

Article

Chaos creator: Misinformation inoculation in information literacy instruction

<http://dx.doi.org/10.11645/18.2.13>

Mandi Goodsett

Performing Arts & Humanities Librarian, OER & Copyright Librarian, Cleveland State University.
Email: a.goodsett@csuohio.edu. ORCID: [0000-0003-0477-2828](https://orcid.org/0000-0003-0477-2828).

Abstract

Combating the spread of misinformation is a struggle that has inspired considerable research in the fields of psychology, education, political science, and information science, among others. Such research has found that “prebunking” or “inoculation” techniques—strategies that reduce the acceptance of misinformation before one has encountered it—have had marked success. However, there is little evidence that librarians are deliberately employing inoculation techniques in their information literacy (IL) instruction. Via a quasi-experimental study, this research explores the effect of prebunking techniques in an IL instruction session on undergraduate students’ ability to recognise misinformation. The prebunking techniques are delivered through a competitive game called Chaos Creator, based on the Bad News game developed by researchers at Cambridge University. Results of the study show that misinformation inoculation techniques are more effective than the popular source evaluation tool, the CRAP test, in helping students identify misinformation. However, misinformation inoculation techniques can backfire, causing students to become overly sceptical of trustworthy messages.

Keywords

inoculation; information literacy; library instruction; misinformation; prebunking

1. Introduction

In early 2022, the world was saddened to learn of the passing of talented actress and comedian Betty White. However, within hours of the news, rumours and dubious claims about her death began circulating on social media. One viral Twitter post claimed, “She got the booster on Dec 28th ... oooopps!”, implying that the COVID-19 vaccination booster shot was somehow related to White’s death (Emery, 2022). The Associated Press wrote in an article a short time later that Betty White had not had her booster shot on December 28th and that the quote was false (Phan, 2022). This did not prevent hundreds of Twitter users from engaging with the tweet and

This [Open Access](#) work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#), allowing others to share and adapt this content, even commercially, so long as the work is properly cited and applies the same license. Copyright for the article content resides with the authors, and copyright for the publication layout resides with the Chartered Institute of Library and Information Professionals, Information Literacy Group.

claiming it supported their beliefs that the COVID-19 booster was unsafe. This claim about Betty White is just one example of how false information, especially about topics that frighten, alarm, or enrage social media users, can spread quickly and become widely adopted.

In the past decade, concern over misinformation and its negative effect on individuals and society has surged. Misinformation is “false information that is spread either by mistake or with the intent to mislead” and it can result in bad decision-making at the individual or societal level, with potentially significant consequences (Lewandowsky et al., 2020, p. 4). The intuitive solution to the spread of misinformation is to disseminate corrections to the false messages; however, it can be very difficult to convince individuals to change their beliefs, and, even if this is successful, the effect may not last long. Misinformation can continue to influence behaviour even after the individual claims to have updated a previously held belief. This is known as the “continued influence effect” (Lewandowsky et al., 2012, p. 113). For this reason, misinformation is thought to be “sticky,” and considerable time and care may be necessary to successfully debunk false claims (Lewandowsky et al., 2020, p. 6).

As concern about the spread of misinformation has increased, librarians have promoted information literacy and library instruction as important contributors to improved misinformation-detection skills (Becker, 2016; De Paor & Heravi, 2020; Gibson & Jacobson, 2018). However, many of the debunking strategies that librarians have been using do not always align with psychologists’ understanding of mental models. For example, sharing tips for fact-checking information is unlikely to alter beliefs that have already been incorporated into someone’s mental model and that align with their worldview (Sullivan, 2019). Librarians may have more success trying “prebunking” techniques, such as raising scepticism about the persuasion techniques commonly used in misinformation messages (a tactic called misinformation “inoculation”). While evidence of the success of these inoculation techniques has existed for years, there is little exploration in the literature of their use in library instruction. The present study attempts to address this knowledge gap by exploring the impact of prebunking techniques in library instruction on students’ ability to recognise misinformation. The study tries to answer the question: To what extent does incorporating misinformation inoculation techniques into information literacy instruction improve first year college students’ ability to recognise misinformation?

The three hypotheses that this study attempts to test are as follows:

- H1) Students who receive the treatment condition (misinformation inoculation) will see a greater improvement from their pre- to their post-test scores than the students who receive the control condition (CRAP test).
- H2) Students who receive the treatment condition will perform better on their post-test than on their pre-test.
- H3) Students who receive the treatment condition will have a greater increase in their confidence levels after the intervention than students who receive the control condition.

