How to Open Science: A Principle and Reproducibility Review of the Learning Analytics and Knowledge Conference

Aaron Haim
ahaim@wpi.edu
Worcester Polytechnic Institute
Worcester, Massachusetts, USA

Stacy T. Shaw
sshaw@wpi.edu
Worcester Polytechnic Institute
Worcester, Massachusetts, USA

Neil T. Heffernan
nth@wpi.edu
Worcester Polytechnic Institute
Worcester, Massachusetts, USA

ABSTRACT

Within the field of education technology, learning analytics has increased in popularity over the past decade. Researchers conduct experiments and develop software, building on each other's work to create more intricate systems. In parallel, open science — which describes a set of practices to make research more open, transparent, and reproducible — has exploded in recent years, resulting in more open data, code, and materials for researchers to use. However, without prior knowledge of open science, many researchers do not make their datasets, code, and materials openly available, and those that are available are often difficult, if not impossible, to reproduce. The purpose of the current study was to take a close look at our field by examining previous papers within the proceedings of the *International Conference on Learning Analytics and Knowledge*, and document the rate of open science adoption (e.g., preregistration, open data), as well as how well available data and code could be reproduced. Specifically, we examined 133 research papers, allowing ourselves 15 minutes for each paper to identify open science practices and attempt to reproduce the results according to their provided specifications. Our results showed that less than half of the research adopted standard open science principles, with approximately 5% fully meeting some of the defined principles. Further, we were unable to reproduce any of the papers successfully in the given time period. We conclude by providing recommendations on how to improve the reproducibility of our research as a field moving forward.

All openly accessible work can be found in an Open Science Foundation project[1].

CCS CONCEPTS

• **General and reference → Reference works**; **General conference proceedings**; *Validation*; *Reliability*; *Verification.*

KEYWORDS

Open Science, Peer Review, Reproducibility

ACM Reference Format:
Aaron Haim, Stacy T. Shaw, and Neil T. Heffernan. 2023. How to Open Science: A Principle and Reproducibility Review of the Learning Analytics and Knowledge Conference. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK 2023), March 13–17, 2023, Arlington, TX, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3576050.3576071

[1]https://doi.org/10.17605/osf.io/74bzs

LAK 2023, March 13–17, 2023, Arlington, TX, USA

ACM ISBN 978-1-4503-9865-7/23/03.
https://doi.org/10.1145/3576050.3576071

1 INTRODUCTION

Research using learning analytics and open science practices has both increased in popularity over the past decade, with learning analytics aimed at improving our understanding of learning and education and open science focused on improving scientific practices. Learning analytics has developed the field of education technology through applications and integration in Learning Management Systems (LMS), collecting student data to produce better models in a circular development cycle [3, 10]. Open science, which describes practices to make research more open, transparent, and reproducible, has been receiving widespread attention and increased adoption in fields such as Psychology and Statistics[28, 35]. However, learning analytics and open science tend to run parallel to one another, occasionally overlapping. Conferences such as the *International Conference on Learning Analytics and Knowledge* do not actively promote open science within the field, which is disappointing, as open science practices are designed to improve the rigor and reproducibility of research. But even when researchers are aware of open science, they may not have enough knowledge or experience to properly make their work meet open science principles.

The goal of this study is to take a deep look at our field in learning analytics and document the rate of open science adoption and reproducibility of our research. To do this, we examine research papers from the proceedings of the *International Conference on Learning Analytics and Knowledge* (LAK). As this work represents an initial starting point, we relied only on the information provided in each paper (and did not reach out to the authors) and limited our time to only 15 minutes per paper, which was informed by both logistical constraints of investigating 133 research articles and the idea that openness and reproducibility should, when done well, require little time to identify and reproduce. By pointing out the shortcomings in open science and reproducibility of our work and others, the authors hope to stimulate the education technology community to develop and share their practices.

Specifically, this work aims to accomplish the following tasks:

(1) Document and analyze which papers within the proceedings of the *International Conference on Learning Analytics and Knowledge* (LAK) meet the open science principles and subfields defined by this work.
(2) Determine and document whether the results within the papers are reproducible within a 15-minute timeframe.

2 BACKGROUND

2.1 Open Science

Open Science is an 'umbrella' term used to describe that the methodologies, datasets, analysis, and results of any piece of research are openly accessible[15, 35]. While each subcategory has existed for decades, open science itself became prevalent near the beginning of the 2010s with an explosion in popularity near the middle of the decade[28]. In the early 2010s, many researchers were finding issues when peer reviewing others' work: replication failures, unclear practices, overused materials, etc. The open science movement was starting to gain traction as a way to ameliorate these issues by combining mitigations for each part of the problem into a collection of practices, standards, and recommendations. Large-scale studies in the mid-2010's cast doubt on the reproducibility and replicability of research, including replication efforts in psychology (Open Science Collaboration, 2015)[4] and other fields (Baker, 2016)[2]. As a result, many fields began to actively use and improve open science standards, elevating research standards to provide greater clarity. This work will make use of some of the principles of open science in addition to its reproducibility catalyst to evaluate the selected papers.

Open science practices are designed to improve the rigor and reproducibility of our research. Beyond serving a larger goal of cumulative science, where access and openness are no longer a barrier, and materials and data are widely shared, open science aims to makes our work better. The inspiration behind this project came one afternoon, when during a lab meeting we realized that we could not figure out how a previous lab member obtained certain published results. It was unclear which variables they used, what the variable names meant, what code they used, and what packages / libraries they had used on their computer to run the analyses. After lengthy email exchanges, we got to the bottom of it and were able to successfully reproduce the work, but not before asking ourselves an important question: why don't we do a better job of making our published work more reproducible? Admitting our own lack of reproducibility first and foremost, our goal of the current paper is to take a snapshot of the current practices of our field, and document the rate of reproducibility. Our hope is that through this work, we can better understand our field and how to improve our openness and reproducibility in future years.

