0	1	0	3	3	0	**9**
Psychology, Health and Medicine	1	0	1	0	0	0	0	1	0	0	0	0	**2**
Translational Behavioral Medicine	2	0	2	1	1	1	2	1	1	1	1	2	**13**
Total	**15**	**2**	**15**	**16**	**10**	**8**	**21**	**14**	**3**	**11**	**12**	**4**	

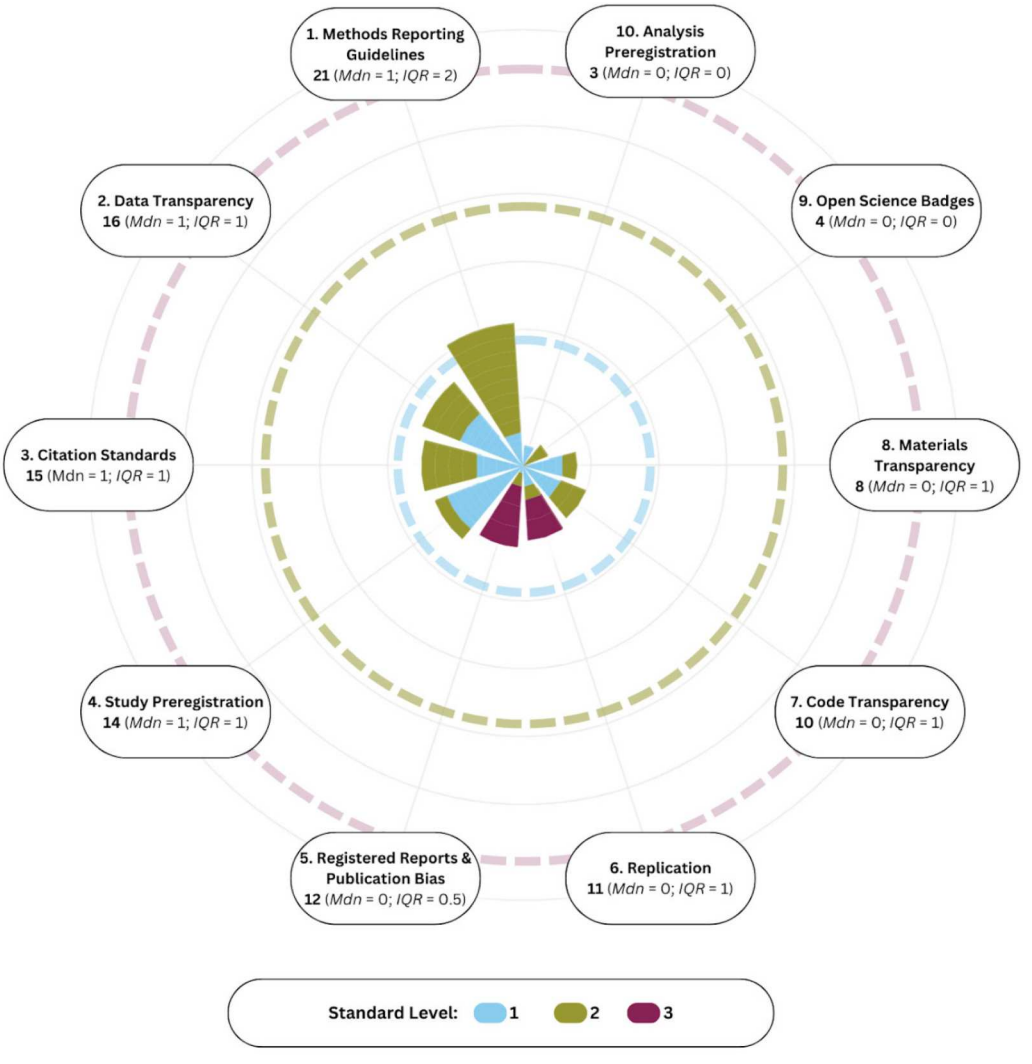

Figure 3. Radar plot indicating overall adoption of TOP standards.

Note: Colours indicate different levels of adoption of TOP standards (stringency; see legend). Solid bars indicate overall adoption per TOP standard across all journals. Dashed lines indicate overall levels of adoption if all journals had adopted standard level 1 (light blue), level 2 (reef gold), or level 3 (lipstick; except the maximum level for OS badges of 2). Mdn indicates median values, IQR indicate the interquartile range.

