

Guideline for Reporting Standards of Eye-tracking Research in Decision Sciences

Susann Fiedler

Max Planck Institute for Research on Collective Goods

Michael Schulte-Mecklenbeck

University of Bern and Max Planck Institute for Human Development

Frank Renkewitz

University of Erfurt

Jacob L. Orquin

Aarhus University and Reykjavik University

Author Note

The authors thank Lina Falkenberg, Sarah Piechowski, Marlene Rösner, Marie Hellmann, Tabea Stegers, and Eveline Söder for invaluable assistance in collecting the data as well as Anita Lyubenova for great support in developing the Shiny App. Correspondence concerning this article should be addressed to Susann Fiedler, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, 53113 Bonn (Germany). E-mail: fiedler@coll.mpg.de

Abstract

We develop a reporting guideline for eye-tracking research in the behavioral sciences. To this end, we coded 215 articles on behavioral decision-making published between 2009 and 2017 and extracted a list of reported items. The coded articles were from a broad range of disciplines linked to judgement and decision making, such as cognitive science, marketing, economics, developmental research, vision research, and human–computer interaction. We then had a panel of eye-tracking experts rate the necessity of each item for reproducing a reported study. From these two sources, we generated a guideline containing 31 items that are judged as 'necessary' by the majority of experts for reproducing an eye-tracking study. None of the 215 coded articles report all identified items and approximately 70 percent of the articles report less than 50 percent of the 'necessary' items. We provide the data and list of recommendations as a hands-on shiny app to allow for an easy adoption of the proposed reporting guideline to improve transparency and reproducibility in eye-tracking research.

Keywords: reporting standards, reproducibility, eye-tracking, judgment and decision-making

Guideline for Reporting Standards of Eye-tracking Research in Decision Sciences

Reproducibility is an important element in ensuring that insights from scientific studies stand the test of time (Open Science Collaboration, 2015). A critical component of reproducibility is transparency in reporting how data were collected, aggregated and analyzed. Specifically, the method section should provide information on “how the study was conducted, including conceptual and operational definitions of the variables used in the study.” (VandenBos, 2010, p. 29).

In recent years, reproducibility has gained increased focus, and several calls have been made for more comprehensive and transparent reporting (Asendorpf et al., 2013; Eich, 2014; Nosek et al., 2015; Spellman, 2015; Vazire, 2015). One way to improve the desired transparency and reproducibility is to standardize reporting guidelines. Such guidelines typically contain a list of items that inform authors about what technical and analytical details should be reported (e.g., JARS, 2008). While some reporting guidelines are general, for instance, the number of participants, each discipline and area is likely to require more specific reporting. As a consequence, several specialized guidelines have been developed over the years, for instance, reporting of fMRI and MRI studies (Poldrack et al., 2008, Nichols, Das, Eickhoff, Evans et al., 2017), experiments conducted on the internet (Reips, 2002), computational studies (Stodden et al., 2016) and meta-analyses (Moher, Liberati, Tetzlaff, & Altman, 2009).

Despite this diversity in the guidelines, many disciplines are still lacking such specialized reporting standards. One of these areas is eye-tracking research. In recent years, the application of the eye-tracking methodology has grown rapidly in the behavioral sciences. Lower costs of eye-trackers, easier handling and a broader interest in process models in general have all contributed to this development (Schulte-Mecklenbeck et al., 2017). Nevertheless, eye-tracking requires complex apparatus and provides many degrees of freedom in terms of experimental and stimulus design, technical configuration, data pre-processing, and analysis (Orquin & Holmqvist, 2017). With a growing community of researchers and many degrees of freedom in running eye-tracking studies, we believe that eye-tracking is in particular need of specialized reporting standards.

Reporting guidelines tend to be constructed through discussions among small groups of experts, and the criterion for inclusion or exclusion of items is agreement in discussion (e.g.,

JARS, 2008). While expert discussions may go a long way in determining important reporting issues, we aimed for an additional step to allow us to document the identification and in- or exclusion of each reported item. This approach has two distinct advantages. 1) The selection process is better documented; we are able to provide a full list of items that are of interest to eye-tracking researchers (i.e., that have been reported in published papers), and that we considered for inclusion in the final reporting guideline. The informed reader may therefore decide for herself whether to report otherwise excluded items. 2) A systematic approach to developing reporting guidelines can easily be adopted as a blueprint by other disciplines in need of specialized guidelines.