2. Literature Review

As the media and academia have paid increased attention to the problem of misinformation, librarians have explored their role in combatting false information spread online. One study found over one hundred articles about “fake news” had been published in academic library and information science journals from 2017 to 2020 alone (Sahoo et al., 2021). Some librarians have argued that teaching information literacy (IL) is one of the most effective strategies for

combatting misinformation (Batchelor, 2017). In 2018, the Chartered Institute of Library and Information Professionals (CILIP) revised their own definition of IL to include information contexts outside of the classroom, and they emphasised the value of “the ability to think critically and make balanced judgements about any information we find and use” (CILIP, 2018). This revised definition reflects the understanding of many librarians that part of the librarian’s role is to instil IL skills in students and the public, so that they are better able to discern and combat misinformation (De Paor & Heravi, 2020).

The strategies used by librarians to combat misinformation using IL include credit-bearing courses (Eva & Shea, 2018), interactive workshops (Eva & Shea, 2018; Lefkowitz, 2017), news literacy training (Banks, 2016), checklists and acronyms (Batchelor, 2017; Jacobson, 2017), and collection development strategies (De Paor & Heravi, 2020; Finley et al., 2017). Sullivan (2018) points out that at least some of these methods, such as simply providing access to quality information through collection development, are unlikely to be effective, given psychological research findings about how people’s prior beliefs impact their willingness to consider conflicting information. De Paor and Heravi (2020) argue that IL instruction may be helpful, but it isn’t sufficient, and that librarians should reframe their approach to include other literacies, such as metaliteracy (Mackey & Jacobson, 2011), as well. The strategies that some libraries use to combat misinformation may be effective, but one area that librarians might explore more is in using strategies promoted by psychology and education researchers who study misinformation.

For years, psychology and education researchers have been attempting to answer the questions, “How can we decrease acceptance of misinformation?” and “Why does misinformation seem to be so difficult to correct?” (Compton et al, 2021). In the process, they have uncovered many debunking strategies—some that have been shown to be effective, and some that have not. While initial efforts focused on correcting misperceptions, researchers argue that individuals create a mental model of an event or situation and are reluctant to modify the model with new information when the existing belief has sufficient explanatory power (Ecker, Lewandowsky, Swire, & Chang, 2011; Swire & Ecker, 2018). When misinformation is integrated into a mental schema for a topic, unless corrective information is presented simultaneously, it becomes very challenging to change the mental model later (Walter & Tukachinsky, 2020).

However, research that focuses on prebunking techniques—strategies that reduce the acceptance of misinformation before it is encountered—has proved more promising (Lewandowsky, 2020). Prebunking techniques that slow down thought processes and increase scepticism can help people more accurately identify misinformation that is encountered later (Lewandowsky et al., 2012). Simply encouraging reflection and the engagement of metacognitive skills can improve misinformation detection (Salovich & Rapp, 2020). Research has even found that prebunking techniques do not need to be specific to a single issue to be effective (sometimes called an “issue-based intervention”); “technique-based interventions” raise participants’ scepticism about misinformation techniques used across a variety of topics, improving the transferability of the intervention (Iles et al., 2021).

Generally, prebunking is thought to consist of a forewarning that what follows is false information (thereby raising scepticism), followed by a statement of the misinformation that it refutes, as well as an explanation of how the misinformation misleads (Compton, 2013). An increasingly common approach to prebunking is called misinformation “inoculation” which uses vaccination as a metaphor for increasing misinformation resistance (Basol et al, 2021; Cook et al, 2017; Iles et al, 2021). In an inoculation intervention, the participant is given a warning about

how others may try to trick them, which can activate the person's immune system defences against the misinformation persuasion tactic to come (Iles et al, 2021). Then, the participant is given an explanation of the misinformation in a weakened form in the context of the learning environment, which can help them exhibit an immune response to the same type of genuine misinformation encountered in their everyday lives (Cook et al, 2017). Inoculation can raise scepticism about information and helps to make people more alert to the quality of the information being shared.

One of the most successful prebunking interventions is the game *Bad News*, developed by Roozenbeek and van der Linden (2019a). In the game, participants try to gain as many followers as possible on their social media platform. To do this, players engage in the strategies of misinformation-creators: they fabricate information, sow doubt about credible public figures and organisations, and promote belief in conspiracy theories, all of which increases engagement with the players' posts. In the process, participants learn how misinformation is created and how appealing it can be, which raises their scepticism and improves their ability to identify misinformation (Basol, Roozenbeek, & van der Linden, 2020). The *Bad News* game is a technique-based inoculation strategy that teaches about misinformation-spreading techniques without limiting the intervention to certain topics, which gives it an advantage over ones that focus on a particular subject (Roozenbeek & van der Linden, 2019a; Roozenbeek, Maertens, McClanahan, & van der Linden, 2021). Studies of the effectiveness of the game have found that, regardless of political ideology, gender, and age group of players, this gamified approach to inoculation is associated with better detection of misinformation persuasion tactics (Roozenbeek & van der Linden, 2019a). The game has even been shown to be effective in other languages with participants from a variety of countries (Roozenbeek, van der Linden, & Nygren, 2020).