The analysis code transparency standard was less commonly adopted than data transparency, being adopted by only eight journals (42.1%). Of these, six were adopted at level 1 (31.6%) and only two were adopted at level 2 (10.5%). The total sum score for analysis code transparency was 10, with a median (IQR) of 0 (1.0).

The materials transparency standard was adopted by seven of the 19 journals (36.8%). Of these, six adopted at level 1 (31.6%) and one at level 2 (5.3%). The total sum score for materials transparency across journals was 8, with a median (IQR) of 0 (1.0).

3.2.3. Reporting guidelines standard

The reporting guidelines standard was adopted by 13 journals in total (68.4%), with five (26.3%) adopting this at level 1 and eight (42.1%) at level 2. The sum score for reporting guidelines across the journals was 21, while the median (IQR) was 1.0 (2.0).

3.2.4. Pre-registration standards – study and analysis plans

The study pre-registration standard was adopted by 13 journals (68.4%). Twelve (63.2%) adopted this at level 1 and one at level 2 (5.3%). The sum score across journals was 14 and the median (IQR) was 1.0 (1.0).

The analysis plan pre-registration standard was less common, being adopted in only three journals (15.8%) and all of these adopted this standard at level 1. The sum score was therefore 3, with a median (IQR) of 0 (0).

3.2.5. Replication standard

Replication was adopted to some extent by six journals (31.6%). Three adopted this at level 1 (15.6%); one at level 2 (5.3%) and two at level 3 (10.5%). The sum score across journals was 11, and the median (IQR) was 0 (1.0).

3.2.6. Registered Reports and publication bias standard

The Registered Report and publication bias standard was adopted by 5 journals (26.3%). One adopted this at level 1 and one at level 2 (5.3% each respectively), and three adopted this at level 3 (15.8%). The sum score across journals was 12, and the median (IQR) was 0 (0.5).

3.2.7. Open science badges

Open Science badges were adopted by only two journals (10.5%), both of whom adopted this at level 2. The sum score was therefore 4, with a median (IQR) of 0 (0).

3.3. Inter-rater agreement and inter-rater reliability of ratings across journals

Across all journals, IRA ranged from 55% to 100% but was higher than 82% for 12 out of all 19 journals assessed (Table 4). In contrast, IRR as indicated by ICCs was *poor* for 58% of the journals, and *moderate* for the remaining 10%, according to standards of Koo & Li (2016). IRR was not deemed to be *excellent* for any journal due to estimation problems with the ICCs as previously outlined (Figure 1).

3.4. Inter-rater agreement and inter-rater reliability of ratings across TOP standards

Overall, IRA across TOP standards ranged from 68% to 100% but was over 84% for more than half of all TOP standards assessed (Table 5). In contrast, out of the 10 standards, IRR was *poor* for six and *good* for two standards. Only the Materials Transparency standard was rated with both poor agreement and reliability. In most other cases, the agreement between raters was high (and above chance) while reliability was considered *poor* or *moderate*. Four domains (Data Transparency, Materials Transparency, Design and Analysis Reporting Guidelines and Analysis Plan Preregistration) had the most disagreements between raters (five in each) (Supplementary File 3). Through team discussion, it was deemed that in seven instances this was due to simple human error, e.g., one rater just missing information during review that the other rater had identified (potentially due to time pressures on the day). However, it was felt that these human errors were often driven by the challenges in finding the relevant information within journal policies, and in additional 11 instances a rater missing information was specifically deemed to be due to the lack of clarity and consistency within the journal policies, e.g., where the same standard was referred to in more than once place, where data and materials were explicitly referred to in some areas but not in others or where requirements were in place for some types of manuscripts but not others. In five instances, the disagreement was hypothesised to be also attributable to raters' interpretation of the policy assessment process, for example getting confused between the scoring of requirements for the availability of data or materials versus requirements for statements of data or materials availability.

Table 4. Inter-rater agreement and inter-rater reliability of ratings across journals.

Displayed are measures of Inter-Rater Agreement (IRA) and Intra-Rater Reliability (ICC) with 95% confidence intervals (CIs) and overall rating interpretations for all sets of ratings per journal policies and procedures.

