Using the STROBE statement: survey findings emphasized the role of journals in enforcing reporting guidelines

Melissa Sharp, Lorenzo Bertizzolo, Roser Rius, Elizabeth Wager, Guadalupe Gómez, Darko Hren

Using the STROBE statement: Survey findings emphasized the role of journals in enforcing reporting guidelines

Melissa K Sharp*[1,2], Lorenzo Bertizzolo[2], Roser Rius[3], Elizabeth Wager[4,5], Guadalupe Gómez[3], Darko Hren[1]

[1] University of Split, Department of Psychology, Faculty of Humanities and Social Sciences. Split, Croatia
[2] Université de Paris, CRESS, INSERM, INRA, F-75004 Paris, France
[3] Universitat Politècnica de Catalunya-BarcelonaTech, Departament d'Estadística i Investigació Operativa. Barcelona, Spain
[4] University of Split, School of Medicine. Split, Croatia
[5] Sideview. Princes Risborough, United Kingdom

Corresponding Author:
Melissa K Sharp
Poljička cesta 35 Split, Croatia 2100
msharp@unist.hr; melissa.sharp@etu.parisdescartes.fr
+1 248 701 8109; +385 097 634 4546

Word Count: 3101

Abstract

Objective: To identify factors affecting the use of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement, specifically authors' attitudes towards and experiences with it.

Study Design and Setting: An online survey was distributed to authors of observational studies recruited via social media, personal network snowballing, and mass-mailings using targeted search strategies. Data on demographics, awareness, motivators, and usage were collected in conjunction with a modified Unified Theory of Acceptance and Use of Technology (UTAUT) scale on which confirmatory factor analysis (CFA) was performed.

Results: 1015 participants completed the survey. Of these, 185 (18.2%) indicated they had never heard of STROBE nor used it previously, 195 (19.2%) had heard of it but never used it, and 635 (62.6%) had used it. Journals promoting STROBE were both key motivators and awareness mechanisms; peers and educational workshops were also important influencing factors to a lesser degree. The internal consistency of the modified UTAUT scale was strong (Cronbach's alpha = 0.94). CFA supported a 4-factor model with 23 questions.

Conclusion: The endorsement of STROBE by journals is key to authors' awareness and use of the guideline. We tested and validated our scale which can guide future research on reporting guidelines.

Word Count: 197

Keywords: Observational studies, Guidelines as topic, Epidemiologic research design, Information dissemination/methods, STROBE

Running Head: STROBE Statement Online Survey for Authors

Introduction

Reporting guidelines provide a protective "cognitive net" against the fallibility of human memory and support the skills of expert professionals [1]. Authors of biomedical manuscripts are generally unaware of the existence or utility of reporting guidelines (RG) and those responding to peer reviewers often have problems adhering to the methodological standards proposed [2–4]. Many journals do not require a relevant RG checklist to be submitted with a manuscript, therefore, there is often no incentive for authors to complete one [5].

Some authors reject reporting guidelines, claiming that RG can be condescending and rigid [6,7]. It is unclear what maintains these attitudes. Therefore, it would be useful to understand factors affecting use [8]. This study was designed to explore how researchers view and interact with one reporting guideline, the STrengthening the Reporting of Observational studies in Epidemiology (STROBE) Statement. STROBE was created in 2007 to improve the reporting of observational studies (e.g., cross-sectional, cohort, case-control).

Many journals promote STROBE by requiring or recommending its use during the manuscript preparation process. However, endorsement rates are relatively low [9–13] and there is a diffusion of responsibility amongst journal editors, authors, and peer reviewers for RG compliance [3]. In order to better understand the current situation facing authors, we aimed to identify the personal and environmental facilitators, barriers, and motivators to using the STROBE Statement. With this information, we hoped to extend the practical value of STROBE and perhaps other RG.

Methods

Survey Design

We followed the CHERRIES guideline for online surveys (Supplemental File 1) [14]. Prior to distribution, we piloted the survey within the Methods in Research on Research (MiRoR) network [15], allowing collaborators to give feedback on content and functionality [16]. The University of Split School of Medicine Ethical Review Committee granted ethical approval.

The survey flow is presented in Figure 1 and the survey is in Supplemental File 2. All questions were forced response except for one optional open-ended question and mistakenly, the question asking about the respondent's country. After consenting to participate, adaptive questioning branched the survey based on participant's level of awareness and use of STROBE (i.e., never heard of, never used; heard of, never used; heard of, have used). After branching, participants were presented with questions about their interactions with STROBE (e.g., real or theoretical timing of use: writing a grant, or peer reviewing an article).

Next, all participants were presented 25 questions informed by the Unified Theory of Acceptance and Use of Technology (UTAUT) scale [17,18]. UTAUT is an amalgamation of eight dominant psychological and health technology assessment (HTA) theories and models that attempts to explain one's intention to use a piece of technology *and* their subsequent use behavior. The scale aims to explain information system usage behavior by measuring: Performance Expectancy (PE), Effort Expectancy (EE), attitude toward using technology, Social Influence (SI), Facilitating Conditions (FC), self-efficacy, anxiety, and behavioral intention to use the tool [17]. HTA systematically evaluates direct and indirect consequences of using a piece of health technology. It can tap into whether the technology works, for whom, and at what cost. [19]

We rephrased questions to be relevant to STROBE and kept the scale's four core constructs (PE, EE, SI, FC) (Figure 2). Each subscale contained several items to ensure reliability and validity. The final version contained nine Likert scale items from PE, six from EE, five from EE, four from FC, and one assessing the intention to use STROBE. Respondents rated statements on seven-point Likert-type scales from "strongly disagree" to "strongly agree" (Supplemental File 2).