We approached the development in two steps. First, we developed an inclusive list of reporting items by coding a large set of eye-tracking papers in behavioral decision-making published between 2007 and 2017. The list was enriched with items from the literature on eye-tracking methodology. Second, this comprehensive list of items was evaluated by an expert panel who judged the necessity of each item for reproducibility. These steps result in (1) an extensive list including all items identified through our literature coding procedure. This list serves as a reference for conducting eye-tracking studies. It highlights the large set of decisions that have to be made during the design and data collection stage of an eye-tracking experiment. (2) The output of the expert panel is a condensed reporting guideline which includes only those items deemed necessary by a majority of experts. The reporting guideline is easier to apply than the complete list and we hope this will promote its adoption and thereby enhance the reproducibility and transparency of eye-tracking research. To ensure that the basic data are available to the interested reader we programmed a Shiny-App (<https://decisionlab.shinyapps.io/iGuidelines/>) that allows access to all the included citations and codings we report here.

Method

The reporting guideline was developed in five steps shown in Figure 1 below. The main idea of this approach is to take advantage of the large body of literature to identify reporting items that might have been overlooked by individual experts while also allowing for the possibility that important items may have been ignored in the literature.

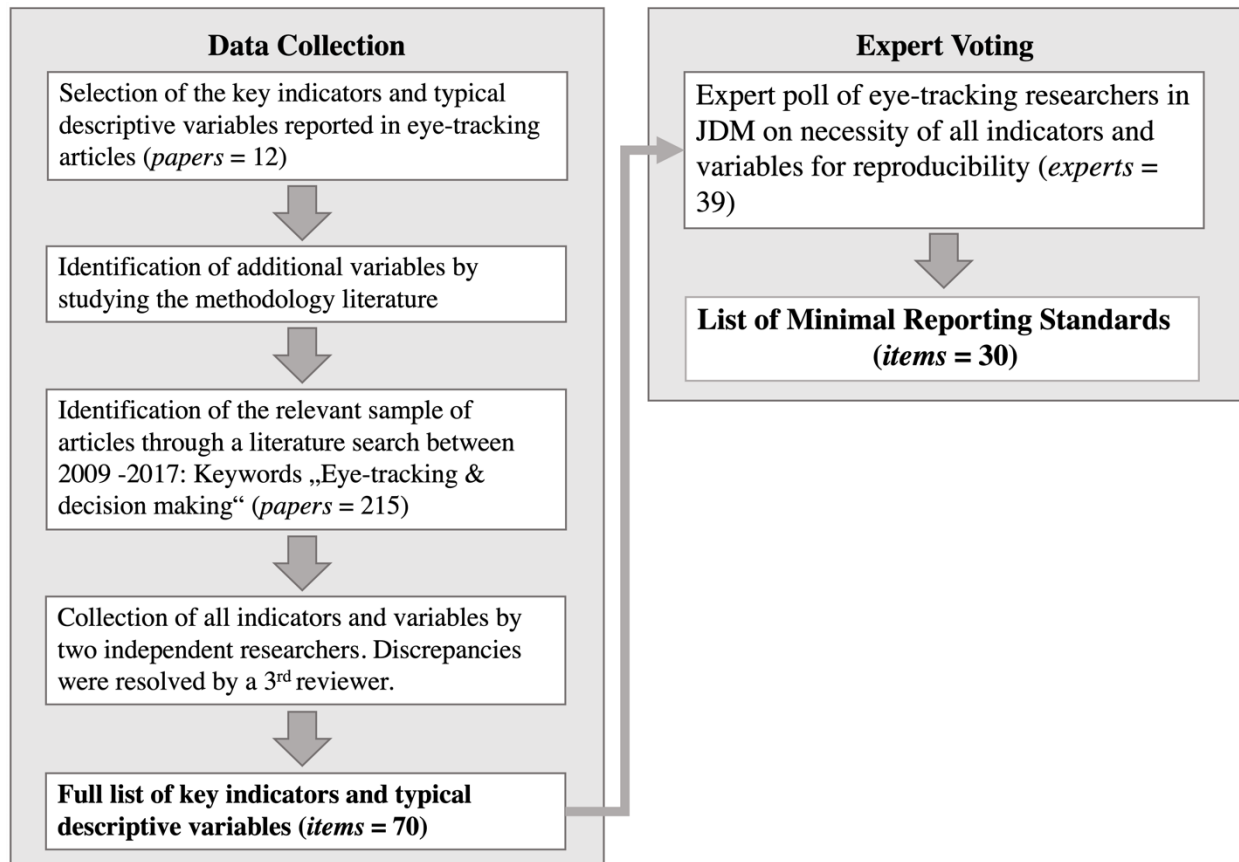


Figure 1. Overview of the steps in the development of the reporting guideline

Developing the coding scheme. Based on an initial set of 12 eye-tracking articles, a list of criteria was developed and revised in four rounds through discussions between all four authors. This initial list was enriched by an additional set of items identified in the eye-tracking methodology literature. The full list contains 70 items related to eye-tracking methodology or analysis (see Table A1 in the Appendix).