	IRA measure	IRR measure			
Journal Name	Overall agreement percentage	ICC [95% CI]	F value	*p* value	Interpretation
Annals of Behavioral Medicine	0.82	0.83 [0.49; 0.95]	11.0	<.001	poor
Applied Psychology-Health and Wellbeing	1.00	*	*	*	*
Behavioral Medicine	0.91	0.76 [0.33; 0.93]	7.4	.002	poor
Behavioral Sleep Medicine	0.82	0.8 [0.38; 0.95]	9.0	.002	poor
British Journal of Health Psychology	0.82	0.85 [0.54; 0.96]	12.5	<.001	moderate
European Journal of Health Psychology	0.64	0.56 [−0.02; 0.86]	3.5	.03	poor
Health Psychology	0.73	0.83 [0.49; 0.95]	10.8	<.001	poor
Health Psychology and Behavioral Medicine	1.00	*	*	*	*
Health Psychology Open	0.91	0.67 [0.15; 0.90]	5.0	.009	poor
Health Psychology Report	0.64	0.41 [−0.19; 0.8]	2.4	−.09	poor
Health Psychology Review	0.73	0.73 [0.29; 0.92]	6.2	.004	poor
International Journal of Behavioral Medicine	1.00	*	*	*	*
International Journal of Clinical and Health Psychology	1.00	*	*	*	*
Journal of Behavioral Medicine	0.73	0.57 [−0.01; 0.86]	3.7	.03	poor
Journal of Health Psychology	1.00	*	*	*	*
Journal of Occupational Health Psychology	1.00	*	*	*	*
Psychology and Health	0.73	0.14 [−0.47; 0.66]	1.3	.33	poor
Psychology, Health and Medicine	0.91	0.86 [0.58; 0.96]	13.8	<.001	moderate
Translational Behavioral Medicine	0.55	0.66 [0.15; 0.90]	5.0	.009	poor

Note: IRA was assessed as total percent of agreement for each two raters per journal. ICC indicates the intra-class correlation as a measure of inter-rater reliability. Interpretations are based on the lower bound of the 95% CI: < 0.50 = poor; 0.50–0.75 = moderate; 0.75–0.90 = good; >0.90 = excellent (Koo & Li, 2016). * indicates no variability between raters´ assessments, an ICC could not be estimated (see Figure 1).

4. Discussion

Overall, we found that the TOP guidelines have to date been poorly adopted within health psychology and behavioural medicine journal policies. Standards related to use of reporting guidelines and data transparency were the most addressed in policies, compared to standards related to pre-registration of study analysis plans and citation of code which were the least addressed. While raters mostly agreed in their assessment of journal policies, the overall low adoption rates of TOP guidelines resulted in a low inter-rater reliability of scoring.

Our findings regarding the implementation of TOP standards align with previous similar research into health-related journal policies. In a 2019 evaluation of pain journal policies' adoption of TOP standards (Cashin et al., 2021), the study pre-registration and reporting guidelines standards was also the most addressed, by 8/10 and 7/10 journals respectively. The pre-registration of analysis plans standard was not incorporated into any journal policies. Similarly, in an evaluation of sport science journal policies (Hansford et al., 2022), the reporting guidelines standard was the most addressed, with 55% of the journals addressing the standard at either Level 1 or 2. Again no journal adopted the pre-registration of analysis plans standard to any extent. In contrast, an assessment of journal policies from multiple disciplines across the Scopus content database (Pataržić & Stojanovski, 2022) found that data citation was the most adopted standard addressed by 70% journals, followed by data transparency (19%). When the study compared TOP adoption across disciplines, it found that health-specific journals ranked lower than multidisciplinary or social science journals (Pataržić & Stojanovski, 2022).

For those working in health psychology and behavioural medicine research, there are various aspects of TOP standards that require further consideration. For example, some standards were adopted more frequently than others, with the reporting guidelines standard being the most

Table 5. Inter-rater agreement and inter-rater reliability of ratings across TOP standards.

Displayed are measures of Inter-Rater Agreement (IRA) and Intra-Rater Reliability (ICC) with 95% confidence intervals (CIs) and overall rating for all sets of ratings per TOP standard across all journals.