3. Methodology

This quasi-experimental study employed a pre-/post-test design to test the efficacy of a misinformation inoculation game in conferring resistance to misinformation. The independent variable consisted of either a treatment condition, in which participants (n=154) received 50 minutes of library instruction that centred on a misinformation inoculation game called *Chaos Creator* (based on the *Bad News Game* developed by Roozenbeek and van der Linden), or a control condition, in which participants (n=114) received 50 minutes of library instruction that centred on the source evaluation tool called the CRAP test (Blakeslee, 2004). The CRAP test (originally CRAAP, but modified for this study), is a commonly used source evaluation acronym that stands for "Currency, Relevance, Authority, Accuracy, and Purpose" (Blakeslee, 2004). Testing the control condition allows for better confidence that the presence of a librarian teaching about source evaluation alone does not influence differences in the dependent variable. The CRAP test was used as the subject in the control condition because it is a commonly used tool for teaching students source evaluation (LeBlanc & Quintiliano, 2015), but it does not rely on misinformation inoculation techniques. To some extent, the researcher was hoping to determine whether alternative, more evidence-based techniques for source evaluation instruction would be more effective than the CRAP test. Students in the control condition were not limited to social media sources but were encouraged to use any relevant source.

In creating the *Chaos Creator* game, the researcher modified the *Bad News Game* by presenting it as a teams-based in-person game, rather than an online, individual game. She also modified the questions to be more conducive to an in-person, teams-based teaching

environment (e.g. reducing the overall number of questions, providing fewer options for players to choose from when moving through the game, etc.). Finally, she modified the subject of some of the questions in the game to make them more relevant to timely, news-related issues and to clarify for a primarily young-adult audience (e.g. taking out references to celebrities that are less familiar to them, like Warren Buffett).

The dependent variable was an assessment of the reliability of eight fictional tweets, six associated with a misinformation technique and two serving as controls (see Appendix A and B). These tweets were developed by Roozenbeek and van Der Linden (2019a) as part of an openly licensed instrument that has been administered to over 15,000 participants. The instrument has also been analysed for reliability and item effects by Roozenbeek, Maertens, and van Der Linden (2020).

While there were two item effects identified (item set order effects and differences in psychometric properties) that can impact effect size, the researchers found that the tests were still reliable for detecting the effectiveness of misinformation inoculation approaches. For this reason, the researcher elected to use this evaluated instrument rather than creating a new one. The instrument and procedures of this study were reviewed and approved by the Cleveland State University (CSU) Institutional Review Board before data collection took place.

Following a library instruction session, participants in both the control and treatment groups were asked to indicate reliability of each tweet on a seven-point Likert-style scale. The pre- and post-test fictional tweets were different but were associated with the same set of six misinformation persuasion techniques. The tweets are fictional to avoid memory confounds (in case participants had seen any real examples selected), to exert more control over the tweet content to match the misinformation techniques being tested, and for ethical reasons (such as to avoid unintentionally spreading authentic misinformation messages) (Greene et al., 2022).

The study population consisted of first year students at a mid-sized, urban, American higher education institution in the Midwest. To sample this population, purposive sampling methods were employed. Students in 21 sections of a First Year Writing course at Cleveland State University (ENG 102) were included in the study. Students were excluded from the study if they were under the age of 18, did not complete all the questions on the pre- and post-tests, or did not indicate understanding of the informed consent form. To solicit sections for inclusion in the study, the researcher emailed all First Year Writing instructors explaining the study and soliciting volunteers. Eight instructors agreed to participate. Most instructors who volunteered taught multiple sections of the course, allowing for some consistency of instructors across the control and treatment groups (thereby reducing the effect of the instructor on the results). All sessions were taught by the researcher, who emphasised that participation in the study would not impact students' grades in the course. The entire intervention, including administration of the pre- and post-tests, took place in 50 minutes (the typical class session time period at CSU). To improve the validity of the study, sections were randomly assigned to either the treatment or control group. The minimal sample size for observing an effect is 100, and the number of valid participants in this study was 258.