2.2 Education Technology and Learning Analytics

Education Technology is a field of study investigating the development and evaluation of a learner, their process, and their materials[6, 9]. Progress within the field is cyclical: researchers develop new technology or processes for teachers and learners to use, teachers and learners provide data and feedback, and researchers use the data to better improve their technology or process to begin the cycle over again. With a broad definition, education technology covers a wide variety of topics, most of which create or integrate with Learning Management Systems (LMS)[5]. One such subfield that focuses more on the data collection, modeling, and analysis side is learning analytics.

Learning Analytics is the process by which data collected from learners is investigated to improve learning[3, 10, 34]. The definition and research associated with the topic have received more exposure since the early 2010s, as its integration within education has become more pronounced. Additionally, with the creation of the *International Conference on Learning Analytics and Knowledge* (LAK) in 2011, a common definition was adapted along with a platform for this subfield of research to become independent[29]. As the subfield focuses more on topics that are in line with the open science principles, papers relating to learning analytics submitted to the LAK conference will be used as the dataset for this work.

3 METHODOLOGY: OPEN SCIENCE PEER REVIEW

To answer RQ1, we evaluated every research article, short paper, and poster recorded in LAK proceedings of the last two years: the *12th International Learning Analytics and Knowledge Conference*[31] and the *11th International Learning Analytics and Knowledge Conference*[32]. This work only looks at these two conferences for two reasons. First, open science has slowly expanded since the early 2010s, which is around the time of the first LAK conference[28]. As such, it is impractical to assume that papers submitted to earlier LAK conferences would meet all defined open science principles from this work when they were not originally present. Second, the older a paper, the greater likelihood that the work would face challenges to reproducibility: old packages and libraries, lost data, etc. Thus, for the first look, we examined only the previous two years.

As the LAK conference publishes papers to the Association for Computing Machinery Digital Library (ACMDL), this work uses a Hypertext Transfer Protocol (HTTP) request to query the papers specifically sent to the LAK conference[2]. The documentation for each paper was collected using a Qualtrics[3] survey, which contained the open science principles defined for each paper. In addition to those fields, the survey captures additional metadata: the paper's Digital Object Identifier (DOI) to use as a unique identifier for each row, the proceedings the paper is a part of, and the ACMDL paper classification (Research Article, Short Paper, or Poster). We took 15 minutes to complete the survey for each paper. This 15-minute timeframe was chosen due to logistical constraints (e.g. not spending hours searching for data), as well as the underlying notation that open and reproducible work should be easy to locate and execute. A supplemental document was additionally provided to defend the choices made within the survey (e.g. whether the study was preregistered, had links to open data, had documentation for open materials, etc.). If the links observed in the paper no longer reference their associated resources, a mark will be reported in the survey along with a summary in the documentation about which link has degraded.

[2]The query used is here: https://dl.acm.org/action/doSearch?ConceptID=118647&expand=all&rel=nofollow&sortBy=Ppub_desc&pageSize=50
[3]https://www.qualtrics.com/

3.1 Open Methodology

Open Methodology is an uncommonly used 'umbrella' term under open science, which represents the details of the methods and evaluation used by the authors, including but not limited to the setup, logic flow, aggregations of results, etc. [15]. The methodologies provide additional information that goes beyond the original paper, such that reviewing researchers are able to replicate or reproduce the study themselves.

Papers submitted to the LAK conference tend to follow a similar structure and almost always contain a section dedicated to the methodology–data collection and analysis performed–within the paper. As such, documentation will mark a paper as open methodology if the paper is available on ACMDL[4]. This is considered a naive approach as papers typically do not contain all the information in some structured process to reproduce a given result [13, 16]. The naivete is a limitation of the 15-minute timeframe assigned to each paper, as it would be impossible and impractical to conduct a full review of each paper without the necessary background on the topic covered.

Each paper in ACMDL can be marked with additional metadata tags to define the access provided to each paper. The two metadata tags this work will focus on are the 'Open Access' and 'Public Access' tags. 'Open Access' is used to define a paper that is made freely available online to anyone [1]. Most researchers are required to pay a premium to make a work openly accessible on ACMDL[5]. 'Public Access' meanwhile is a requirement assigned to a paper that must be made freely available within one year of its publication (e.g. publications from the United States National Institutes of Health (NIH)[6]. In both cases, the paper is accessible by anyone on ACMDL. For observational purposes, the documentation will mark whether the paper has an associated access tag in addition to whether it is available on ACMDL.

The scope of this field excludes whether the paper itself has an associated content license. Each paper contains a usage policy that allows a researcher to use the paper for references or analysis purposes, provided that the user does not break ACMDL's policy (e.g. making the paper openly available without paying an appropriate feed).

3.2 Open Data

Open Data is used to define a single or multiple datasets which can be used and made openly accessible without restriction, or with restrictions that do not prevent its use by the recipient [17, 18]. These datasets are provided with a permissive license, terms of use, or released into the public domain, such that the assigned rights are available and automatically hold for the user. This documentation will mark a paper as containing open data if the paper contains a link, or a link to another paper with a link, to the dataset. If a paper does not use or collect data, then this field will be marked as non-applicable.

Not all data can be openly released, however. Some data may contain personally identifiable information that could not be anonymized or made openly available in its entirety. In these cases, an additional option was added for data that can be obtained on request from the author. Our requirement was that this must be explicitly stated within the paper; otherwise, the documentation marks the field as containing no open data.