	IRA measure	IRR measure			
TOP Standard	Overall agreement percentage	ICC [95% CI]	F value	*p* value	Interpretation
Citation Standards	0.95	0.95 [0.89; 0.98]	43.0	<.001	good
Data Citation	0.95	0.95 [0.89; 0.98]	43.0	<.001	good
Code Citation	0.95	0 [−0.44; 0.44]	1.0	.50	poor
Data Transparency	0.74	0.77 [0.49; 0.90]	7.6	<.001	poor
Code Transparency	0.84	0.83 [0.61; 0.93]	10.8	<.001	moderate
Research Materials Transparency	0.68	0.46 [0.02; 0.75]	2.7	.02	poor
Design & Analysis Transparency	0.74	0.80 [0.56; 0.92]	9.1	<.001	moderate
Registration of Studies	0.79	0.74 [0.42; 0.89]	6.6	<.001	poor
Registration of Analysis Plans	0.74	0.37 [−0.10; 0.69]	2.1	.06	poor
Replication	0.84	0.45 [0.01; 0.75]	2.6	.02	poor
Registered Reports & Publication Bias	1.00	*	*	*	*
Open Science Badges	1.00	*	*	*	*

Note: IRA indicates 'inter-rater agreement' and was assessed as total percent of agreement for each two raters per journal. ICC indicates the intra-class correlation as a measure of inter-rater reliability. Interpretations are based on the lower bound of the 95% CI: < 0.50 = poor; 0.50–0.75 = moderate; 0.75–0.90 = good; >0.90 = excellent (Koo & Li, 2016). Data Citation and Code Citation as subdomains of Citation Standards are highlighted in italic. * indicates no variability between raters´ assessments, an ICC could not be estimated (see Figure 1).

widely adopted. The majority of included journals (68%) scored at least 1 on this standard. This may be due to the availability and awareness of resources regarding reporting guidelines, such as the EQUATOR network (https://www.equator-network.org), an international initiative involving researchers, journal editors, peer reviewers and funders established in 2006 focusing on promoting transparency and accurate research reporting through use of reporting guidelines. As such, the importance of reporting guidelines is increasingly recognised, with several research funders now endorsing the use of reporting guidelines (Diong et al., 2021; National Institute for Health and Care Research. Reporting guidelines, 2024). This success likely reflects the importance of collaboration between all relevant stakeholders to bring about change, and the need for investment and adequate resourcing of supports to enable this. As such, facilitating greater implementation of TOP standards into health psychology and behavioural medicine journal policies will likely require collaboration between all relevant parties. In other words, journal editors and publishers that are willing to review and lead an update of their policies could work alongside health psychology and behavioural medicine researchers and open science advocates that are willing to offer help and capacity to support this. This would also help ensure coherence between journal and publisher-level policies and thus aid clarity for researchers in adhering to these. For example, the journal *Health Psychology and Behavioral Medicine* recently introduced data notes and registered reports as a new article format (Norris et al., 2024). To support the implementation of this within the journal, the co-chairs of the EHPS Open Science Special Interest Group were invited to join the editorial board to establish and publicise a special call for papers, including the publication of guidance on these new formats (Norris et al., 2024).

The next most addressed standards within health psychology and behavioural medicine journals were those related to research data (i.e., data transparency and citation). This may reflect a growing emphasis in recent times towards open data, for example with the advent of FAIR data policies and increasing mandates from national and international funding bodies to ensure that data is made available (European, C., R. Directorate-General for, and Innovation, Horizon Europe, open science – Early knowledge and data sharing, and open collaboration, 2021; The White House, 2022). For example, Horizon Europe funding requires all data from funded studies to be made available (European Commission, 2024). Although this is an important shift and is likely to have been a key impetus towards improving journals' policies regarding data, data itself is also only one element of the wider

picture. For example, ensuring transparency of code as well as study materials and analysis procedures are just as important to ensure maximum impact and usability of research data and study findings. In our study, although 11 out of 19 journals addressed data citation standards and 13/19 addressed data transparency, only one journal (*European Journal of Health Psychology*) included sufficient detail regarding code citation, while eight addressed code transparency and seven addressed materials transparency. The reasons for this low adoption may be partly due to limitations of the TOP guidelines themselves in that there are discrepancies across how they are described in different documents and rubrics (Grant et al., 2024), with the data citation standard sometimes referring only to data, and other times referring to date, code and materials. As TOP standards are updated to address this and other discrepancies (Grant et al., 2024), one relatively easy interim change that many journals could make is to extend their existing data policies to also include reference to code. Again reflecting the importance of collaborative efforts in achieving change, other relevant stakeholders such as research funders should also consider all of the elements of research transparency beyond just data and extend their requirements and resources accordingly.