Recruitment

Eligible participants were researchers involved in manuscript writing (within the past 10 years) reporting the results of observational studies. The survey was distributed from March 5 to August 31, 2018.

Survey recruitment used several snowball and purposive sampling routes. Firstly, MKS invited her professional network and those involved in the Methods in Research on Research (MiRoR) consortium [15] to participate. Next, the survey was promoted through social media, primarily Twitter. We then emailed the editors of 257 biomedical journals identified in another study [9,20] and asked them to invite their authors to participate (e.g., via e-mail list-servs, Twitter, LinkedIn, etc.). Up to three e-mails were sent if they did not respond. When initial recruitment methods failed to provide sufficient respondents, we used Python to scrape emails of corresponding authors from an observational study corpus which examined endorsement of seven STROBE extensions [9,20,21]. To broaden the scope, we also included other journals primarily focused in Epidemiology. We identified 75 English language journals from the "Epidemiology" Broad Subject Term (BST) in the National Library of Medicine (NLM) [22], 122 endorsing journals from the STROBE Statement website [23], and 98 top-ranked journals in

the Scimago Journal and Country 2017 "Medicine" ranking [24]. We ran an Ovid MEDLINE observational study search filter from the same previous study [20] on all journals, deleted non-relevant publication types (e.g., case summaries, editorials), and restricted the search to English-language articles published within the past year (to reduce bounced emails). Supplementary File 3 details search strategies and journals searched. We de-duplicated e-mails and sent up to two emails to each author.

Statistical Analyses

General information on demographics, STROBE extension awareness, research stage usage, and awareness referral mechanisms are presented as counts and percentages in the aggregate and per subgroup. Likert scale responses are reported as means and standard deviations. Completion/dropout rates were calculated overall and per group based on completion of the final forced-response question.

As we used a modified UTAUT scale (Table 3), we had a priori assumptions about our model and its latent factors (Figure 1). Essentially, we were testing the health technology assessment (HTA) theory in our setting. Thus, Confirmatory Factor Analysis (CFA) was used to test a four-factor model of intention to use STROBE in the overall sample and subgroups (Figure 1). Rather than simply comparing average attitudes between groups, CFA allows us to test a theory and whether we captured relevant indicators and how they relate to each other (e.g., that we captured the key influencing factors that affect one's likelihood to use STROBE and furthermore, that we are comprehensive with our questioning and not redundant). All questions were scored from 1 to 7 and treated as continuous variables (Supplemental File 2). Three negatively worded questions from the Effort Expectancy scale were reverse coded prior to

calculating Cronbach's alpha and conducting CFA (see Table 4). For judging internal consistency, or the estimate of the reliability indicating the degree to which items measure different aspects of the same concept, we used Cronbach's alpha and considered ≧ 0.7 an acceptable value (28, 29).

All analyses were performed in R version 3.4.0. The R Markdown file, containing code and output, is available on Open Science Framework [27]. The model was fit using lavaan version 0.6-3 [28]. Maximum likelihood estimation (MLM specification) with robust standard errors was used to account for non-normality sample variance-covariance matrices and provide scaled test statistics. Latent factors were standardized, allowing for free estimation of all factor loadings. As suggested by Hu and Bentler [29], we considered Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values ≥0.90 for acceptable and ≥0.95 for good fit, root mean squared error of approximation (RMSEA) values ≤0.06 (poor fit > .10), and standardized root mean squared residual (SRMR) values ≤ 0.08 to indicate a good fit between the model and data. When conducting multiple-group CFA, convergence issues are common [30]. When they occurred, we investigated the model within subgroups to detect issues with modification indices, individual factor loadings, and covariances between latent factors.

Results

Of the 257 editorial offices contacted, 65 (25.3%) responded after three attempts. Of those who responded, 20 (30.8%) reported that they would invite their authors to participate (via Twitter, LinkedIn, listserv, blog, etc.), 42 (64.6%) declined to participate, and 3 (4.6%) reported individual-level participation. Reasons for declining included: no access to a list, no time, a desire to remain neutral, the inability to contact authors due to General Data Protection

Regulation restrictions [31], a belief that the journal did not publish observational studies (although we contend that it did), and a belief that the survey was flawed.

After accounting for 2,304 invalid addresses, 14,621 e-mails were sent to authors; we sent a second reminder e-mail to non-respondents. Over 100 authors (n = 109) informed us of participation, 23 declined (giving no reason or stating no time/ interest), and 23 reported ineligibility (i.e., did not work in observational research). Another 145 were ineligible as they were unreachable during the recruitment period (e.g., family, sick, sabbatical leaves) or unreachable permanently (e.g., left job, retired, died).

As the survey was anonymous and recruitment methods used network snowballing, social media, and mass mailings, we cannot estimate the total number of people that read the survey invitation. However, we know that 1293 visitors read the informed consent page and 1265 (97.8%) agreed to participate. After evaluating free-text responses, seven indicated ineligibility (e.g., *"I do not do observational research"*). Of the 1258 eligible participants, 1015 (80.7%) completed the survey. Nearly 20% (*n = 195*) indicated they never heard of STROBE nor used it prior to the study *(group 1)*, 18% (*n = 185*) had heard of it but never used it *(group 2)*, and over half (63%, *n = 635*) had heard of and used it *(group 3)* (Figure 3, Table 1). The completion rates were 67% for group 1, 81% for group 2, and 97% for group 3 (Figure 3).