In addition to the open data field, a separate field was added to determine whether there was data documentation available for the dataset. Data documentation contains a one-to-one mapping of the field name in the dataset to a description of what the field is and what values it may contain (e.g., a 'data dictionary', table, specification, etc.). The documentation will mark a paper as having data documentation if the paper or a separate document contains all fields in the dataset. If a paper does not use or collect data, then this field will also be marked as non-applicable.

The scope of this field is broadened from the original definition in a few instances. First, the dataset is not required to have a permissive license, terms of use, or be released into the public domain to be considered open data. This assumes that researchers are not familiar with copyright law in their country, or that there exists dataset licenses. Dataset licenses also tend to be complex and generally ambiguous, making it even more difficult to determine whether the dataset can be openly used [12, 25]. Second, the data documentation can be marked as partial, indicating that there exists some documentation on at least one field that may show up in the dataset. This assumes that even if a dataset is provided, the researcher may not provide any documentation for the fields or briefly glance over it in a table or subsection within the paper itself.

3.3 Open Materials

Open Materials is another 'umbrella' term used to define technologies that can be used and made openly accessible without restriction, or with restrictions that do not prevent its use by the recipient. Within the literature, open materials tend to be synonymous with **Open Source** as it is defined within the context of software development [11, 24]. Open materials define a slightly broader scope: not only containing open source software, but also technologies that are free-to-use (e.g. Google Sheets[7]). This documentation will mark a paper as containing open materials if the paper contains a link, or a link to another paper with a link, all of the materials and source used. If a paper does not use a material (e.g. theoretical or argumentative papers), then this field will be marked as non-applicable.

In addition to the open materials field, three more subfields are added. First, a materials documentation field is added, which determines whether the materials and source provided contain documentation for their usage. Documentation is necessary to understand how to use material without having to perform an individual analysis of the source, especially if the source is not present [7]. Second, a README field is added, which determines whether a README is present within the source. READMEs usually contain an introduction to the source in addition to information on the documentation

[4]Typically, all research articles, short papers, and posters as defined by ACMDL would meet this requirement. However, some papers from previous conferences, such as the *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, are classified as research articles[30]. To address this issue, a misclassified field is added to mark whether the paper type provided by ACMDL is incorrect

[5]Learn more about open access through ACMDL on their website: https://www.acm.org/publications/openaccess

[6]Learn more about NIH public access here: https://publicaccess.nih.gov/

[7]https://docs.google.com/spreadsheets

or setup necessary to use the provided source [14]. Finally, a license field is added, which determines whether the source has a permissible software license, such that the source can be used openly by any user [8, 24, 27]. These fields will be marked as non-applicable only if the open materials field is marked as non-applicable.

The scope of this field has also been broadened from its original definition. The open materials and materials documentation fields can be marked as partial, indicating that at least one of the materials or materials documentation provided is openly accessible. This assumes that a researcher may not use only openly accessible materials in favor of in-house or organization provided tools. In addition, paid commercial products are excluded from the materials determinant. There is no prior literature on whether commercial products are considered open materials, and the discussion on such a topic is highly complex, so this paper would like to avoid any claims until some established standard has been determined.

3.4 Preregistration

Preregistration is a term used to define the steps conducted for the paper without knowledge of the research outcomes [20, 21, 33]. These are typically made prior to the initial starting point of an experiment, such that any prediction or result obtained is unbiased by any observations made by the researcher. Additionally, if the experiment needs to be altered, an update or additional preregistration can be made to preserve the initial methodologies. This registration is made openly accessible and is held on some independent registry (e.g., the Open Science Foundation[8]). Standard preregistration practices involve citing the preregistration within the body of the paper and identifying which analyses/aspects of the research were preregistered. This documentation will mark a paper as containing a preregistration if the paper contains a link to the preregistration or a project containing the preregistration. If a paper does not need to use a preregistration (e.g. theoretical or argumentative papers), then this field will be marked as non-applicable.

4 METHODOLOGY: REPRODUCIBILITY

Reproducibility is a term used to define performing the exact same experiment with the exact same method to see if the reviewer can achieve the exact results reported in the paper[19, 22, 23]. However, in context, reproducibility can have different levels of completeness. As such, this work will define reproducibility as the ability to take the dataset collected and the analysis method used and return the exact same results and figures given in the paper. It is impractical to collect the data from scratch or write a similar implementation based on pseudocode in the paper, so if the dataset or the source is not provided, then the paper will be marked as non-reproducible. The reproducibility field will only be marked as non-applicable if either the dataset or materials fields is also marked as non-applicable.

To answer RQ2, a reproducibility test was conducted across all papers in which open data and open materials were available. While the requirement could contain additional fields, such as whether a README is present for necessary setup or if a license is provided to legally use the source, this work assumes that the minimum required to reproduce a paper is the dataset and source which consumes it. Each reproducibility test will run for a maximum of 15 minutes in addition to the 15 minutes assigned for the open science principles assessment. This timeframe is arbitrarily chosen as a reasonable length to run a provided command or script as this work assumes that, at most, only a few setups, preprocessings, and analyses need to be run. The 15-minute timeframe only applies to a sequence of non-author actions to perform. This means that if the author has laid out the necessary actions to perform, then even if the 15-minute timeframe is exceeded, if the exact results are produced, then it will be marked as reproducible. Additionally, the 15-minute timeframe will only be exceeded once the last running action of a sequence has finished executing.

As the reproducibility of a work should not be tied to any specific platform, the authors' personal machine will be used as a testing site. The authors' machine runs Windows 10[9] with a PowerShell[10] shell. If the machine is not powerful enough, then the same attempt will be performed using a machine built for large quantities of data and computationally heavy processing. The big data machine runs Ubuntu[11] with a Bash[12] shell. As there is great variability in the tools needed to reproduce a paper, the machine will start with a clean instance of whatever software is used (e.g. the base software with no addons or packages). From there, the test will follow any setup instructions provided by the author, if present, and then attempt to run the source using the associated method.