The inoculation intervention consisted of a game similar to the Bad News Game described in the literature, modified to be presented synchronously and in person. The game, called Chaos Creator, incorporated the prebunking strategies of raising scepticism by describing common misinformation persuasion tactics. The prebunking activity employed a technique-based

strategy that does not address a specific issue but instead exposes students to persuasion techniques used by misinformation-generators about a wide variety of topics. The six misinformation persuasion tactics that are described and explored in the game are: impersonation, polarisation, discredit, trolling, emotion, and conspiracy.

Figure 1: Common Misinformation Persuasion Tactics.

Persuasion Tactic	Description
Impersonation	Impersonating a real person or organisation; falsely posing as a legitimate news site or as an expert in a particular subject
Emotion	Using content that deliberately plays on emotions like fear, anger, empathy, or vindication
Polarisation	Deliberately trying to divide viewers along lines of political ideology
Conspiracy	Using content that promotes the belief that events are organised by secret groups or organisations
Discredit	Attacking the person sharing a claim or argument rather than addressing the validity of the claim
Trolling	Deliberately evoking emotional responses from viewers using bait

Note. Identified tactics come from the Bad News Game developed by Jon Roozenbeek and Sander van der Linden at Cambridge University (2019a).

To play the game, students were divided into groups of 3–4 players and instructed to decide as a team what misleading tweets to share in order to gain followers in the game. Each decision was followed by a breakdown of which manipulation tactic was being employed in the example. The team with the most followers at the end of the game won. In addition to playing the game, at the beginning of the session, students were introduced to various common types of misinformation and given a chance to reflect on why these misleading messages are shared. The lesson plan and a link to the slides for the misinformation inoculation game, Chaos Creator, can be found in Appendix C.

The control group instruction consisted of a traditional source evaluation library session focusing on the use of the CRAP test. First the students were asked to identify characteristics of popular and scholarly sources and discuss their usefulness in research. Then, they were introduced to the CRAP test as a method of evaluating sources in academic settings. Students were asked to apply the CRAP test to one source of their choice in a Google Form. The lesson plan and a link to the slides for the control group instruction session can be found in Appendix D.

4. Results

The data obtained in this study was analysed using the SPSS data analysis software. In the pre-test, participants were asked about gender, age, political association, use of social media, and use of Twitter in particular. These responses were used to determine the diversity of participants

on these measures. Of particular interest were the political affiliation, social media use frequency, and Twitter use frequency of the participants. Political affiliation was fairly evenly divided, although about half of the participants indicated that their affiliation was undecided or unknown (see Table 2). Anecdotally, this question was the one students most frequently asked about when completing the pre-test; the concept of “right” and “left” political leanings was unfamiliar to many.

Figure 2: Political Affiliation for all Participants.

Political Affiliation	Response Frequency	Percent
Conservative Right	16	6.2
Moderate Right	26	10.1
Moderate Left	39	15.1
Progressive left	41	15.9
Undecided/unsure	127	49.2
Other	9	3.5
Total	258	100.0

About 53% of students indicated that they use social media multiple times per day (see figure 3), although only about 9% use Twitter at this frequency (see figure 4). About 60% of students indicated that they rarely or never use Twitter, which could impact the ability for students to apply their misinformation detection skills to social media environments that differ from Twitter.

Figure 3: Social Media Use Frequency for all Participants.

Social Media Use Frequency	Response Frequency	Percent
Never	1	0.4
Rarely	7	2.7
Monthly	0	0.0
Once per week	3	1.2
About 3 times per week	11	4.3
Daily	99	38.4
Multiple times per day	137	53.1
Total	258	100.0

Figure 4: Twitter Use Frequency for all Participants.

Twitter Use Frequency	Response Frequency	Percent
Never	101	39.1
Rarely	55	21.3
Monthly	6	2.3
Once per week	25	9.7
About 3 times per week	25	9.7
Daily	23	8.9
Multiple times per day	23	8.9
Total	258	100.0

A multiple regression analysis was also conducted to determine if political affiliation, social media use, or Twitter use were predictors of misinformation reliability scores. No significant relationship was found.

Each participant was given a unique code to input for both pre- and post-test; this allowed the researcher to compare individual responses to the pre- and post-tests in the statistical analysis while maintaining anonymity. A paired sample t-test was used to compare pre- and post-test scores for both control and treatment groups, and an independent sample t-test was used to compare the treatment and control group. The paired sample t-test was chosen because it determines the difference between two continuous variables for the same subject—in this case, the pre- and post-test scores for the control and treatment groups. The independent sample t-test was used because it is meant to compare the means of two independent groups—in this case, the post-test scores of the control and treatment groups.