We found fairly equal distributions for demographic categories across groups (Table 1, Supplemental File 4). The top five countries responding were: The United States (21.6%), United Kingdom (9.8%), Italy (6.8%), Canada (6.4%), and Australia (4.9%). To account for the multidisciplinary nature of research, we allowed up to three selections for area of work. Participants working in public health and epidemiology were well represented with 470 (46%) and 247 (24%) respectively choosing those options as one of their primary fields of work.

Table 1. Sample Demographics				
	Total Sample	**Never Heard of STROBE, Never Used** [Group 1]	**Heard of STROBE, Never Used** [Group 2]	**Heard of STROBE, Have Used** [Group 3]
	N *(%)*	N *(%)*	N *(%)*	N *(%)*
	1015 *(100)*	195 *(19)*	185 *(18)*	635 *(62)*
Time Spent in Research				
1 – 10 years	332 *(33)*	57 *(29)*	65 (*35)*	210 *(33)*
11 – 30	362 *(36)*	107 (*55)*	95 *(51)*	372 *(59)*
31 +	86 *(10)*	30 *(15)*	25 *(14)*	48 *(8)*
I do not work in research	3 *(<1)*	1 *(0)*	0 *(0)*	2 *(<1)*
Prefer not to say	3 *(<1)*	0 *(0)*	0 *(0)*	3 *(<1)*
Age				
18 – 34	185 *(18)*	36 *(19)*	38 *(21)*	111 *(1)*
35 – 54	589 *(58)*	101 *(52)*	83 *(45)*	405 *(64)*
55 +	235 *(23)*	58 *(30)*	64 *(35)*	113 *(18)*
Prefer not to say	6 *(<1)*	0 *(0)*	0 *(0)*	6 *(<1)*
Gender				
Woman	469 *(46)*	97 *(50)*	82 *(44)*	289 *(46)*
Man	525 *(52)*	94 *(48)*	101 *(55)*	329 *(52)*
Trans	3 *(<1)*	0 *(0)*	0 *(0)*	3 *(<1)*
Prefer not to say	20 *(2)*	4 *(2)*	2 *(1)*	14 *(2)*
Region				
Africa	22 *(2)*	5 *(3)*	2 *(1)*	15 *(2)*
Asiatic region	31 *(3)*	7 *(4)*	4 *(2)*	20 *(3)*
Eastern Europe	33 *(3)*	12 *(6)*	5 *(3)*	16 *(3)*
Latin America	54 *(5)*	14 *(7)*	10 *(5)*	30 *(5)*
Middle East	26 *(3)*	11 *(6)*	6 *(3)*	9 *(1)*
Northern America	283 *(28)*	58 *(30)*	57 *(31)*	168 *(27)*
Pacific Region	54 *(5)*	4 *(2)*	10 *(5)*	40 *(6)*
Western Europe	465 *(46)*	69 *(35)*	83 *(45)*	313 *(49)*
Not reported	47 *(5)*	15 *(8)*	8 *(4)*	24 *(4)*

Motivators of Use

When asked about what factors would or have motivated use of STROBE, the journal submission process and mandatory reporting guideline use were the most frequently chosen options. After journal policies, self-motivation was among the top-ranked reported influences. Around half (53%) of those who were aware of STROBE but had not used it (*group 2*) reported

that this was because journals did not require it. The next most frequently reported reason was that their co-authors did not use it (24.3%). (Table 2).

Usage Timing and Frequency

Participants who used STROBE (*group three*) most commonly did so during the manuscript writing process (*n = 451)* or after completing their draft to check that all relevant information had been reported (*n = 439*). Participants who had not used STROBE before (*groups one and two; n = 380*), most frequently reported that they would most likely use it during the manuscript writing process (*62.6%; 68.1%)* or during the protocol/study design stage (*64.6%; 59.5%*). For those who previously used STROBE (group 3), 48% used it for less than a quarter of their manuscripts, while 11.5% used it for all of their manuscripts. (Table 2)

Awareness

Of those who were aware of STROBE prior to the survey (groups 2 and 3, *n = 820*), the most frequently reported route that made them aware of STROBE was a journal requiring or recommending it [*group 2: n = 55, 29.7%; group 3: n = 234, 36.9%*]. The other options (peers, superiors, courses, or online) ranged from 12.6 to 16.7%. A majority of participants (70.7%, *n = 718* indicated that they were not aware of any STROBE extension.

Table 2. Motivators, Usage, and Awareness Descriptives

	Never Heard of STROBE, Never Used	**Heard of STROBE, Never Used**	**Heard of STROBE, Used**
	Group 1 *n = 195*	Group 2 *n = 185*	Group 3 *n = 635*