Literature search. To narrow down the available literature, we limited ourselves to eye-tracking studies on behavioral decision-making. Through a literature search in the Web of Knowledge and Google Scholar databases, we identified 215 articles including a total of 268 empirical studies. Search keywords were combinations of 'eye-tracking', 'decision-making', 'choice' and 'judgment'. We identified additional articles by searching the reference lists of the respective articles. Only experimental research published between 2009 and 2017 was included.

The articles identified covered a broad range of disciplines, such as cognitive science, marketing, economics, developmental research, vision research, and human–computer interaction.

Coding the literature. Each of the 215 articles included were coded on the 70 items by at least two independent coders. Disagreements between coders were resolved by a third coder. All four authors and three research assistants contributed to coding the articles. To train the coders and ensure coder reliability, an initial set of articles were coded and discussed by the entire team. Codes were entered into an online form to limit errors.

Expert panel and final guideline. The 70 items were evaluated by 39 eye-tracking experts, who indicated the importance of each item for the evaluation and reproducibility of eye-tracking studies. The response options were 'necessary', 'helpful', 'not helpful', and 'no opinion'. We identified an expert as a researcher who had published at least one scientific article about eye-tracking in a peer-reviewed journal (a criterion that excluded five participants in the survey). Respondents in the final sample had published on average five articles on eye-tracking ($M = 4.7$, $SD = 5.4$, range = 1:20). The final reporting guideline consists of those items rated as necessary by the majority of experts (see Table 1).

Results

Summary statistics for the coding of all included articles is provided in Table A1, that shows the percentages of articles reporting an item from the comprehensive item list and, additionally, the percentage of experts judging the items as necessary for the minimal reporting guideline. Based on the 30 items in the minimum reporting guideline we coded all 215 papers on whether an item was reported or not. The results of this coding are illustrated in Figure 2. None of the articles coded reported enough information for reproducibility based on the minimal requirements we suggest. Approximately 70 percent of all papers reported less than 50 percent of the necessary items. Interestingly, there is little agreement between what is reported and what experts deem important. The correlation between the proportion of articles reporting an item and the proportion of experts indicating the item as necessary is $r = -.02$, $p = .89$.