The independent t-test showed that the treatment group performed significantly better on their post-tests ($p < .001$), and the mean difference between scores was 3.04 points. A Cohen's D calculation indicated that the effect size was moderate (point estimate=0.565). Levene's test for equality of variances showed that equal variances could be assumed, so the values for this assumption are used.

However, the paired sample t-test showed that, across all participants, pre-test scores ($M=42.39$, $SD=4.79$) were actually lower than post-test scores ($M=40.73$; $SD=5.835$) for tweet reliability ratings. The same was true for a paired sample t-test run for just the control group and just for the treatment group, although the decrease in post-test scores was greater for the control group ($M=38.82$, $SD=6.10$) than for the treatment group ($M=41.94$, $SD=5.36$).

Figure 5: Pre- and Post-Test Scores Mean Differences for Control and Treatment Groups.

	Control (n=114)	Treatment (n=154)	Overall (both Treatment and Control)
Pre-Test Scores Mean	42.28	42.2	42.39
Post-Test Scores Mean	38.82	41.94	40.73
Difference	3.46	0.26	1.66

Paired sample t-test were also run for each pair of pre- and post-test questions that addressed the same misinformation persuasion tactic (impersonation, emotion, polarisation, conspiracy, discredit, and trolling, as well as two controls with reliable messages). The mean scores for the control group showed that students performed worse on the post-test for impersonation, emotion, conspiracy, and trolling (all tactics except discredit and polarisation). However, when the treatment group was isolated, students performed significantly worse on the tweets related to conspiracy ($p < .001$) and trolling ($p = 0.009$), and performed better on post-test scores for all other persuasion tactics. For all tactics except conspiracy, the treatment group's post-test mean scores were higher overall than the treatment group's scores.

Figure 6: Pre- and Post-Test Item Level Scores by Misinformation Persuasion Tactic.

Persuasion Tactic	Control Group				Treatment Group			
	Mean Pre- Test Score	Mean Post- Test Score	Difference	p	Mean Pre- Test Score	Mean Post- Test Score	Difference	p
Impersonation	6.25	4.48	-1.77	0.001	5.88	6.12	0.24	0.014
Emotion	5.34	4.59	-0.75	<0.001	5.09	5.57	0.48	<0.001
Polarisation	5.58	5.66	0.08	<0.001	5.64	6.12	0.48	<0.001
Conspiracy	6.37	6.22	-0.15	0.004	6.33	6.16	-0.17	<0.001
Discredit	5.62	5.79	0.17	0.002	5.64	5.95	0.31	<0.001
Trolling	6.29	5.51	-0.78	<0.001	6.22	6.18	-0.04	0.009

When it came to the two control "reliable" tweets in the pre- and post-tests, the treatment group performed significantly better comparing the other post-test reliable tweet with the pre-test reliable tweet scores. However, students in the treatment performed very poorly on one reliable

tweet in their post-tests (see figure 7). This isolated poor score could have influenced the students' overall scores on the post-test.

Figure 7: Pre- and Post-Test Scores for Reliable/Control Tweets, Treatment Group Only.

Test	Tweet Type	Mean Score
Pre-Test	Reliable #1	4.27
Pre-Test	Reliable #2	3.23
Post-Test	Reliable #1	2.64
Post-Test	Reliable #2	3.30

Students were also asked to indicate their level of confidence in their answers at the end of the pre- and post-tests. Across both control and treatment, the mean confidence score increased from 4.67 to 5.31 from pre- to post-test. When isolating just the treatment group, their scores increased from 4.79 to 5.49, a difference of 0.7 points. The control group confidence scores increased from 4.56 to 5.10, a difference of 0.56 points. The control group's confidence scores were lower than both pre- and post-test scores of the treatment group, and the scores increased less after the implementation of the library session. However, it is important to note that the mean confidence scores of all participants increased between pre- and post-test, despite the fact that students' misinformation identification accuracy scores went down.

5. Discussion

Hypothesis 1 of this study was that the students in the treatment group would see a greater improvement from pre- to post-test scores than those in the control group. This hypothesis is supported by the data. Overall performance was worse for students who experienced the control group setting, in which they learned about the CRAP test and did not learn any misinformation identification techniques. Students in the treatment group were still overly sceptical but were more accurate in judging the reliability of the tweets overall than the students in the control group.