Motivator of Use *(Past/theoretical)* †*			
Self	128 *(66)*	55 *(30)*	308 *(49)*
Co-authors	57 *(29)*	72 *(39)*	116 *(18)*
Mentor/supervisor encouraged	40 *(21)*	-	-
Social norm	44 *(23)*	-	-
Journal submission process	104 *(53)*	134 *(72)*	376 *(59)*
Journal peer review	82 *(42)*	90 *(49)*	77 *(12)*
Incentivized in workplace	-	28 *(15)*	-
Immediate feedback	-	35 *(19)*	-
Free text	8 *(4)*	8 *(4)*	44 *(7)*
Reasons for Not Using†			
My writing wouldn't benefit	-	26 *(14)*	-
Do not want strict rules	-	25 *(14)*	-
Hard to understand	-	11 *(6)*	-
Word count	-	20 *(11)*	-
Format is difficult	-	16 *(9)*	-
Co-authors don't use it	-	45 *(24)*	-
May result in more negative review	-	1 *(<1)*	-
Journals don't require it	-	98 *(53)*	-
Not applicable for study type	-	25 *(14)*	-
Other	-	35 *(19)*	-
Frequency of Current Use			
Don't currently use	-	-	5 *(<1)*
Less than a quarter of manuscripts	-	-	305 *(48)*
Roughly half of manuscripts	-	-	134 *(21)*
Roughly 75% of manuscripts	-	-	118 *(19)*
All applicable manuscripts	-	-	73 *(12)*
Research Stage of Use *(Past/theoretical)* †			
Did not consider	22 *(11)*	10 *(5)*	NA
Protocol/design stage	126 *(65)*	110 *(60)*	239 *(38)*
Grant	76 *(39)*	61 *(33)*	89 *(14)*
Manuscript	122 *(63)*	126 *(68)*	451 *(71)*
After completing the article to check	98 *(50)*	80 *(43)*	439 *(69)*
Evaluating the article	70 *(36)*	69 *(37)*	243 *(38)*
Awareness Mechanism			
Peer/colleague	-	31 *(17)*	97 *(15)*
Boss/mentor/supervisor	-	16 *(9)*	90 *(14)*
Journal	-	55 *(30)*	234 *(37)*
Course/workshop	-	32 *(17)*	105 *(17)*
Online	-	37 *(20)*	66 *(10)*
Other	-	14 *(8)*	43 *(7)*

** Columns/items are blank as not all questions were presented to all branches*

† Question allowed for multiple responses to be selected

Confirmatory Factor Analysis on Modified UTAUT Scale

We attempted confirmatory factor analysis (CFA) on the overall sample but it would not converge. Therefore, we investigated the model within subgroups to identify convergence issues; it converged in all subgroups (Table 3). There were three recurring issues across groups: 1) The third Facilitating Conditions item (FC3) appeared to not belong to the FC scale; 2) The FC and EE covariance was very high (.88 - .91); and 3) Two pairs of items (EE4: EE5 and PE4:PE5) had significant shared variance, with the highest modification indices across all subgroups.

Model fit statistics and accompanying step-by-step descriptions are in Supplemental File 4, Table 6. Succinctly, the FC3 item phrasing was redundant with EE1. When FC3 was deleted, the model converged. This also reduced the high covariance between the FC and EE factors. The EE4 and EE5 items, along with EE6, were negatively worded so we allowed them to covary to account for method effects [32]. Items PE4 and PE5 were also allowed to covary as they were both related to academic publishing suggesting that they could covary for reasons other than the shared influence of the latent factor.

The 4-factor model addressing these issues was the best fit model for our data (Table 3). The CFI (.91) and TLI (.90) reached the "acceptable" cut point of .90. The SRMR (.07) was below its cutpoint of .08. The RMSEA (.08) was not less than .06 though. All factor loadings were statistically significant (all $ps \leq .001$) and salient (.437 to .909) (Table 5). The internal consistency reliability of all four subscales was strong (Cronbach's alpha ≥ 0.94 for all). Our items were parsimonious, functional, and internally consistent.

An overall pattern between groups was seen where those who had used STROBE before (*group 3*) had the highest scores, those who had never heard of STROBE before the survey (*group 1*) had second highest scores, and those who had heard of STROBE but never used it (*group 2*) most often had the lowest scores. (Additional File 4, Table 4; Figure 3).

Discussion

To the best of our knowledge, this is the first project to ask authors about their attitudes towards and experiences with STROBE, especially using health technology assessment (HTA) framework. Our project used a broad and multi-faceted sampling strategy which created a diverse sample of observational study authors. We also engaged nearly 200 participants who previously had never heard of STROBE, making our survey an awareness intervention itself.

The large sample enabled us to test a modified UTAUT scale on our entire sample and within subgroups. With an acceptable fit between our model and the data, we expect that this instrument may be useful for evaluating interactions with other reporting guidelines. Our results confirm the applicability of an HTA approach to RG, reveal important factors impacting STROBE use, and highlight a unique additional aspect of use, which may separate it from other pieces of technology – the academic publishing environment. Since our model-data fit was only "acceptable" and we needed to address shared error variance of two publishing-related items (PE4 and PE5), we believe that these two Performance Expectancy items might signal an unaccounted latent factor related to publishing.

Our CFA should be considered complementary to the descriptive results which emphasize the key role that journals have in raising awareness, motivating, and enforcing use. Journals were the most typical medium by which participants originally became aware of STROBE. Moreover, journals not requiring STROBE was the top reason why authors did not use it. We recommend that future work explores this concept more deeply. We suggest building upon the most parsimonious model (Table 3), not including FC3, accounting for method effects on the Effort Expectancy scale, and addressing the shared error variance of PE4 and PE5.

With regard to limitations, estimating a sample size was not tenable as there is no clearly defined participant pool. Also, we used mass mailings, thus, potentially, some e-mails were likely blocked by spam filters [33]. Additionally, we had differential dropout rates between groups which is expected as it is conceptually harder to think in theoretical terms (e.g., when would you consider using STROBE which you were just introduced to vs. when have you used it). The differences in participation rates between groups and non-response and self-selection biases also could have skewed our responses to be more positive towards STROBE. Additionally, the introduction to STROBE may have not been detailed enough and/or the participants may have not spent enough time on it.

Despite these limitations, overall, participants reported positive views towards STROBE, considering it useful, clear, and relatively easy to use. They also thought it would increase manuscript quality and the chances of getting published. However, they were not as positive regarding time requirements, reporting effects on productivity, speed and ease of writing. Our results should be reassuring to journal editors who fear losing authors to other journals with less strict requirements for publication [3]. These fears may be unfounded as participants indicated that, despite time costs, there are benefits of using an RG such as increasing the quality of their manuscripts and the chances of being published. Furthermore, they thought that the publishing environment (i.e., journals) would or do support its use.