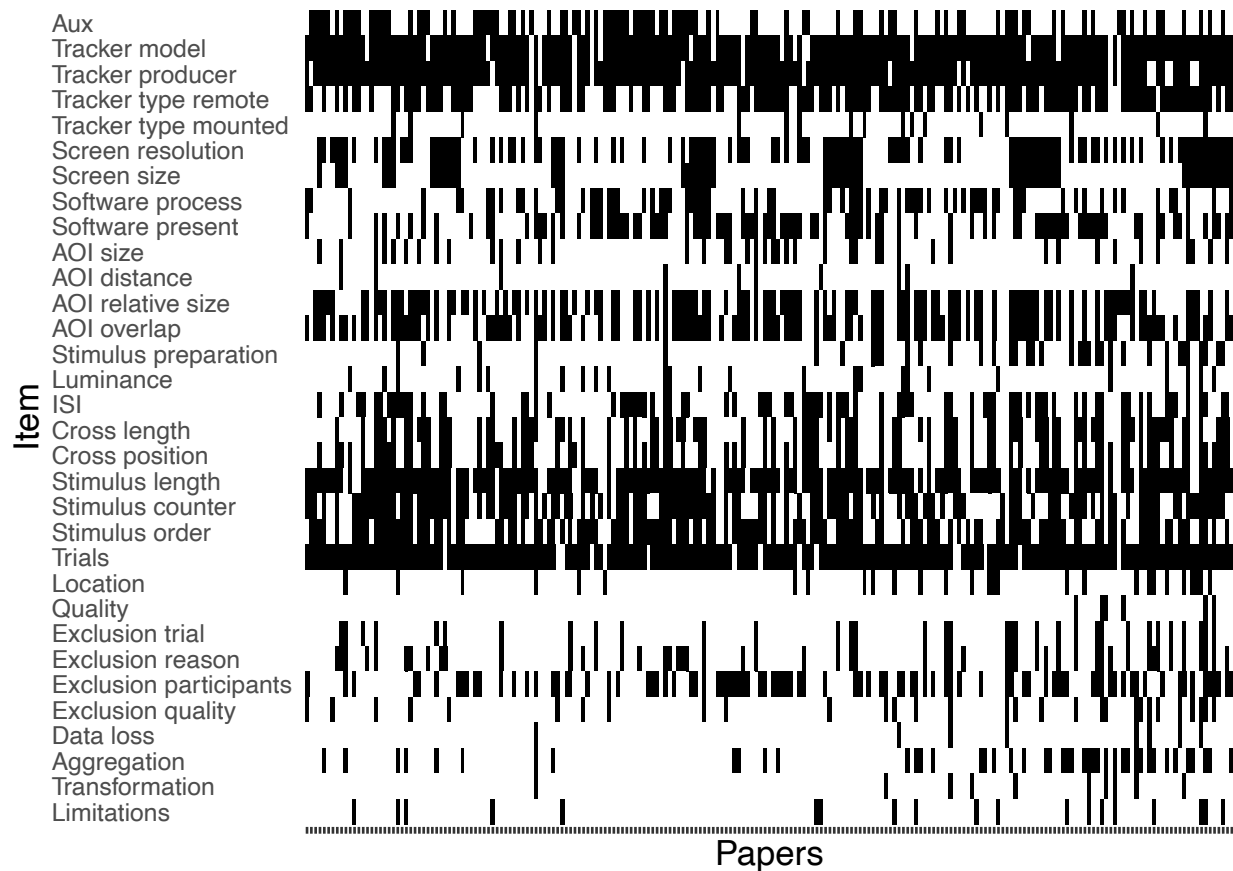


Figure 2. Overview of reporting practices in eye-tracking studies. Each paper is represented by a column (tick on the x-axis), a black rectangle denotes that the item given on the y-axis was reported. The list of items corresponds to the Minimal Guideline presented below, these items are explained in more detail below.

Reporting guideline

Despite this rather sobering picture in terms of what actually gets reported we believe that introducing a standard for reporting might increase these numbers substantially. Our suggested reporting guideline is shown in Table 1 below. In what follows we explain all the included items in more detail and discuss reasons why the items are essential for reproducibility.

Table 1. *The reporting guideline for eye-tracking research in behavioral sciences.*

Introduction

State auxiliary assumptions about the underlying processes of dependent variables

Methods

Apparatus

Description of the eye-tracking device

Model (e.g., Tobii 1000)

Producer/ brand

Type (remote, head mounted)

Description of the monitor

Resolution

Size

Description of the Software

Software used to pre-process the eye-tracking data

Stimulus presentation software

Material

Description of how AOIs were defined

Absolute size of the AOIs

Relative size of AOIs and content within the AOIs

Minimal distance between AOIs

Overlap between the AOIs

Description of the stimulus

Method for stimulus preparation

Luminescence matched

Procedure

Setup

Inter stimulus interval

Length of fixation cross presentation

Position of the fixation cross

Length of stimulus presentation

Counter balancing of the position

Order of stimulus presentation

Number of trials

Settings and locations where data was collected

Results

Data quality

Monitoring of data quality during experiment

Proportion of trials excluded for the analysis

Reasons for exclusion

Number of participants excluded from the analysis

Quality threshold for data exclusion

Percentage of lost data

Dependent measures

Aggregation method for fixations

Additional transformation of the data

Discussion

Limitations due to the use of eye-tracking methodology

Introduction (auxiliary assumptions)

The recording of eye-movements objectively measure the spatial orientation of the eye and pupil size at a given point in time. Observable gaze behavior is often interpreted as an indicator of cognitive processes; in many behavioral experiments, the eye movement is not the objective of the study so much as the observable output of an underlying psychological process. When making inferences about unobserved cognitive processes based on eye movements, we typically have to make assumptions about the relation between observed and unobserved variables. These auxiliary assumptions may be more or less reasonable. For instance, researchers often wish to assume a strong relationship between the eye and the mind, the so-called eye-mind assumption (Just & Carpenter, 1980) which states that stimuli that what is currently fixated on is being processed cognitively. However, this assumption has been falsified in various instances (Huettig, Rommers & Meyer, 2011, Orquin & Holmqvist, 2017). Therefore, the *auxiliary assumptions* linking the dependent measurement with the construct under investigation must be reported. By explicitly stating auxiliary assumptions, other researchers may decide for themselves whether a dependent measure is construct valid and whether to trust the conclusions in the article.

Describe the Devices and Software

Eye-tracker. One of the most constitutive aspects of eye-tracking data is the eye-tracker. The abundance of sampling techniques, sampling rates (e.g., 1000 vs. 30 Hz) and software has resulted in very diverse eye-tracking devices in terms of temporal and spatial precision and accuracy. Stating the *model*, *producer*, and *type* of eye-tracker is therefore no trivial matter as it will allow other researchers to consult documentation about the hardware. However, it is important not to rely on the accessibility of online documentation from hardware producers as these may cease to exist and the websites with documentation may be deleted.

Monitor. Usually eye-tracking devices can be linked up to a wide set of monitors and projectors. As *screen sizes* and *resolution* may vary, their measures are required to make sense of the stimulus size in degrees of visual angle and the distance between specific areas of interest (AOIs). To enable reproducibility, measures on screen size and resolution must be available. This will allow other researchers to mimic the original test and to obtain data on, for instance, saccades of similar length and AOIs of similar size. If a study is re-done with equipment of a different size or resolution, the outcome will most likely be imprecise stimuli and a different information search process.