Although journals scored higher on data transparency and citation standards compared to other standards, they still achieved only half of the total possible score and therefore provide much room for improvement. Naaman et al. (2023) recently surveyed 88 journal editors regarding the implementation of TOP guidelines within their journals, including a qualitative exploration of their perceptions of the barriers and enablers to doing this. The study found that different types of study designs can pose challenges for editors regarding the implementation of TOP standards into policies (Naaman et al., 2023). For example, much of the existing open science and open data guidance and resources were originally developed with quantitative research studies in mind, and are less appropriate for qualitative studies (Campbell et al., 2023). For journals that publish a substantial amount of qualitative research, this may represent a challenge to implementing data-related standards. However in recent times, there has been much progress in relation to qualitative research and open science (Branney et al., 2023). There is ongoing debate about the extent to which mandated open data practices in journals should apply to qualitative research (Prosser et al., 2024), yet significant strides have been made in the increased uptake of pre-registration of qualitative research (Evans et al., 2023; Haven & Van Grootel, 2019). As such, journal policies should be updated on an ongoing basis to reflect the pace of change in research itself.

It is important to acknowledge that full adoption (i.e., Level 3) of all standards may not be achievable or even desirable for many journals or disciplines. However, it is also important to acknowledge that the TOP guidelines were developed with the need for flexibility (Nosek et al., 2015), and the intention that journals could implement them flexibly according to disciplinary norms. Nonetheless, it is clear from the results of our study that health psychology and behavioural medicine journals have substantial room for improvement. However, achieving change in any complex system such as a journal editorial process is never straightforward. For many journal editors, implementing TOP standards within their journal's policies may not be an issue of not wanting to, but rather may reflect a lack of capacity to do so within their role. For example, the study by Naaman et al. (2023) showed that most editors were supportive of implementing TOP within their journals, but that it was not a priority compared to other editorial responsibilities. Additionally, many felt they lacked the time, authority and resources to make changes (Naaman et al., 2023). There is a substantial workload involved in editorial roles and making any change to policies typically involves substantial timeframes. However, it must be acknowledged that many scores in our study could be improved relatively easily by improving clarity and reducing ambiguity in relation to author submission guidelines. The TRUST policy evaluation tool proved very useful in many aspects in reducing the ambiguity of ratings, however, we still struggled to conduct some ratings due to wording in journal policies such as 'authors are expected to' or 'authors should' which do not clarify whether a particular standard is recommended or mandatory, and may have resulted in policies being rated at a lower level. While we identified many of these issues during our training and thus ensured clarity within our team for rating in advance, there were many other areas in policies lacking clarity and consistency

which affected our inter-rater agreement. As such, ensuring clarity and coherence in policies both at journal and publisher-level would likely be of benefit for improving the accuracy of future ratings as well as for improving the efficiency of journal submission systems, both for authors and for editors handling submissions. For example, any differences between journal and publisher-level policies would be clearly delineated and researchers would know exactly what they needed to do in order from the policy wording.

4.1. Strengths and limitations

The current study has several strengths and limitations that need to be acknowledged. Firstly, when it comes to our methodology, we opted to use the TRUST tool instead of the TOP Rubric due to its robust development, previous testing, and structured approach to aid inter-rater agreement (Kianersi et al., 2023; Mayo-Wilson et al., 2021). Furthermore, our raters undertook careful training to use this tool and to assess journal guidelines and policies as accurately as possible. This is the first study to use the TRUST form outside of the original development team. However, we identified some challenges in the use of the TRUST tool, such as an issue with the scoring algorithm for the study pre-registration standard. Additionally, as previously mentioned there are some issues with the TOP standards themselves, such as the consistency of how terminology is applied (e.g., disclose, require, verify), and the combination of several components in one standard (e.g., data citation and code citation within the Citation Standard.

Challenges with implementing the TRUST tool and interpretation of TOP standards may in part explain our low reliability scores across journals, together with challenges in interpreting journal policy summaries consistently. Beyond these procedural challenges, another reason for our low inter-rater reliability scores might be the overall low levels of TOP standard adoption that led to mostly zero ratings. This, in turn, led to low levels of variability between raters that hindered or biased the estimation of ICC values (e.g., ratings for Code Citation); however this is similar to previous studies findings (Kianersi et al., 2023). Another limitation of our study is the fact that we focused solely on the adoption of TOP standards within policy documents, and did not examine the actual implementation of these standards within journal submission systems or published articles published in the journal.