Despite this perceived benefit of an increase in quality, we caution that, empirically speaking, the research in this regard is mixed. Recent work (2019) demonstrated that having a methodological reviewer dedicated to looking for missing reporting guideline items (not only STROBE) increased the number of citations that an article received by 43% [34]. This could be perceived as a proxy for higher quality or impact. Conversely, other authors have found no effect

on the reporting of confounding [35] or insufficient evidence to determine an impact on overall completeness of reporting [4,36]. To further assuage editors' (and authors') concerns, more research is needed in this area which focuses on a broad range of journals (i.e., not only high-impact) and which takes endorsement type (i.e., requiring vs. recommending use) into account.

A 2019 scoping review complements our results, highlighting the complexity of reporting guideline adherence and highlighting the need to implement interventions with different stakeholders throughout the research process [37]. Their review showed that most of the evaluated interventions to improve reporting guideline adherence have been conducted in journals. There have been mixed results but promising ones for more active implementation efforts (i.e., requiring a checklist with submission), including editorial assistants trained on reporting issues, and automatic peer review tools.

Widespread interventions are needed to improve RG adherence. Efforts to target research clusters, not just individuals, in order to foster broader support are needed. With increased uptake amongst co-authors completing reporting checklists, the time required may be reduced further, thus making using STROBE more appealing. When reporting guidelines become an expected part of the research process, self-regulation can occur and formal journal and institutional policies can be more fruitful as well [38]. Targeted and widespread promotion of reporting guidelines is needed to improve the completeness of reporting and reduce research waste [39].

Declarations

Ethics approval and consent to participate

Ethical approval was granted by the University of Split (2181-198-03-04-18-0010).

Authors' contributions

All authors have made substantive intellectual contributions to the development of the protocol and this manuscript. MKS conceptualized the study and led the writing of the manuscript. DH led the supervision of the manuscript preparation. MKS and LB managed survey recruitment. MKS, GG, and RR assisted with all analyses. MKS performed analyses. All authors read and approved the final manuscript.

Conflicts of Interest

MKS works with the STROBE Statement as a part of her doctoral studies. DH, GG, and EW provide support and mentoring as a part of the Methods in Research on Research (MiRoR) project. EW was a Fellow of the UK EQUATOR Centre which promotes the use of reporting guidelines (this was an unpaid position).

Source of Funding

This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie [grant agreement No 676207].
Coauthor GG was supported by work from Ministerio de Economía y Competitividad, Grant/AwardNumbers: MTM2015-64465-C2-1-R, MDM-2014-0445.

Availability of data and material

The final R Markdown code used for the current study will be made available on the Open Science Framework (https://osf.io/2fkny/) and in the Zenodo repository in the Methods in Research on Research (MiRoR) community (https://zenodo.org/communities/miror/).

Acknowledgements

The authors would like to acknowledge the Methods in Research on Research (MiRoR) network for their support and guidance. We would also like to thank Dr. Michail Kovanis for assistance gathering author emails and Ms. Alice Biggane for her assistance with the recruitment strategy.

Additional Files

Additional File 1. Completed CHERRIES Checklist
Additional File 2. Copy of Survey
Additional File 3. List of Journals Contacted and Searched
Additional File 4. Additional Tables and Analyses
Additional File 5. Recruitment Materials

References

[1] Gawande A. The checklist manifesto: How to get things right. Metropolitan Books New York; 2010.

[2] Costa BR da, Cevallos M, Altman DG, Rutjes AWS, Egger M. Uses and misuses of the STROBE statement: bibliographic study. BMJ Open 2011;1:e000048. doi:10.1136/bmjopen-2010-000048.

[3] Grindlay DJ, Dean RS, Christopher MM, Brennan ML. A survey of the awareness, knowledge, policies and views of veterinary journal Editors-in-Chief on reporting guidelines for publication of research. BMC Vet Res BMC Vet Res 2014;10, 10:10, 10–10. doi:10.1186/1746-6148-10-10, 10.1186/1746-6148-10-10.

[4] Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. BMJ 2011;343:d6783. doi:10.1136/bmj.d6783.

[5] Larson EL, Cortazal M. Publication guidelines need widespread adoption. J Clin Epidemiol 2012;65:239–46. doi:10.1016/j.jclinepi.2011.07.008.

[6] MacMahon B, Weiss NS. Is There a Dark Phase of This STROBE? Epidemiology 2007;18:791. doi:10.1097/EDE.0b013e3181570f6e.

[7] Rothman KJ, Poole C. Some Guidelines on Guidelines: They Should Come With Expiration Dates. Epidemiology 2007;18:794. doi:10.1097/EDE.0b013e3181571259.

[8] Hopewell S, Boutron I, Altman DG, Barbour G, Moher D, Montori V, et al. Impact of a web-based tool (WebCONSORT) to improve the reporting of randomised trials: results of a randomised controlled trial. BMC Med 2016;14:199. doi:10.1186/s12916-016-0736-x.

[9] Sharp MK, Tokalić R, Gómez G, Wager E, Altman DG, Hren D. A cross-sectional bibliometric study showed suboptimal journal endorsement rates of STROBE and its extensions. J Clin Epidemiol 2018;0. doi:10.1016/j.jclinepi.2018.11.006.

[10] Hua F, Walsh T, Glenny A-M, Worthington H. Surveys on Reporting Guideline Usage in Dental Journals. J Dent Res 2016;95:1207–13. doi:10.1177/0022034516657803.