As there are particular nuances and problems that are specific to a piece of software, this work will also define some common problems in popular software and the steps taken towards providing a reproducibility test.

4.1 Python

Python[13] is a programming language used commonly by researchers for integration to read data, conduct analyses, and generate results in a consumable format. The language itself is under an open source license and contains thousands of libraries, known as packages, tailored to most needs. Although a highly useful tool, there are a number of reproducibility issues if no steps are taken to freeze the author's current environment.

4.1.1 Python 2 vs Python 3. Python has two major versions which are used by researchers: Python 2, which has been discontinued as of January 1st, 2020, and Python 3, which is the recommended version to use. Although Python 2 is no longer supported, there are numerous repositories that contain sources written for that language. This is especially true in the research field, where many sources are never updated after publication. Additionally, unless updated explicitly, the author would not have the most recent release of a Python version, meaning that authors using the same machine would likely be using the same outdated version. Furthermore, if an author is not familiar with Python and has both versions installed, it may not be clear to them which version they are using. The command `python` may refer to Python 2 or Python 3 depending on the installation and operating system. Therefore, it is likely that a paper may be using Python 2 compared to Python 3.

[8]https://osf.io/registries

[9]Windows 10: Version 10.0.19044 Build 19044
[10]PowerShell: Version 5.1.19041.1682
[11]Ubuntu: Version 20.04.5 LTS
[12]Bash: Version 5.0.17(1)-release
[13]https://www.python.org/

4.1.2 Within Python 3.x. The same issue between Python 2 and Python 3 can be seen with the minor versions of Python 3 as well. If a feature in a later minor version is used within a Python script, if the end user contains an earlier version of Python 3, then an error will be thrown. This is even more likely than the Python 2 issue above as versions 3.7-3.10 are still maintained in 2022. Therefore, the survey will mark a paper as not reproducible if an explicit Python version is not provided.

4.1.3 Beta Packages. Many of the commonly used packages within Python for researchers do not currently or did not have a stable release during the prior two conferences. While stable versions are likely to contain the method used within the author's source, beta versions make no guarantee the method would still be present in the exact same place in a later version. Additionally, some environment setup tools use beta packages because of the fragility of certain packages. Therefore, the documentation will mark a paper as not reproducible if a package is missing or if a package is not given an explicit version.

4.1.4 Minor: Platform-Dependent. There are some instances within Python where using the `os` module, the behavior of a script may change depending on the operating system it is run on. This is a minor concern, as it is unlikely a researcher would stumble into such a scenario. However, for brevity, the machine used to test reproducibility runs Windows 10[9] through PowerShell[10], with a second machine running Ubuntu[11] through Bash[12] if the script failed and used the `os` module. If the source failed on one platform but not on the other, it will be mentioned in the documentation. Otherwise, nothing was mentioned on the platform itself.

4.2 R

R[14] is a programming language commonly used by researchers for statistical computing and graphics. The language is licensed as free-to-use software and contains numerous packages written by statisticians and published in peer-reviewed journals. However, like Python, R can have a number of reproducibility issues, especially as it does not have a standardized way of managing package versions.

4.2.1 Hidden Language and Tools. One of the most common things researchers miss when initially setting up an R workspace is that there might be additional tools and language binaries that need to be downloaded to properly download the packages within the script itself. Additionally, as R is provided with some Linux operating system distributions, some setups may not even be applicable to some users. When reproducing R scripts, it is recommended to have a development environment setup (e.g., `Rtools` for Windows machines, `Xcode` for MacOS machines). To mitigate this, `devtools` will be installed as part of the initial setup process for R environments.

4.2.2 Version Management. As mentioned in the Python version sections, using different versions of packages or the base language itself can cause unintended errors for the end user. However, unlike Python, R does not have a standard method of managing packages built into the language itself. As such, additional care must be taken that the package manager used is consistent in addition to the packages. Therefore, the documentation will mark a paper as not reproducible if a package is missing, a package is not given an explicit version, or if the package manager is not provided.

4.3 JavaScript: Node.js

Node.js[15] is a JavaScript runtime used to create network applications using an asynchronous, event-driven model. Node is used by researchers to create front-end applications to record or collect data from users within an experiment. Using npm[16], Node has access to a large number of packages that can be version locked and easily transferable to other machines, almost nullifying the version management issue in other languages. As long as a version of Node.js and npm is provided, then it is highly likely that the front end will be reproducible.

4.3.1 Minor: Local Host Ports. If multiple connections are running on a local machine, the ports required by the front-end application or any connections may be mismatched. In these cases, the machine will make sure that all running applications using ports are closed and provide an additional configuration in case a mismatch occurs.

5 RESULTS

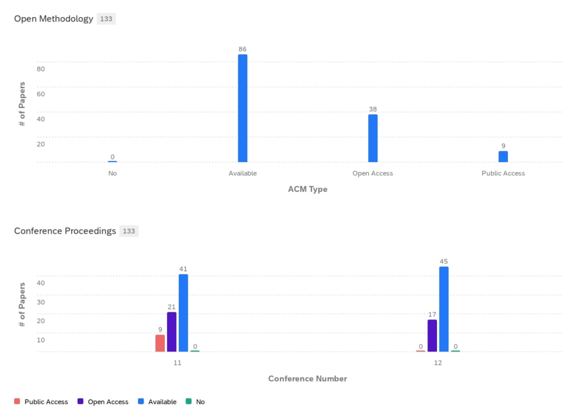

Figure 1: A breakdown of the open methodologies of papers submitted to the 11th and 12th conference. Shows the total (top) and breakdown per conference (bottom)..