4.2. Future research

The TOP standards are currently due to be updated to address these concerns (Grant et al., 2024), and as such the specific scores for journals included in our study are likely to change. However, we believe the specific scores themselves are of less importance than the systematic identification of areas across health psychology and behavioural medicine journals where improvements can be made. Additionally, we believe this study overall highlights the potential contributions that stand to be made by journals in working towards improving open science practices in this discipline. As evidenced by Naaman et al. (2023), journal editorial teams face a set of barriers and enablers regarding the implementation of TOP guidelines within their journals. Awareness of such barriers and enablers is needed to facilitate greater adoption of guidelines across the board. Although Naaman et al highlighted several barriers that appear to be applicable to journals independent of their field (e.g., guidance and resources being predominantly developed for quantitative studies and therefore harder to implement by journals that publish qualitative studies as well), there are likely to be barriers and enablers more specific to the discipline of health psychology. For example, prior to the International Committee of Medical Journal Editors endorsing clinical trial registration, it would have likely been challenging for any single journal to implement a more stringent standard for study pre-registration. Making a joint decision on this front would likely have sent a clear normative message to researchers in the field, thus facilitating adoption in a manner similar to how the EQUATOR network facilitated the implementation and adoption of the reporting guidelines standard. Further qualitative research amongst editorial team members of health psychology and

behavioural medicine journals specifically could help identify similar opportunities for other standards where creating change as a single journal might be difficult or impossible, but a joint endorsement in a specific field may remove some of the barriers. Such research could also provide insight into what has helped early and more stringent adopters of these standards, which may facilitate adoption by other journals.

5. Conclusion

Given the direct relevance of health psychology and behavioural medicine research to the health and well-being of individuals and the common use of public funds within these fields, the transparency and trustworthiness of our science is crucial. Scientific journals play a key role in the verifiability and robustness of this research, and the TOP Guidelines provide a comprehensive and structured starting point for journals to achieve this. Given the potential challenges that journal editorial boards may face in implementing changes, identifying 'smaller wins' such as improved clarity in wording of author guidelines and coherence with their publisher's policies may be an easier place to start, as well as working collaboratively with all research stakeholders including funders and research communities to encourage greater awareness and practice of the wider spectrum of open science behaviours.

Disclosure statement

Elaine Toomey and Emma Norris are Associate Editors of the journal *Health Psychology and Behavioral Medicine*.
Ilona McNeill is an Editorial Board member of the journal *Psychology and Health*.
Felix Naughton is a Senior Editor of *Addiction* and an Editorial Consultant of the *British Journal of Health Psychology*.
Sean P. Grant is Chair of the TOP Advisory Board and Elaine Toomey is a member of the TOP Advisory Board.

Funding

This work was not supported by a research grant, however the European Health Psychology Society Special Interest Group fund covered costs for workshop catering and venue hire.

Authors' contributions

All authors made substantive intellectual contributions to, and were involved in drafting or revising the manuscript. All authors have read and approved the final manuscript. The following specific contributions are described according to the CREDIT taxonomy: *Conceptualization:* ET; *Data curation* and *Investigation:* All authors; *Formal analysis:* CJ, MK, AOM; *Methodology:* All authors; *Project Administration:* ET, *Software:* CJ, MK, AOM, SG; *Validation:* CJ, MK, AOM; *Visualization:* All authors; *Writing:* All authors.

References

Branney, P. E., Brooks, J., Kilby, L., Newman, K., Norris, E., Pownall, M., Talbot, C. V., Treharne, G. J., & Whitaker, C. M. (2023). Three steps to open science for qualitative research in psychology. *Social and Personality Psychology Compass*, *17*(4), e12728. https://doi.org/10.1111/spc3.12728

Campbell, R., Javorka, M., Engleton, J., Fishwick, K., Gregory, K., & Goodman-Williams, R. (2023). Open-Science guidance for qualitative research: An empirically validated approach for De-identifying sensitive narrative data. *Advances in Methods and Practices in Psychological Science*, *6*(4), 25152459231205832. https://doi.org/10.1177/25152459231205832