[11] Kunath F, Grobe HR, Rücker G, Engehausen D, Antes G, Wullich B, et al. Do Journals Publishing in the Field of Urology Endorse Reporting Guidelines? A Survey of Author Instructions. Urol Int 2011;88:54–9. doi:10.1159/000332742.

[12] Meerpohl JJ, Wolff RF, Niemeyer CM, Antes G, Elm E von. Editorial Policies of Pediatric Journals: Survey of Instructions for Authors. Arch Pediatr Adolesc Med 2010;164:268–72. doi:10.1001/archpediatrics.2009.287.

[13] Toews I, Binder N, Wolff RF, Toprak G, Elm E von, Meerpohl JJ. Guidance in author instructions of hematology and oncology journals: A cross sectional and longitudinal study. PLOS ONE 2017;12:e0176489. doi:10.1371/journal.pone.0176489.

[14] Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004;6:e34. doi:10.2196/jmir.6.3.e34.

[15] Projet MiRoR | An innovative and ambitious joint doctoral training programme n.d. http://miror-ejd.eu/ (accessed July 16, 2018).

[16] Welcome to SurveyMonkey! n.d. https://www.surveymonkey.com/home/ (accessed March 17, 2017).

[17] Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Q 2003;27:425–78.

[18] Williams MD, Rana NP, Dwivedi YK. The unified theory of acceptance and use of technology (UTAUT): a literature review. J Enterp Inf Manag 2015;28:443–88. doi:10.1108/JEIM-09-2014-0088.
[19] Technology (U.S.) NIC on HSR& HC. HTA 101: Introduction to Health Technology Assessment n.d. https://www.nlm.nih.gov/nichsr/hta101/ta10103.html (accessed July 2, 2019).
[20] Sharp MK, Utrobičić A, Gómez G, Cobo E, Wager E, Hren D. The STROBE extensions: protocol for a qualitative assessment of content and a survey of endorsement. BMJ Open 2017;7:e019043. doi:10.1136/bmjopen-2017-019043.
[21] Sharp MK. Methods in Research on Research (MiRoR). STROBE Guidelines (ESR9). 2017. osf.io/u75gb.
[22] Broad Subject Terms for Indexed Journals n.d. https://wwwcf.nlm.nih.gov/serials/journals/index.cfm (accessed July 16, 2018).
[23] STROBE Statement: Endorsement n.d. https://www.strobe-statement.org/index.php?id=strobe-endorsement (accessed July 16, 2018).
[24] Journal Rankings on Agricultural and Biological Sciences n.d. https://www.scimagojr.com/journalrank.php?area=2700 (accessed July 16, 2018).
[25] Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951;16:297–334. doi:10.1007/BF02310555.
[26] Bland JM, Altman DG. Statistics notes: Cronbach's alpha. BMJ 1997;314:572. doi:10.1136/bmj.314.7080.572.
[27] Sharp MK. Use of reporting guidelines as an educational intervention for teaching research methods and writing. Open Sci Framew n.d. 10.17605/OSF.IO/HP623 (accessed November 16, 2017).
[28] Rosseel Y. lavaan: An R Package for Structural Equation Modeling. J Stat Softw Vol 1 Issue 2 2012 2012.
[29] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6:1–55. doi:10.1080/10705519909540118.
[30] Brown, Timothy A. Confirmatory Factor Analysis for Applied Research. Second Edition. New York, New York, USA: The Guilford Press; n.d.
[31] EU GDPR Information Portal. EU GDPR Portal n.d. http://eugdpr.org/eugdpr.org-1.html (accessed August 10, 2018).
[32] Maul A. Method Effects and the Meaning of Measurement. Front Psychol 2013;4. doi:10.3389/fpsyg.2013.00169.
[33] Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Comput Hum Behav 2010;26:132–9. doi:10.1016/j.chb.2009.10.015.
[34] Vilaró M, Cortés J, Selva-O'Callaghan A, Urrutia A, Ribera J-M, Cardellach F, et al. Adherence to reporting guidelines increases the number of citations: the argument for including a methodologist in the editorial process and peer-review. BMC Med Res Methodol 2019;19:112. doi:10.1186/s12874-019-0746-4.
[35] Pouwels KB, Widyakusuma NN, Groenwold RH, Hak E. Quality of reporting of confounding remained suboptimal after the STROBE guideline. J Clin Epidemiol 2016;69:217–24. doi:http://dx.doi.org.ezproxy.cul.columbia.edu/10.1016/j.jclinepi.2015.08.009.

[36] Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. BMJ 2014;348:g3804. doi:10.1136/bmj.g3804.
[37] Blanco D, Altman D, Moher D, Boutron I, Kirkham JJ, Cobo E. Scoping review on interventions to improve adherence to reporting guidelines in health research. BMJ Open 2019;9:e026589. doi:10.1136/bmjopen-2018-026589.
[38] Fuller T, Pearson M, Peters J, Anderson R. What Affects Authors' and Editors' Use of Reporting Guidelines? Findings from an Online Survey and Qualitative Interviews. PLOS ONE 2015;10:e0121585. doi:10.1371/journal.pone.0121585.
[39] Glasziou P, Altman DG, Bossuyt P, Boutron I, Clarke M, Julious S, et al. Reducing waste from incomplete or unusable reports of biomedical research. The Lancet 2014;383:267–76. doi:10.1016/S0140-6736(13)62228-X.