However, in this study, all students performed worse on the post-test Tweet reliability ratings, indicating that while they were more sceptical of false messages, they were also unnecessarily sceptical of reliable messages. This rejects Hypothesis 2, which proposes that the students in the treatment group will perform better on the post-test scores than the pre-test scores. As has been recognised previously in the literature (Hoes et al, 2023; Roozenbeek & van der Linden, 2019b), teaching students about identifying misinformation carries a significant risk of raising their overall scepticism to all messages, false or not. In addition, the confidence level scores show that students were more confident in their responses to the post-test, despite performing more poorly; this could be an indicator of the Dunning Kruger effect, in which those who have insufficient knowledge overestimate their abilities (Dunning, 2011). Perhaps with more training, the students would become more accurate in their responses, as well as less overly confident.

It is important to note that the treatment group did better at identifying misinformation persuasion tactics in the post-test tweets for all tactics except conspiracy and trolling.

Conspiracy theory thinking is associated with an overreach of scepticism into cynicism (Harford, 2021), and this excessive scepticism is reflected in the treatment group's overall scores, so it could explain the worse scores for this particular tactic. Where the treatment group's scores suffered most was in their assessment of reliable tweets, which they generally did not trust. These findings show that the treatment group is in less danger of missing false messages and in more danger of rejecting accurate ones.

The results of this study demonstrate that source evaluation instruction of any kind runs a risk of increasing students' cynicism, and it should, whenever possible, be paired with instruction discussing the reasons why some sources deserve students increased trust. When it comes to reliably identifying misinformation messages, a strategy that uses misinformation inoculation techniques is more effective than traditional, CRAP test approaches, although it still can fail without a discussion of trust in reliable messages, even if it is brief. This evidence, in addition to other studies that have criticised the CRAP test as a mechanism for teaching source evaluation, supports the conclusion that librarians should look to alternatives to the CRAP test in their teaching when possible.

The final hypothesis (Hypothesis 3) was that students in the treatment group would have a greater increase in their confidence levels after the intervention than students in the control group. This hypothesis was supported by the data, although the poor post-test scores of the treatment group make this increased confidence more troubling. If students have more confidence without improved skills, it is imperative that librarians take additional measures to balance students' increased scepticism with appropriate trust of authoritative sources.

It is also interesting to note that the librarian researchers' experience implementing the Chaos Creator game with students was very positive. Students seemed to enjoy the competitive, interactive nature of the game, and even the most lacklustre class became animated as the game went on. Some even passionately argued their decisions with the librarian, trying to increase their scores. Students in the CRAP test control group did not exhibit this level of enthusiasm.

6. Conclusion

The ubiquity of misinformation in today's world, in addition to the widespread acceptance of false claims presented by misinformation sources, has caused many to become alarmed. Among those who want to reduce the acceptance of misinformation are librarians, who have long held access to accurate information to be a foundational value. However, librarians do not seem to be adopting the misinformation-fighting strategies that have been tested and facilitated by researchers in other fields, methods such as inoculation techniques that attempt to head-off misinformation acceptance before false claims are encountered.

This study involved deploying a pre-test and post-test to a treatment group of first year students who received library instruction *with* inoculation strategies employed, as well as to a control group of first year students who received library instruction *without* inoculation strategies employed. While students who received the misinformation inoculation intervention performed worse overall after the intervention, they performed better in comparison to the control group, and they performed better for most of the misinformation persuasion tactics explored in the instruction session. Modification of the instruction to include more information about how to identify reliable information messages is advised for future uses of misinformation inoculation in

library instruction. However, the study results encourage the exploration of misinformation inoculation strategies in IL instruction, as well as continued research in this area.

There were several limitations to the generalisability and validity of this study. Although class sections were randomly assigned to the control and treatment groups, the individual students in each class were not, so there may have been aspects of the class dynamics that impacted the students' performance on the pre- and post-test. Because examples in the pre- and post-test were limited to fabricated tweets, the tests did not assess students' ability to identify misinformation in real world contexts. The fictional nature of the tweets also confused some students, who did not understand how to assess the reliability of a statement that was fake. This study also did not test the longevity of the effects observed, so whether they are long-lasting is unknown.

In future iterations of the study, collaboration with librarian researchers across multiple institutions could be used to increase the broad applicability of the study's outcomes. This would also provide an opportunity to try other variations on misinformation inoculation teaching techniques, as well as observe the impact of pairing misinformation inoculation strategies with instruction that encourages trust in reliable information sources. Additional research in this rapidly advancing area of study is encouraged.

Declarations

Ethics approval

This study was approved by the Institutional Review Board (IRB) of Cleveland State University (IRB-FY2022-240).