Cashin, A. G., Bagg, M. K., Richards, G. C., Toomey, E., McAuley, J. H., & Lee, H. (2021). Limited engagement with transparent and open science standards in the policies of pain journals: A cross-sectional evaluation. *BMJ Evidence-Based Medicine*, *26*(6), 313–319. https://doi.org/10.1136/bmjebm-2019-111296

Castro, E., Crosas, M., Garnett, A., Sheridan, K., & Altman, M. (2017). Evaluating and promoting open data practices in open access journals. *Journal of Scholarly Publishing*, *49*(1), 66–88. https://doi.org/10.3138/jsp.49.1.66

Cole, N. L., Reichmann, S., & Ross-Hellauer, T. (2023). Toward equitable open research: Stakeholder co-created recommendations for research institutions, funders and researchers. *Royal Society Open Science*, *10*(2), 221460. https://doi.org/10.1098/rsos.221460

De Angelis, C., Drazen, J. M., Frizelle, F. A., Haug, C., Hoey, J., Horton, R., Kotzin, S., Laine, C., Marusic, A., Overbeke, A. J. P. M., Schroeder, T. V., Sox, H. C., & Weyden, M. B. V. D. (2004). Clinical trial registration: A statement from the international committee of medical journal editors. *New England Journal of Medicine*, *351*(12), 1250–1251. https://doi.org/10.1056/NEJMe048225

Diong, J., Kroeger, C. M., Reynolds, K. J., Barnett, A., & Bero, L. A. (2021). Strengthening the incentives for responsible research practices in Australian health and medical research funding. *Research Integrity and Peer Review*, *6*(1), 11. https://doi.org/10.1186/s41073-021-00113-7

European Commission. (2024). Open science in Horizon Europe. 22 October 2024]; Available from: https://rea.ec.europa.eu/open-science_en.

European, C., R. Directorate-General for, and Innovation, Horizon Europe, open science – Early knowledge and data sharing, and open collaboration. (2021). Publications Office of the European Union.

Evans, T. R., Branney, P., Clements, A., & Hatton, E. (2023). Improving evidence-based practice through preregistration of applied research: Barriers and recommendations. *Accountability in Research*, *30*(2), 88–108. https://doi.org/10.1080/08989621.2021.1969233

Gamer, M. (2019). Irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.1.

Gardener, A. D., Hick, E. J., Jacklin, C., Tan, G., Cashin, A. G., Lee, H., Nunan, D., Toomey, E. C., & Richards, G. C. (2022). Open science and conflict of interest policies of medical and health sciences journals before and during the COVID-19 pandemic: A repeat cross-sectional study: Open science policies of medical journals. *JRSM Open*, *13*(11), 20542704221132139. https://doi.org/10.1177/20542704221132139

Grant, S., Corker, K. S., Mellor, D. T., Stewart, S. L., Cashin, A., Lagisz, M., Mayo-Wilson, E., Moher, D., Umpierre, D., Barbour, V., & Buck, S. (2024). TOP 2025: An update to the transparency and openness promotion guidelines. MetaArXiV PrePrints.

Grant, S., Mayo-Wilson, E., Kianersi, S., Naaman, K., & Henschel, B. (2023). Open science standards at journals that inform evidence-based policy. *Prevention Science*, *24*(7), 1275–1291. https://doi.org/10.1007/s11121-023-01543-z

Hagger, M. S. (2022). Developing an open science 'mindset'. *Health Psychology and Behavioral Medicine*, *10*(1), 1–21. https://doi.org/10.1080/21642850.2021.2012474

Hansford, H. J., Cashin, A. G., Bagg, M. K., Wewege, M. A., Ferraro, M. C., Kianersi, S., Mayo-Wilson, E., Grant, S. P., Toomey, E., Skinner, I. W., McAuley, J. H., Lee, H., & Jones, M. D. (2022). Feasibility of an audit and feedback intervention to facilitate journal policy change towards greater promotion of transparency and openness in sports science research. *Sports Medicine - Open*, *8*(1), 101. https://doi.org/10.1186/s40798-022-00496-x

Haven, T., & Van Grootel, L. (2019). Preregistering qualitative research. *Accountability in Research*, *26*(3), 229–244. https://doi.org/10.1080/08989621.2019.1580147