Table 3. Comparisons of Model Fit

Model	**Group (*n*)**	**x^2**	**df**	**RMSEA (90% CI)**[a]	**SRMR**[b]	**TLI**[c,d]	**CFI**[c,d]	**AIC**
4-Factor Model	Overall *(1015)*	--	--	--	--	--	--	--
	1 *(195)*	776.900	266	.109 (.100, .118)	.087	.818	.838	12593.900
	2 *(185)*	730.552	266	.108 (.099, .117)	.085	.797	.802	12305.731
	3 *(635)*	1582.699	266	.102 (.097, .107)	**.077** [b]	.813	.834	42959.805
FINAL MODEL: 4 Factor Model, No FC3, Method Effects	Overall *(1015)*	1931.539	717	.078 (.074, .082)	**.072** [b]	.895	**.909** [d]	64488.934
	1 *(195)*	489.527	239	.079 (.069, .089)	**.077** [b]	**.904** [d]	**.917** [d]	11915.867
	2 *(185)*	496.303	239	.084 (.074, .095)	**.075** [b]	.877	.894	11734.401
	3 *(635)*	927.172	239	.076 (.071, .081)	**.070** [b]	.897	**.911** [d]	40838.666

x^2: Chi-squared; df: degree of freedom; RMSEA: Root Mean Square Error of Approximation with 90% Confidence Intervals; SRMR: Square Root Mean Residual; TLI: Tucker-Lewis Index; CFI: Comparative Fit Index; AIC: Akaike Information Criterion

a) Within the range (≤ 0.06) indicating a good fit between the model and the data

b) Within the range ($\leq .08$) indicating a good fit between the model and the data

c) Within the range ($\geq .95$) indicating a good fit between the model and the data

d) Within the range ($.90 \leq x \leq .95$) indicating an acceptable fit between the model and the data

*Method Effects addresses the reverse-coded items EE4, EE5, and EE6 and the high covariance between PE4 and PE5

Table 4. Factor Loadings of Final Model *(n = 1015)*

Items		*Performance Expectancy (PE)*			*Effort Expectancy (EE)*			*Social Influence (SI)*			*Facilitating Conditions (FC)*		
		1[a]	*2*[b]	*3*[c]	*1*[a]	*2*[b]	*3*[c]	*1*[a]	*2*[b]	*3*[c]	*1*[a]	*2*[b]	*3*[c]
PE1	STROBE will be/is useful in my job	.728	.710	.665									
PE2	Using STROBE will enable/enables me to write papers more quickly	.868	.821	.818									
PE3	STROBE will increase/increases my productivity	.865	.817	.775									
PE4	If I use STROBE, I (will) increase my chances of getting published	.766	.699	.511									
PE5	If I use STROBE, I will get a more positive peer review of my paper	.737	.670	.553									
PE6	Using STROBE will make/makes it easier for me to write papers	.903	.861	.864									
PE7	Using STROBE will improve/improves the quality of my manuscripts	.803	.675	.767									
PE8	Using STROBE will make/makes my manuscript writing more efficient	.849	.867	.870									
PE9	Using STROBE increases the quality of my output for the same amount of effort	.850	.831	.804									
EE1	I think STROBE will be/is easy to use				.841	.771	.887						
EE2	I think STROBE's content is clear and understandable				.869	.833	.866						
EE3	I think that it will be/is easy for me to become skillful at using STROBE				.793	.797	.693						
EE4	Using STROBE will take/takes too much time compared to my normal writing process*				.437	.464	.604						
EE5	STROBE is so complicated, it will be/is difficult to understand what to do*				.622	.579	.671						
EE6	Will take/takes too long to learn how to properly use STROBE to make it worth the effort*				.569	.542	.598						
SI1	My peers will think/think that I should use STROBE							.848	.909	.870			
SI2	My superiors will think/think that I should use it							.621	.562	.639			
SI3	The research climate is helpful in promoting the use of reporting guidelines like STROBE							.887	.890	.831			
SI4	In general, I think that journals will support/support the use of STROBE							.649	.461	.473			
SI5	I will use STROBE because a lot of scientists in my field are using it							.531	.553	.549			
FC1	I have the knowledge necessary to use STROBE										.599	.490	.567
FC2	STROBE is compatible with my current workflow										.785	.777	.817
FC4	Using STROBE fits well with the way I like to work										.852	.878	.843

1[a] Subgroup One: Never heard of STROBE and never used it (*n = 195)*
2[b] Subgroup Two: Heard of STROBE but never used it (*n = 185)*
3[c] Subgroup Three: Heard of STROBE and have used it (*n = 635)*