Software. Eye-tracking studies can be organized using a variety of experimental software, much of which is offered by the equipment manufacturers (e.g., Experiment Center). Lately, however, open source alternatives, featuring interfaces for many eye-tracking devices, have emerged (e.g., Psychopy). All of this presentation software has its specific characteristics, such as randomization and latencies, which may affect the data. When collecting data, specifically when generating individual fixations or pupil dilations from raw gaze data, it takes quite some pre-processing of the data before the data set is usable. The various procedures of filtering, aggregation and event detection can be carried through using a software add-on with the eye-tracking device. These software packages, or if the researchers' own pre-processing programs are used, will determine the quality of the data set to be analyzed subsequently. The settings for pre-processing programs (e.g., what is the maximum dispersion of a fixation, how is the pupil size being normalized) differ across eye-tracking providers and program versions. If individually scripted software is applied in a study, the source code should be published because the original software is the key to completely reproducing an experiment and any pre-processing of the data.

Material

Areas of Interest. Most eye-tracking research relies on areas of interest (AOI) for data aggregation and analysis. AOIs are defined as an area around an object of interest, and all fixations falling inside the area are assigned to the object of interest. Research has shown that how the AOI is specified both in terms of space and place may affect the data set and the outcome (Orquin, Ashby, Clarke, 2016). As a result, the *absolute* and *relative size of the AOI*, the *distance between the AOIs* and their potential *overlap* must be stated. Furthermore, even the slightest details such as the interaction between the individual eye-tracker characteristics (i.e., accuracy and precision) and the size of the AOI are so sensitive that the data set and results may change.

Stimuli. Eye-tracking studies and the visual stimuli used are interlinked. Therefore the stimuli used as well as the *stimulus preparation method* must be specified. Also, whether stimuli are *luminance matched* must be reported for pupillometric eye-tracking research. Luminance matching reduces the volatility of the pupil size due to changes in light conditions and is of particular importance when we want to measure pupil dilation. If differences in light sources can be ruled out, it is easier to interpret pupil data and it also reduces noise when collecting eye position measures.

Procedure. To give a full account of how to conduct an eye-tracking study, a number of questions concerning the length and form of the stimuli presentation as well as the experiment setup must be answered. All these elements have an impact on the validity of the study and its conclusions. The *length of the inter-stimulus interval* (the time between the stimulus presentations) should be stated as this piece of information is crucial in terms of resting phases and potential baseline measures, because it allows assessing the relative differences in pupil size between when the eye is stimulated and when it is at a baseline. Since pupil contraction has some latency, it needs time to return to its baseline (approximately 4 s). A fixation cross before the stimulus presentation serves to focus or divert attention on or away from a particular position on the screen. Information on its *position* (e.g., at the middle of the screen) and *presentation length* therefore improves the evaluation of the meaning of the first fixation. Another equally important item in the interpretation of the data is the *length of the stimulus presentation*; for example, a fixed exposure time of 2000 ms vs free exposure time permitting participants to decide when to

continue, calls for radically different interpretations of the same analysis. With free exposure time, the number and duration of fixations can easily be compared between AOIs or participants, while the fixed exposure time makes it easier to compare when AOIs are fixated. In free exposure, trials are typically of different lengths, and comparing when AOIs are fixated raises the question of timing relative to the beginning or end of the trial or the proportion of trial length.

When multiple AOIs are involved, there is the issue of *counterbalancing* – that is, do the objects of interest switch position (e.g., between the top and bottom or left and right side of the screen). If counterbalancing is neglected, the interpretation of the eye movement distribution may be distorted as the eye movements under question may be confounded with the natural reading direction (Orquin & Holmqvist, 2017). Temporal *ordering* of the stimuli (e.g., random or fixed presentation order) may also be an issue as are the *number of trials*. At the outset of eye tracking studies, decision times tend to be long and more information tends to be scrutinize in comparison to later on in the study (e.g., Fiedler, Glöckner, Nicklisch & Dickert, 2013). Orquin, Bagger and Mueller Loose (2013) used this observation to demonstrate that there is an increase in top-down control of eye movements over time. Particularly for within subject designs, it is relevant to know the type of stimuli and their temporal ordering. For the purpose of understanding the surroundings of the data collection and how it was controlled, all details about the location used must be described (e.g., in a controlled lab environment versus a public area, etc.).

Data Pre-processing

Dependent measures. Before the data analysis most researchers pre-process their data with one or more data aggregations and transformation steps, for instance, *aggregation of fixations* (i.e., defining the minimum fixation length and maximum dispersion) or *pupil size* (i.e., algorithm used to calculate pupil size e.g., as a difference measure to baseline or as a relative change in size) from raw data. Eye-tracking equipment providers often supply pre-processing programs that can handle these aggregations. Data characteristics must be reported as the provider-set defaults are not standardized, and decisions should be made in accordance with the stimuli used (Manor & Gordon, 2003). Besides the standard algorithms, many analyses are based on within-AOI fixation aggregations or other transformations such as computation of transition indices or metrics. The computation of these metrics will, of course, affect subsequent analyses;

sometimes the correct computation can be disputed, for instance, in terms of transition index development (Böckenholt & Hynan, 1994).