Funding

Not applicable.

AI-generated content

No AI tools were used.

Acknowledgements

The researcher wishes to acknowledge Marie Kennedy, Kristine Brancolini, and the members of the 2022 IRDL cohort for their invaluable support and feedback.

References

- Banks, M. (2016, December 27). [Fighting fake news](#). *American Libraries Magazine*.
- Basol, M., Roozenbeek, J., & van der Linden, S. (2020). [Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news](#). *Journal of Cognition*, 3(1).
- Batchelor, O. (2017). [Getting out the truth: The role of libraries in the fight against fake news](#). *Reference Services Review*, 45(2), 143–148.

Becker, B. W. (2016). [The librarian's information war](#). *Behavioral & Social Sciences Librarian*, 35(4), 188–191.

Blakeslee, S. (2004). [The CRAAP test](#). *LOEX Quarterly*, 31(3), 4.

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). [Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation](#). *Psychological Science*, 28(11), 1531–1546.

CILIP. (2018). [What is information literacy?](#) CILIP.

Compton, J. (2013). Inoculation theory. In J. P. Dillard, & L. Shen, *The SAGE handbook of persuasion: Developments in theory and practice*, 2nd ed, (pp. 220–237). SAGE Publishing.

Compton, J., van der Linden, S., Cook, J., & Basol, M. (2021). [Inoculation theory in the post-truth era: Extant findings and new frontiers for contested science, misinformation, and conspiracy theories](#). *Social and Personality Psychology Compass*, 15(6).

Cook, J., Lewandowsky, S., & Ecker, U. K. (2017). [Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence](#). *PloS One*, 12(5).

De Paor, S., & Heravi, B. (2020). [Information literacy and fake news: How the field of librarianship can help combat the epidemic of fake news](#). *The Journal of Academic Librarianship*, 46(5), 102218.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 44, pp. 247–296). Academic Press.

Ecker, U. K., Lewandowsky, S., Swire, B., & Chang, D. (2011). [Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction](#). *Psychonomic bulletin & review*, 18(3), 570–578.

Emery, D. (2022). [Did Betty White say she got COVID Booster 3 days before she died?](#) Snopes.com.

Eva, N., & Shea, E. (2018). [Amplify your impact: Marketing libraries in an era of “fake news”](#). *Reference & User Services Quarterly*, 57(3), 168–171.

Gibson, C., & Jacobson, T. E. (2018). [Habits of mind in an uncertain information world](#). *Reference & User Services Quarterly*, 57(3), 183–192.

Greene, C., Murphy, G., de Saint Laurent, C., Prike, T., Hegarty, K., & Ecker, U. (2022). [Best Practices for Ethical Conduct of Misinformation Research: A Scoping Review and Critical Commentary](#).

Harford, T. (2021, Mar. 6). [What conspiracy theorists don't believe](#). *The Atlantic*.

- Hoes, E., Aitken, B., Zhang, J., Gackowski, T., & Wojcieszak, M. (2023). [Prominent misinformation interventions reduce misperceptions but increase scepticism](#). *PsyArXiv*.
- Iles, I. A., Gillman, A. S., Platter, H. N., Ferrer, R. A., & Klein, W. M. (2021). [Investigating the potential of inoculation messages and self-affirmation in reducing the effects of health misinformation](#). *Science Communication*, 43(6), 768–804.
- LeBlanc, R. E., & Quintiliano, B. (2015). [Recycling C.R.A.P.](#) *Pennsylvania Libraries: Research & Practice*, 3(2), 115–121.
- Lefkowitz, M. (2017). [Library tackles fake news with workshops, resources, advice](#). Cornell University Library. Retrieved July 4th.
- Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E. Rand, D. G., Rapp, D. N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., Vraga, E. K. ... Zaragoza, M. S. (2020). [The debunking handbook 2020](#).
- Lewandowsky, S. (2020). The 'post-truth' world, misinformation, and information literacy: A perspective from cognitive science. In S. Goldstein & L. J. Hinchliffe (Eds.), *Informed societies—Why information literacy matters for citizenship, participation and democracy* (pp. 69–88). Facet Publishing.
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). [Misinformation and its correction: Continued influence and successful debiasing](#). *Psychological Science in the Public Interest*, 13(3), 106–131.
- Mackey, T. P., & Jacobson, T. E. (2011). [Reframing information literacy as a metaliteracy](#). *College & Research Libraries*, 72(1), 62–78.
- Phan, K. (2022, January 2). [Fake vaccine quote circulates after Betty White's death](#). The Associated Press.
- Rapp, D. N., & Kendeou, P. (2007). [Revising what readers know: Updating text representations during narrative comprehension](#). *Memory & Cognition*, 35(8), 2019–2032.
- Roozenbeek, J., Maertens, R., McClanahan, W., & van der Linden, S. (2021). [Disentangling item and testing effects in inoculation research on online misinformation: Solomon revisited](#). *Educational and Psychological Measurement*, 81(2), 340–362.
- Roozenbeek, J., & van der Linden, S. (2019a). [The fake news game: Actively inoculating against the risk of misinformation](#). *Journal of Risk Research*, 22(5), 570–580.
- Roozenbeek, J., & van der Linden, S. (2019b). [Fake news game confers psychological resistance against online misinformation](#). *Palgrave Communications*, 5(1), 1–10.