For RQ1, as shown in Figure 1, there were 133 papers within the LAK conference proceedings, 71 submitted to the 11th LAK conference and 62 to the 12th, which are classified as 'Research Article', 'Short Paper', and 'Poster'.

In Figure 2, out of the submitted papers, 9 are publicly accessible, 38 are openly accessible, and the remaining 86 are available but not open. None were classified incorrectly by ACMDL.

In Figure 3, we can see that only 5% of all papers make their raw dataset available with 2% explicitly mentioning that the data can be requested. For documentation, 2% of all papers fully document their work, while 38% mention some form of dataset documentation at least once in their paper. The 2% of works that fully document their dataset are using publicly available data which already has the

[14] https://www.r-project.org/

[15] https://nodejs.org/

[16] https://www.npmjs.com/

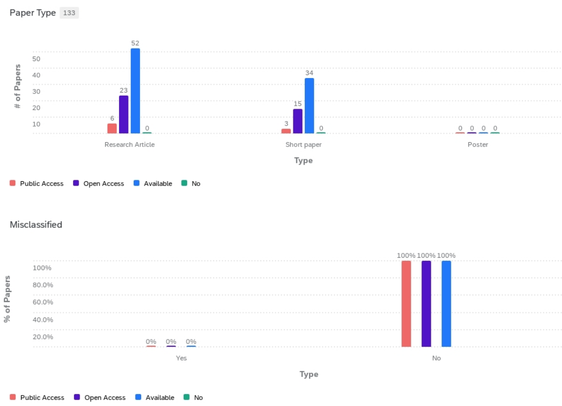

Figure 2: A breakdown of the open methodologies of papers submitted to the 11th and 12th conference. Each paper is broken down into its type (top) and whether it was misclassified by ACMDL (bottom).

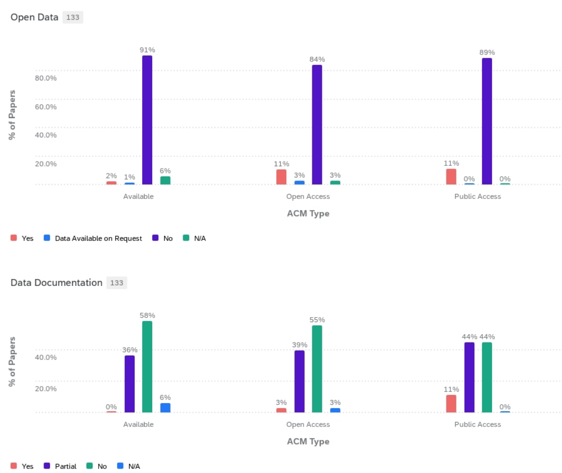

Figure 3: A breakdown of the open data of papers submitted to the 11th and 12th conference by open methodology. Breaks down the requirements into having the data publicly available (top) and the documentation for the dataset (bottom).

available documentation. Open and Public Access papers adhere more to open data principles compared to the other papers, with Public Access papers adhering slightly better.

For open materials in Figure 4, we can see only 5% of papers provide the full materials needed to reproduce the paper, while 29% mention using at least one open material. In combination with the data fields, only 2% of the papers have the full materials available along with the raw dataset needed to produce a reproduction. No papers fully documented their materials. The materials documentation either could not be located within the 15-minute timeframe or was not available for every public method in the provided source. Open access papers tend to be more in line with open materials

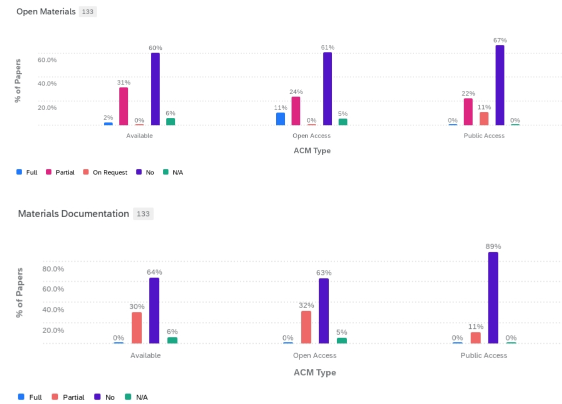

Figure 4: A breakdown of the open materials of papers submitted to the 11th and 12th conference by open methodology. Breaks down the requirements into having the materials publicly available (top) and the documentation for the materials (bottom).

and documentation by a small margin compared to papers only available on ACMDL. Public Access papers were noticeably worse at providing their materials and documentation.

Figure 5: A breakdown of the open materials subfields and preregistration of papers submitted to the 11th and 12th conference by open methodology. Breaks down the requirements into having a README (left), a permissible license (middle), and a preregistration (right).

Looking at the open materials subfields and preregistration in Figure 5, 7% of papers contained a README with 3% of those containing a permissible license to use their source. Only one paper provided a preregistration, though its materials could not be accessed. A paper is only marked as non-applicable for these fields if the materials fields are also non-applicable, so the graphs are conflated to make a point. If only looking at papers with open materials, 75% of the papers had a README while 45% contained a permissible license. Open access papers provided more papers with READMEs and permissible licenses, followed by Available papers, and finally none by Public Access papers. This correctly correlates with the breakdown of the open materials field. The only preregistration is provided by an Available paper.

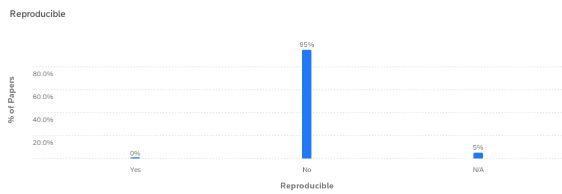

Figure 6: A breakdown of the Reproducibility Test of papers submitted to the 11th and 12th conference.