Data Quality. To understand the conclusions drawn from an experiment, it is important to be able to assess the quality of the data, and further the quality of the results (Orquin & Holmqvist, 2017). Data quality in eye-tracking studies can refer to several, often general measures in eye-tracking studies: the *proportion of participants and trials excluded* from the analysis and the *reasons for these exclusions*, the percentage of data samples for which the eye-tracker could not obtain the gaze position, also referred to as *lost data*, and whether *data quality was monitored* during the experiment.

Discussion - Limitations

In eye-tracking studies, there are certain *limitations* due to the dependent measures and the construct of interest being separated. For example, a fixation cannot always be interpreted as information processing, and the absence of a fixation cannot necessarily be interpreted as the absence of processing as eye-trackers do not measure the use of peripheral vision. Stimulus design, sample and data quality issues must also be reported and discussed to provide a transparency in the warranted and unwarranted conclusions. Conveying these details to readers allow them to draw their own conclusions with regard to the validity of the results and their interpretation.

Discussion

Transparent reporting of scientific research is, without a doubt, one of the stepping stones to replicability and scientific progress. However, it can be difficult to know what information is necessary to report for others to assess and replicate our research. Several guidelines have been proposed to encourage more complete and transparent reporting. These guidelines range from the general, such as the APA reporting guideline that encompasses most psychology articles, to the more specific, such as reporting guidelines for meta-analyses, online studies, or computational studies (Field et al., 2008; Moher, Liberati, Tetzlaff, Altman, & Group, 2009; Poldrack et al., 2008; Reips, 2002; Taylor et al., 2008; Tooth, Ware, Bain, Purdie, & Dobson, 2005).

Here we focus on eye-tracking research, an area undergoing explosive growth, and for which no specific reporting guidelines exist. We approach the development of the guideline in a systematic way. By coding eye-tracking papers in behavioral decision-making from 2007 to

2017, we developed a list of reporting items; we then let a panel of experts judge the necessity of each item for reproducibility. We include all items in the guideline judged as necessary by the majority of experts and present both the complete list of items and the shorter reporting guideline. The complete list is useful when conducting eye-tracking studies in that it helps uncover many of the methodological decisions eye-tracking researchers have to make. The shorter reporting guideline is mainly intended as an assistance for documenting the study in scientific journal articles.

The guideline is not intended as an inflexible requirement for publication but rather as a dynamic list that may change over time as the hardware, software and methodologies evolve. We address the dynamic character of the field in providing a web-based platform that is open to explore the collected data in more detail¹. We do wish to emphasize that the guideline cannot replace critical reflection. A common approach to methods and reporting is to copy previous research, which unfortunately can lead to propagation of errors, such as the exclusion of fixations shorter than 300 msec (Orquin & Holmqvist, 2017). A better approach is to reflect on the particulars of each study and make decisions with regard to methods and reporting that are justified by logical and scientific arguments, not by what previous researchers have done. In relation to this guideline, we therefore hope that it will not replace critical reflection but rather serve as a tool for better methodological reasoning.

¹ <https://decisionlab.shinyapps.io/iGuidelines/>.

References

- Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., ... & Perugini, M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality, 27*(2), 108-119.
- Böckenholt, U., & Hynan, L. S. (1994). Caveats on a process-tracing measure and a remedy. *Journal of Behavioral Decision Making, 7*(2), 103-117.
- Eich, E. (2014). Business not as usual. *Psychological Science, 25*(1), 3-6.
- Fiedler, S., Glöckner, A., Nicklisch, A., & Dickert, S. (2013). Social value orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational behavior and human decision processes, 120*(2), 272-284.
- Field, D., Garrity, G., Gray, T., Morrison, N., Selengut, J., Sterk, P., ... & Ashburner, M. (2008). The minimum information about a genome sequence (MIGS) specification. *Nature Biotechnology, 26*(5), 541-547.
- Huetting, F., Rommers, J., & Meyer, A. S. (2011). Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta psychologica, 137*(2), 151-171.
- JARS. (2008). Reporting standards for research in psychology: why do we need them? What might they be? *The American Psychologist, 63*(9), 839-851.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine, 151*(4), 264-269.
- Nichols, T. E., Das, S., Eickhoff, S. B., Evans, A. C., Glatard, T., Hanke, M., ... & Proal, E. (2017). Best practices in data analysis and sharing in neuroimaging using MRI. *Nature Neuroscience, 20*(3), 299-303.
- Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., ... & Contestabile, M. (2015). Promoting an open research culture. *Science, 348*(6242), 1422-1425.
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.
- Orquin, J. L., Ashby, N. J., & Clarke, A. D. (2016). Areas of interest as a signal detection problem in behavioral eye-tracking research. *Journal of Behavioral Decision Making, 29*(2-3), 103-115.