Kianersi, S., Grant, S. P., Naaman, K., Henschel, B., Mellor, D., Apte, S., Deyoe, J. E., Eze, P., Huo, C., Lavender, B. L., Taschanchai, N., Zhang, X., & Mayo-Wilson, E. (2023). Evaluating implementation of the transparency and openness promotion guidelines: Reliability of instruments to assess journal policies, procedures, and practices. *Advances in Methods and Practices in Psychological Science*, *6*(1), 25152459221149735. https://doi.org/10.1177/25152459221149735

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2017.10.001

Kwasnicka, D., ten Hoor, G. A., van Dongen, A., Gruszczyńska, E., Hagger, M. S., Hamilton, K., Hankonen, N., Heino, M. T. J., Kotzur, M., Noone, C., Rothman, A. J., Toomey, E., Warner, L. M., Kok, G., Peters, G.-J., & Luszczynska, A. (2021). Promoting scientific integrity through open science in health psychology: Results of the synergy expert meeting of the European health psychology society. *Health Psychology Review*, *15*(3), 333–349. https://doi.org/10.1080/17437199.2020.1844037

Lee, H., Lamb, S. E., Bagg, M. K., Toomey, E., Cashin, A. G., & Moseley, G. L. (2018). Reproducible and replicable pain research: A critical review. *Pain*, *159*(9), 1683–1689. https://doi.org/10.1097/j.pain.0000000000001254

Mayo-Wilson, E., Grant, S., Supplee, L., Kianersi, S., Amin, A., DeHaven, A., & Mellor, D. (2021). Evaluating implementation of the Transparency and Openness Promotion (TOP) guidelines: The TRUST process for rating journal policies, procedures, and practices. *Research Integrity and Peer Review*, *6*(1), 9. https://doi.org/10.1186/s41073-021-00112-8

Munafò, M. R., Chambers, C., Collins, A., Fortunato, L., & Macleod, M. (2022). The reproducibility debate is an opportunity, not a crisis. *BMC Research Notes*, *15*(1), 43. https://doi.org/10.1186/s13104-022-05942-3

Naaman, K., Grant, S., Kianersi, S., Supplee, L., Henschel, B., & Mayo-Wilson, E. (2023). Exploring enablers and barriers to implementing the transparency and openness promotion guidelines: A theory-based survey of journal editors. *Royal Society Open Science*, *10*(2), 221093. https://doi.org/10.1098/rsos.221093

National Institute for Health and Care Research. Reporting guidelines. (2024). 22 October 2024; Available from: https://www.journalslibrary.nihr.ac.uk/information-for-authors/reporting-guidelines/.

Norris, E., O'Mahony, A., Coyne, R., Varol, T., Green, J. A., Reynolds, J., & Toomey, E. (2024). Demystifying open science in health psychology and behavioral medicine: A practical guide to registered reports and data notes. *Health Psychology and Behavioral Medicine*, *12*(1), 2351939. https://doi.org/10.1080/21642850.2024.2351939

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B.,

Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Patarčić, I., & Stojanovski, J. (2022). Adoption of Transparency and Openness Promotion (TOP) guidelines across journals. *Publications*, *10*(4), 46. https://doi.org/10.3390/publications10040046

Prosser, A. M., Bagnall, R., & Higson-Sweeney, N. (2024). Reflection over compliance: Critiquing mandatory data sharing policies for qualitative research. *Journal of Health Psychology*, *29*(7), 653–658. https://doi.org/10.1177/13591053231225903

Segerstrom, S. C., Diefenbach, M. A., Hamilton, K., O'Connor, D. B., & Tomiyama, A. J. (2023). Open science in health psychology and behavioral medicine: A statement from the behavioral medicine research council. *Health Psychology*, *42*(5), 287–298. https://doi.org/10.1037/hea0001236

Taylor, N. J., & Gorman, D. M. (2022). Registration and primary outcome reporting in behavioral health trials. *BMC Medical Research Methodology*, *22*(1), 41. https://doi.org/10.1186/s12874-021-01500-w

The White House. (2022). OSTP issues guidance to make federally funded research freely available without delay. October 22, 2024; Available from: https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., & Kuhn, M. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*, 1686.

Wolfram, D., Wang, P., Hembree, A., & Park, H. (2020). Open peer review: Promoting transparency in open science. *Scientometrics*, *125*(2), 1033–1051. https://doi.org/10.1007/s11192-020-03488-4