For RQ2, as shown in Figure 8, not a single paper was reproducible within a 15-minute timeframe. The main cause for non-reproducibility was du e to a lack of libraries provided or that the libraries themselves did not have a version number attached to them. The secondary cause was that some of the sources used some form of randomness but did not make the randomness deterministic to accurately reproduce (e.g. fixed seed). One paper additionally provided a Bash script, which would only work on Linux and older versions (newer with some work) of MacOS[17].

Although none of the work was reproducible in a 15-minute timeframe, if given a longer time period, we estimated that the 2% of papers that contain both the raw dataset and source were likely to be reproducible with enough configuration.

6 LIMITATIONS

We did not investigate all aspects of open science. Some principles are simply impossible to achieve: open peer review[26] is not possible as the LAK conference uses a double-blind review method to determine which papers to accept. Other principles are typically not documented explicitly in the paper or its resources: open educational resources[36] are not explicitly mentioned or unclear where to obtain from and how. It is unlikely that these principles could ever be easily interpreted from the majority of papers without interference from the original author. Additionally, some principles such as open data cannot be used with datasets that might identify participants, making it difficult to discern whether the lack of open data is due to ethical considerations or simply because the data were not made open.

[17] As this script simply ran a Python command, it could be stripped out. However, this work assumed that a researcher will simply attempt to run things without further explanation, so it is likely to cause an issue in understanding

This review also did not attempt to contact authors to receive a better understanding of the context or reasoning behind not meeting these principles. As this paper is conducted as an initial starting point for open science in education technology, this limitation was by design. However, it would be useful to replicate this work and contact the authors so that any previously unclear or uncertain principles can be specified and provided. This would also clear up any reproducibility issues without the tedious cycle of repetition testing. Furthermore, this would allow the authors of the replication or the original paper to work with each other to meet other principles and provide an easy method of reproduction.

The selection of the survey itself is subjective from the authors of this paper. While procedures are put in place to limit the number of opinions provided by the authors, there are still a number of instances that had to be defended for why a particular requirement was or was not met. The supplemental document is provided to mitigate these concerns by providing clear reasoning by the author as to any additional explanations not possible in the survey itself. In addition, the dynamic document itself is provided as a way for authors of the review papers to provide a rebuttal or provide additional links to where their work is currently located to meet the defined principles while also being reproducible.

Finally, reproducibility was limited to the machines owned by the authors of this paper. Due to the number of different setups on machines, there is likely at least one other machine from which a paper can be reproduced on. In order to provide a more general view, a clean machine was used to set up, while unclear or incomplete instructions were immediate grounds for non-reproducibility. However, a second replication should be performed to try to find a working solution and provide a clear setup procedure for reproduction on other machines.

7 FUTURE WORK

As learning analytics is one subfield of the education technology field, future work could replicate this metareview for other conference proceedings, such as for the *International Conference on Artificial Intelligence in Education* and the *International Conference on Educational Data Mining*. Future work with conference papers should contain an additional step on the previous replication to more closely entwine and promote open science in the Educational Technology field.

Future work should also include contacting authors after the survey to get any missing information, instructions, or nuances for the available principles and reproducibility. In addition, a separate question can be asked as to whether the author-held data can be made openly accessible.

Another direction for future work will be to improve upon our reproduction setup. This will consist of documenting the required materials and versions used to test the software in addition to providing simple setups to run the reproduction. Each of these steps will then be sourced to the required repository or hosted ourselves if not possible. In addition, a DOI will be assigned to any dynamic resources to provide a permanent identifier in case of reference degradation.

As a supplement for those who want to learn more about open science and reproducibility best practices, workshops will be submitted to the conferences. The first workshop of each conference will cover the concepts of open science and how to make work reproducible. Additional workshops from then on will promote concepts that were not previously reported and papers that provide methodologies, datasets, or materials that are openly accessible. It is not reasonable that every researcher will be able to attend a workshop. Therefore, any resources shown in a workshop will be made openly available afterward for non-attending researchers to consume. Additional backups of the resources will be attached to a DOI link in case the dynamic reference ever degrades.

Finally, the authors will attempt to have conferences promote open science and reproducibility designs within their Call for Papers and submission portals such that their resources can be easily accessed by other researchers looking to use the associated methodologies, dataset, or materials. Work has already begun for having the *International Conference on Educational Data Mining* promote open science and reproducibility designs within their conferences as their papers are self-published and openly accessible.

8 CONCLUSION

Although about 30% of the papers meet some of the very broad partial definitions for the principles of open science, approximately 5% of the papers meet most of the full definitions. Most of the provided subcategories are simply neglected, which could be due to negligence, but most likely is due to inexperience and unclear resources. An average researcher cannot be expected to know anything about the open science principles, much less understand the subcategories required within them. Additionally, most reproducibility metrics tend to occur within networks of people who are familiar with each other, so as long as the researcher understands how to reproduce their own work, they will almost never run into this issue.

In general, a paper does not need to meet all the open science requirements or the reproducibility metric, depending on circumstances. For example, datasets with personally identifiable information may not be released openly. In other cases, works or grants may restrict making any raw datasets or analysis open access, especially if the sponsor is having the researcher develop for a commercial product. Some work may not even be completed by the time the paper is released. In most cases, some commercial tools may provide ease of use while the author may not have distribution access to the raw data. This work actually fails to fully meet the open data and open materials requirements. The raw dataset is papers from the ACMDL site, which may or may not be openly accessible. Additionally, while the survey portion of a Qualtrics survey falls under open materials, the graph reporting tools fall under the commercial product, which, as stated, is ambiguous as to whether it meets the open science requirements.