- Orquin, J. L., Bagger, M.P. & Mueller Loose, S. (2013). Learning affects top down and bottom up modulation of eye movements in decision making. *Judgment and Decision Making*, 8(6), 700-716.
- Orquin, J. L., & Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behavior research methods*, 50(4), 1645-1656.
- Poldrack, R. A., Fletcher, P. C., Henson, R. N., Worsley, K. J., Brett, M., & Nichols, T. E. (2008). Guidelines for reporting an fMRI study. *Neuroimage*, 40(2), 409-414.
- Reips, U. D. (2002). Standards for Internet-based experimenting. *Experimental Psychology*, 49(4), 243-256.
- Schulte-Mecklenbeck, M., Johnson, J. G., Böckenholt, U., Goldstein, D. G., Russo, J. E., Sullivan, N. J., & Willemsen, M. C. (2017). Process-tracing methods in decision making: On growing up in the 70s. *Current Directions in Psychological Science*, 26(5), 442-450.
- Stodden, V., McNutt, M., Bailey, D. H., Deelman, E., Gil, Y., Hanson, B., ... & Taufer, M. (2016). Enhancing reproducibility for computational methods. *Science*, 354(6317), 1240-1241.
- Taylor, C. F., Field, D., Sansone, S. A., Aerts, J., Apweiler, R., Ashburner, M., ... & Brazma, A. (2008). Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nature Biotechnology*, 26(8), 889-896.
- Tooth, L., Ware, R., Bain, C., Purdie, D. M., & Dobson, A. (2005). Quality of reporting of observational longitudinal research. *American Journal of Epidemiology*, 161(3), 280-288.
- VandenBos, G. R. (Ed). (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- Vazire, S. (2016). Editorial. *Social Psychological and Personality Science*, 7 (1), 3-7.

Appendix

Table A1. *Comprehensive list of reporting items.*

#	Variable	Definition	Literature		
			Example	% Reported	No Opinion
Introduction					
1	Importance of auxiliary assumptions	The importance of auxiliary assumptions linking constructs and operational definitions	Assuming that fixations imply cognitive processing	46	0
<hr/>					
Apparatus					
<i>Eye-tracking Device</i>					
2	Name	The eye-tracker used	EyeLink 1000	92	0
3	Producer	The producer of the eye-tracker	Tobii	91	0
4	Type remote Device	The type of eye-tracking device, i.e. remote or head-mounted	Remote	65	0
5	Type of Device	The type of eye-tracking device, i.e. corneal reflection, dual purkinje etc.	Corneal reflection	7	20
6	Lens size	The lens size of the eye-tracker	45mm	<1	30
7	Sampling Procedure	The sampling procedure used	Binocular	47	0
8	Sampling Rate	The sampling rate used	60 Hz	77	0
9	Accuracy	The accuracy of the eye-tracker	0.4°	33	0
10	Precision	The precision of the eye-tracker	0.24°	4	10
11	Temporal Precision	The temporal precision, i.e. the standard deviation of eye-tracker latency in ms	50 ms	1	10
12	Stimulus-Synchronization Latency	Specification of stimulus-synchronization latencies (delay between actual onset of a new stimulus and recorded onset of the stimulus in ms)	100 ms	<1	20
13	Latency	Eye-tracker latency (end-to-end delay from the occurrence of an eye movement to the signal from the recording computer that this movement has taken place)	< 11 ms	<1	10
14	Headbox	Tracking range of the head box in which participants can move without losing data	50 cm x 40 cm	1	30
15	Chin Rest	If a chinrest was used	Chinrest was used	44	10
<i>Monitor</i>					
16	Type	The type of monitor used	LCD monitor	32	0
17	Resolution	The resolution of the used monitor	1280 x 1024	56	0
18	Screen Size	The screen size of the used monitor	19"	30	0
19	Screen Refresh Rate	Screen refresh rate of the used monitor	60 fps	12	0
<i>Software</i>					