Figure 1. Survey Flow

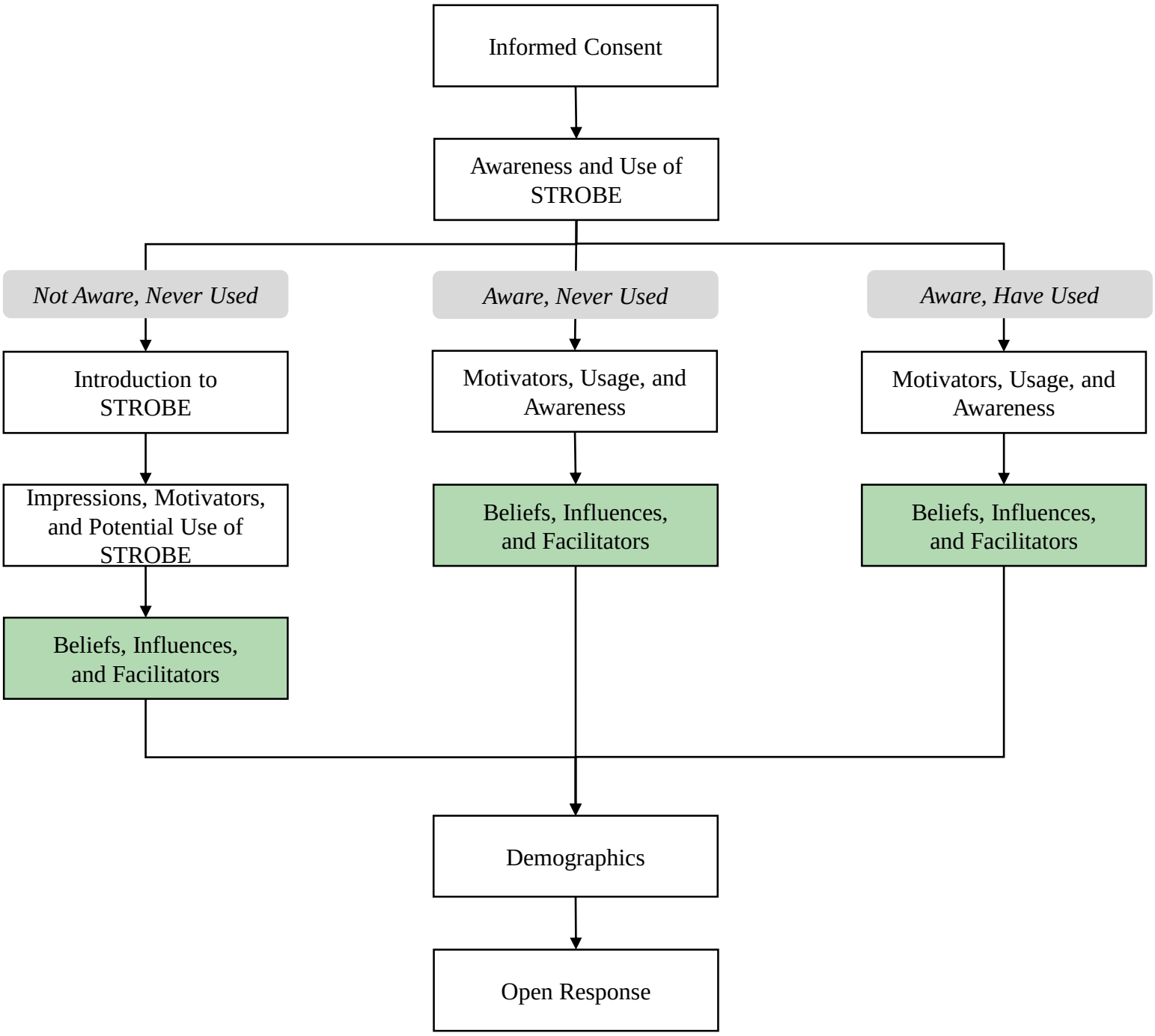
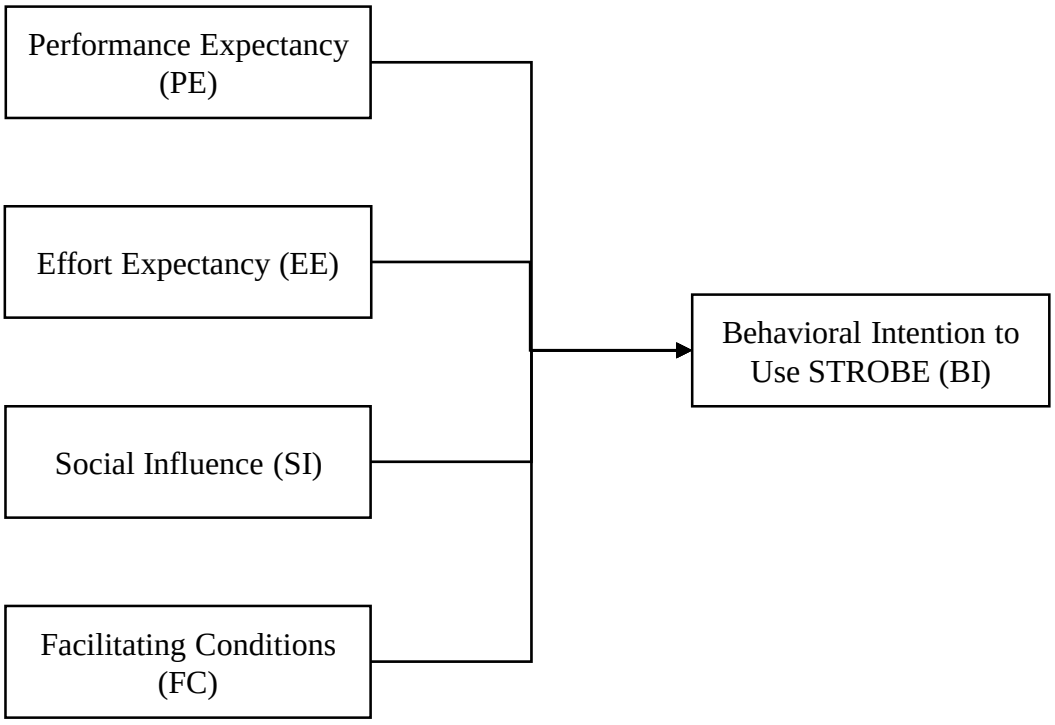


Figure 2. Model and Domain Definitions



Scale Domain Definitions [17]

Performance Expectancy (PE, 9 items): the degree to which an individual believes that using the system will help him or her to attain gains in job performance

Effort Expectancy (EE, 6 items): the degree of ease associated with use

Social Influence (SI, 5 items): the degree to which an individual perceives that important others believe he or she should use the new system

Facilitating Conditions (FC 4 items): the degree to which an individual believes that an organizational and technical infrastructure exists to support use of a system

Figure 3. Participant Flow Diagram