This work is meant as a way to point out some of the complexities when reviewing, reproducing, or even replicating an author's work. The authors of this paper want to promote the sharing of resources and provide a better understanding of research in the past, present, and future. We hope that through a close look at our own work and other researchers, we can encourage the education technology community to develop and share their practices. Using open science principles and reproducibility metrics is one method to preserve the environment of the line of research such that additional branches can build off without having to rediscover the previous understanding.

ACKNOWLEDGMENTS

We would like to thank the NSF (e.g., 2118725, 2118904, 1950683, 1917808, 1931523, 1940236, 1917713, 1903304, 1822830, 1759229, 1724889, 1636782, & 1535428), IES (e.g., R305N210049, R305D210031, R305A170137, R305A170243, R305A180401, R305A120125, & R305R220012), GAANN (e.g., P200A180088 & P200A150306), EIR (U411B190024 & S411B210024), ONR (N00014-18-1-2768), and Schmidt Futures. None of the opinions expressed here are that of the funders. We are funded under an NHI grant (R44GM146483) with Teachly as a SBIR.. We would also like to extend our thanks to the blinded reviewers who provided their suggestions towards improving this paper.

REFERENCES

[1] Karen M Albert. 2006. Open access: implications for scholarly publishing and medical libraries. *Journal of the Medical Library Association* 94, 3 (2006), 253. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1525322/
[2] Monya Baker. 2016. 1,500 scientists lift the lid on reproducibility. *Nature* 533, 7604 (01 May 2016), 452–454. https://doi.org/10.1038/533452a
[3] Doug Clow. 2013. An overview of learning analytics. *Teaching in Higher Education* 18, 6 (2013), 683–695. https://doi.org/10.1080/13562517.2013.827653 arXiv:https://doi.org/10.1080/13562517.2013.827653
[4] Open Science Collaboration. 2015. Estimating the reproducibility of psychological science. *Science* 349, 6251 (2015), aac4716. https://doi.org/10.1126/science.aac4716 arXiv:https://www.science.org/doi/pdf/10.1126/science.aac4716
[5] Miguel Á. Conde, Francisco J. García-Peñalvo, María J. Rodríguez-Conde, Marc Alier, María J. Casany, and Jordi Piguillem. 2014. An evolving Learning Management System for new educational environments using 2.0 tools. *Interactive Learning Environments* 22, 2 (2014), 188–204. https://doi.org/10.1080/10494820.2012.745433 arXiv:https://doi.org/10.1080/10494820.2012.745433
[6] Katie Mcmillan Culp, Margaret Honey, and Ellen Mandinach. 2005. A Retrospective on Twenty Years of Education Technology Policy. *Journal of Educational Computing Research* 32, 3 (2005), 279–307. https://doi.org/10.2190/7W71-QVT2-PAP2-UDX7 arXiv:https://doi.org/10.2190/7W71-QVT2-PAP2-UDX7
[7] Barthélémy Dagenais and Martin P. Robillard. 2010. Creating and Evolving Developer Documentation: Understanding the Decisions of Open Source Contributors. In *Proceedings of the Eighteenth ACM SIGSOFT International Symposium on Foundations of Software Engineering* (Santa Fe, New Mexico, USA) *(FSE '10)*. Association for Computing Machinery, New York, NY, USA, 127–136. https://doi.org/10.1145/1882291.1882312
[8] Arnoud Engelfriet. 2009. Choosing an open source license. *IEEE software* 27, 1 (2009), 48–49. https://doi.org/10.1109/MS.2010.5
[9] Maya Escueta, Vincent Quan, Andre Joshua Nickow, and Philip Oreopoulos. 2017. *Education Technology: An Evidence-Based Review.* Working Paper 23744. National Bureau of Economic Research. https://doi.org/10.3386/w23744
[10] Rebecca Ferguson. 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning* 4, 5-6 (2012), 304–317. https://doi.org/10.1504/IJTEL.2012.051816 arXiv:https://www.inderscienceonline.com/doi/pdf/10.1504/IJTEL.2012.051816
[11] Johndan Johnson-Eilola. 2002. Open Source Basics: Definitions, Models, and Questions. In *Proceedings of the 20th Annual International Conference on Computer Documentation* (Toronto, Ontario, Canada) *(SIGDOC '02)*. Association for Computing Machinery, New York, NY, USA, 79–83. https://doi.org/10.1145/584955.584967
[12] Mashael Khayyat and Frank Bannister. 2015. Open data licensing: more than meets the eye. *Information Polity* 20, 4 (2015), 231–252. https://doi.org/10.3233/IP-150357
[13] K.D. Knorr. 1981. *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science.* Elsevier Science Limited, Los Angeles, CA, USA. https://books.google.com/books?id=HF8cAQAAIAAJ
[14] Miika Koskela, Inka Simola, and Kostas Stefanidis. 2018. Open Source Software Recommendations Using Github. In *Digital Libraries for Open Knowledge*, Eva Méndez, Fabio Crestani, Cristina Ribeiro, Gabriel David, and João Correia Lopes (Eds.). Springer International Publishing, Cham, 279–285. https://doi.org/10.1007/978-3-030-00066-0_24
[15] Peter Kraker, Derick Leony, Wolfgang Reinhardt, and Günter Beham. 2011. The case for an open science in technology enhanced learning. *International Journal of Technology Enhanced Learning*

3, 6 (2011), 643–654. https://doi.org/10.1504/IJTEL.2011.045454 arXiv:https://www.inderscienceonline.com/doi/pdf/10.1504/IJTEL.2011.045454

[16] B. Latour, S. Woolgar, and J. Salk. 1986. *Laboratory Life: The Construction of Scientific Facts*. Princeton University Press, Princeton, NJ, USA. https://books.google.com/books?id=XTcjm0flPdYC