Running head: REPORTING STANDARDS

20	Pre-Processing Program	The software used to pre-process the eye-tracking data	Tobii Studio	34	10
21	Stimulus Presentation	The software used to present the stimuli	Matlab	46	10
<hr/>					
Material					
<i>AOI definition</i>					
22	Size of AOIs	The size of the AOIs in pixel or degrees	100 x 100 pixel	24	10
23	Overlap between the AOIs	Overlap between the AOIs	No overlap	66	10
24	Minimal Distance between AOIs	The minimal distance between AOIs in pixel	0°	6	10
25	Relative Size of AOI and Content within the AOIs	The relative size of AOIs and Content within AOIs	Bigger than stimulus	63	10
<i>Stimulus</i>					
26	Example Image Included	Example image presented in the paper	Yes	76	0
27	Method for Stimulus Preparation	Method for stimulus preparation	e.g. computer generated based on saliency scores	16	10
28	Luminance Matched between Stimuli	Matching of the luminance between the stimuli	Yes	14	20
29	Size of Stimulus	The size of the stimulus (e.g., the picture, the font)	Font size 23	33	0
<hr/>					
Procedure					
<i>Setup</i>					
30	Length of Inter Stimulus Interval	Duration of inter stimulus interval	Self-paced by the participant	64	0
31	Length of Fixation Cross Presentation	Presentation duration of the fixation cross (in ms)	500 ms	66	0
32	Position of Fixation Cross	Position of the fixation cross	Center of screen	62	0
33	Length of Stimulus Presentation	Duration of stimulus presentation (in ms)	Self-paced by the participant	84	10
34	Order of Stimulus Presentation	The order of the stimulus presentation	Random	71	0
35	Counter Balancing of Positions in the Presentation	Counter-balancing of the stimulus in the presentation across positions	No counter balancing	72	10
36	Number of trials	The number of trials in the experiment	One trial	95	0
37	Experimenter	Description of the person running the experiment	Experiment was run by lead author	<1	10
38	Setting and Location of Data Collection	Settings and locations where data were collected	Data were collected in a soundproof room with no direct daylight	12	0

Running head: REPORTING STANDARDS

39	Distance between Participant and Screen	The distance between participants and the screen	60 cm	59	0
Calibration					
40	Number of Calibration Points	Number of points that appeared in calibration	9 calibration points	43	10
41	Background Color of the Calibration	The background color of the calibration	Black	2	10
42	Amount and Time of Recalibration	Time, when the calibration was conducted.	At the beginning	74	0
43	Specification of Calibration Procedure	Specification of the calibration procedure	Automated and system approved calibration	4	10
Participants					
44	Vision of Participants	The vision of the participants	Normal or corrected to normal vision	44	10
45	Proportion of Women	The proportion of women	50% women	76	0
46	Mean Age	The mean age of participants	Mean age = 22	69	0
47	Procedure to test Vision	Procedure for testing visual acuity or color vision (e.g., participants were asked or tested before experiment)	Ishihara Color Vision Test	3	0
48	Color vision	Color vision (e.g., does the participant have full color vision)	Participants were only included if having correct color vision	4	10
<hr/>					
Data Processing					
Data quality					
49	Handling of Artefacts	Procedure for handling of participant artefacts (e.g., drooping eyelids, mascara etc.)	Participants were asked to remove their mascara	6	20
50	Obtained Accuracy	The obtained accuracy of the data (in visual degrees)	0.5°	3	20
51	Monitoring of Data Quality	Monitoring of data quality during experiment	Data quality was monitored constantly by the experimenter	3	0
52	Proportion of Excluded Trials	The proportion of trials excluded for the analysis (in %)	20% of trials had to be removed due to ...	15	0
53	Reasons for Excluding Trials	Reasons for exclusion	Signal loss	23	0
54	Number of Participants Excluded	Number of participants excluded from the analysis	Three participants excluded	38	0

Running head: REPORTING STANDARDS

55	Quality Threshold	The exact quality threshold for exclusion	40% missing observations	12	10
56	Percentage of Lost Data	Percentage of lost data (proportion of samples in which gaze position could not be identified)	An average of 11% lost data	3	10
57	Test of Assumptions for Missing Data	Test of assumptions for missing data	MCAR, MAR, or MNAR	<1	20
58	Methods Addressing Missing Data	Methods for addressing missing data	Censoring or imputation	3	0
59	Denosing, Filtering or Smoothing of Raw Data	The pre-processing of raw data through denosing, filtering or smoothing	Lowpass filtering	8	60
<i>Dependent measures</i>					
62	Other Transformations of Data	Other transformations of data	Search Index	4	10
63	Blink Detection	The algorithm used to identify blinks	SR or Tobii blink algorithm	2	30
64	Fixation Detection	Event detection procedure	I-DT fixation algorithm	27	60
65	Aggregation Method for Fixations during Data Preprocessing	The aggregation method for fixations during data preprocessing used	AOI method or fuzzy AOIs	20	10
66	Artefact Detection	The artefact detection and removal method used	Blinks and missing data removed by linear interpolation	6	40
67	Pupil Aggregation Method	Specification of the algorithm to calculate the pupil size	Average pupil dilation	82	50
68	Pupil Algorithm	The pupil measures used	Pupil dilation	48	10
<hr/>					
Discussion					
69	Importance of Study Specific Limitations	Limitations mentioned due to the eye-tracking methodology, study specific or general stated?	Low data quality attenuated effect sizes	11	0
70	Importance of Limitations in General	Limitations stated due to the eye-tracking methodology in general	Eye movements do not equate attention	8	0