[17] Jennifer C. Molloy. 2011. The Open Knowledge Foundation: Open Data Means Better Science. *PLOS Biology* 9, 12 (12 2011), 1–4. https://doi.org/10.1371/journal.pbio.1001195

[18] Peter Murray-Rust. 2008. Open data in science. *Nature Precedings* 1, 1 (18 Jan 2008), 1. https://doi.org/10.1038/npre.2008.1526.1

[19] E.M. National Academies of Sciences, P.G. Affairs, E.M.P.P. Committee on Science, B.R.D. Information, D.E.P. Sciences, C.A.T. Statistics, B.M.S. Analytics, D.E.L. Studies, N.R.S. Board, D.B.S.S. Education, et al. 2019. *Reproducibility and Replicability in Science*. National Academies Press, Washington, D.C., USA. https://books.google.com/books?id=6T-3DwAAQBAJ

[20] Brian A. Nosek, Emorie D. Beck, Lorne Campbell, Jessica K. Flake, Tom E. Hardwicke, David T. Mellor, Anna E. van 't Veer, and Simine Vazire. 2019. Preregistration Is Hard, And Worthwhile. *Trends in Cognitive Sciences* 23, 10 (01 Oct 2019), 815–818. https://doi.org/10.1016/j.tics.2019.07.009

[21] Brian A. Nosek, Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 2018. The preregistration revolution. *Proceedings of the National Academy of Sciences* 115, 11 (2018), 2600–2606. https://doi.org/10.1073/pnas.1708274114 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.1708274114

[22] Brian A. Nosek, Tom E. Hardwicke, Hannah Moshontz, Aurélien Allard, Katherine S. Corker, Anna Dreber, Fiona Fidler, Joe Hilgard, Melissa Kline Struhl, Michèle B. Nuijten, Julia M. Rohrer, Felipe Romero, Anne M. Scheel, Laura D. Scherer, Felix D. Schönbrodt, and Simine Vazire. 2022. Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology* 73, 1 (2022), 719–748. https://doi.org/10.1146/annurev-psych-020821-114157 arXiv:https://doi.org/10.1146/annurev-psych-020821-114157 PMID: 34665669.

[23] Prasad Patil, Roger D. Peng, and Jeffrey T. Leek. 2016. A statistical definition for reproducibility and replicability. *bioRxiv* 1, 1 (2016), 1–1. https://doi.org/10.1101/066803 arXiv:https://www.biorxiv.org/content/early/2016/07/29/066803.full.pdf

[24] Bruce Perens et al. 1999. The open source definition. *Open sources: voices from the open source revolution* 1 (1999), 171–188. https://www.researchgate.net/publication/200027107_Perens_Open_Source_Definition_LG_26

[25] Gopi Krishnan Rajbahadur, Erika Tuck, Li Zi, Zhang Wei, Dayi Lin, Boyuan Chen, Zhen Ming Jiang, and Daniel Morales German. 2021. Can I use this publicly available dataset to build commercial AI software? Most likely not. *CoRR* abs/2111.02374 (2021), 1–1. arXiv:2111.02374 https://doi.org/10.48550/arXiv.2111.02374

[26] Tony Ross-Hellauer. 2017. What is open peer review? A systematic review [version 2; peer review: 4 approved]. *F1000Research* 6, 588 (2017), 1. https://doi.org/10.12688/f1000research.11369.2

[27] Hendrik Schoettle. 2019. Open Source License Compliance-Why and How? *Computer* 52, 08 (aug 2019), 63–67. https://doi.org/10.1109/MC.2019.2915690

[28] Barbara A. Spellman. 2015. A Short (Personal) Future History of Revolution 2.0. *Perspectives on Psychological Science* 10, 6 (2015), 886–899. https://doi.org/10.1177/1745691615609918 arXiv:https://doi.org/10.1177/1745691615609918 PMID: 26581743.

[29] The Society for Learning Analytics Research 2011. *LAK '11: Proceedings of the 1st International Conference on Learning Analytics and Knowledge* (Banff, Alberta, Canada). The Society for Learning Analytics Research, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2090116

[30] The Society for Learning Analytics Research 2012. *LAK '12: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (Vancouver, British Columbia, Canada). The Society for Learning Analytics Research, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/2330601

[31] The Society for Learning Analytics Research 2021. *LAK21: LAK21: 11th International Learning Analytics and Knowledge Conference* (Irvine, CA, USA). The Society for Learning Analytics Research, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3448139

[32] The Society for Learning Analytics Research 2022. *LAK22: LAK22: 12th International Learning Analytics and Knowledge Conference* (Online, USA). The Society for Learning Analytics Research, Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3506860

[33] Anna Elisabeth van 't Veer and Roger Giner-Sorolla. 2016. Pre-registration in social psychology—A discussion and suggested template. *Journal of Experimental Social Psychology* 67 (2016), 2–12. https://doi.org/10.1016/j.jesp.2016.03.004 Special Issue: Confirmatory.

[34] Olga Viberg, Mathias Hatakka, Olof Bälter, and Anna Mavroudi. 2018. The current landscape of learning analytics in higher education. *Computers in Human Behavior* 89 (2018), 98–110. https://doi.org/10.1016/j.chb.2018.07.027

[35] Ruben Vicente-Saez and Clara Martinez-Fuentes. 2018. Open Science now: A systematic literature review for an integrated definition. *Journal of Business Research* 88 (2018), 428–436. https://doi.org/10.1016/j.jbusres.2017.12.043

[36] David Wiley, T. J. Bliss, and Mary McEwen. 2014. *Open Educational Resources: A Review of the Literature*. Springer New York, New York, NY, 781–789. https://doi.org/10.1007/978-1-4614-3185-5_63