- Roozenbeek, J., van der Linden, S., & Nygren, T. (2020). [Prebunking interventions based on the psychological theory of “inoculation” can reduce susceptibility to misinformation across cultures.](#) *Harvard Kennedy School Misinformation Review*, 1(2).
- Sahoo, J., Sahu, S. C., & Mohanty, B. (2021). [Research on fake news an empirical analysis of selected library and information science journals.](#) *DESIDOC Journal of Library & Information Technology*, 41(4), 268–274.
- Salovich, N. A., & Rapp, D. N. (2020). [Misinformed and unaware? Metacognition and the influence of inaccurate information.](#) *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 47(4), 608–624.
- Sullivan, M. C. (2018). [Why librarians can't fight fake news.](#) *Journal of Librarianship and Information Science*, 51(4), 1146–1156.
- Sullivan, M. (2019). [Libraries and fake news: What's the problem? What's the plan?](#) *Communications in Information Literacy*, 13(1), 91–113.
- Swire, B., & Ecker, U. K. (2018). Misinformation and its correction: Cognitive mechanisms and recommendations for mass communication. In B.G. Southwell, E.A. Thorson, & L. Sheble (Eds.), *Misinformation and mass audiences*, (pp. 195–211). University of Texas Press.
- Walter, N., & Tukachinsky, R. (2020). [A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it?](#) *Communication Research*, 47(2), 155–177.

Appendix A

Pre-Test with Informed Consent

This short quiz is to establish a baseline of the knowledge you have about misinformation detection now. Please don't worry if you are unsure of the answers!

Informed Consent for Participation in a Study

You are being invited to participate in a research study being conducted by Mandi Goodsett, Performing Arts & Humanities Librarian and OER & Copyright Advisor at Cleveland State University. Please read the following information before we begin. Your participation is voluntary and you may ask questions at any time.

Study Details

The purpose of this study is to determine whether certain teaching strategies help students identify misinformation. You will be asked to take two quizzes during the class session today, one before the lesson and one after. Whether you choose to participate in the study will not impact your grade in any way. Pre- or post-quiz responses will be anonymous and information will not be shared with anyone except Mandi.

Potential Risks

Risks of participating in this study are minimal. The risks are no greater than those encountered in everyday life. It is possible you may feel uncomfortable about answering a question. This is unlikely, as the subject of the questions will be of a non-sensitive nature. However, you may ask Mandi Goodsett if you have any concerns. You may also contact the Center for International Services and Programs with any concerns at intlcenter@csuohio.edu. You received a code to participate in the study. In one document, your code is connected to your name. There is a risk that this information could be compromised. However, the researcher has tried to reduce that risk as much as possible.

Confidentiality and Privacy of Data

Access to pre- and post-test responses will be limited to the researcher. However, CSU's Institutional Review Board may review the research records. In any publication resulting from the study, no identifying information will be used. The test results will be stored on the researcher's password-protected work laptop for three years, and then destroyed.

Questions

If you have any questions about what you have read here or about your participation in the study, please contact your instructor or Mandi Goodsett at a.goodsett@csuohio.edu.

By answering the questions about yourself below and proceeding with completing the form, you consent to participate.

This instrument is based on one developed by Jon Roozenbeek, Rakoen Maertens, William McClanahan, and Sander van der Linden and licensed under a CC BY NC license.

- 1) I am at least 18 years of age.*
 - a) Yes
 - b) No

- 2) I understand that if I have any questions about my rights as a research subject, I can contact the CSU Institutional Review Board at (216) 687-3630.*
 - a) Yes
 - b) No
- 3) I have read the contents of this consent form and consent to participate in this study.*
 - a) Yes
 - b) No
- 4) Please enter the special code you received before taking this quiz:*

Demographic Questions

Most of the following questions are optional to answer but help the researcher look for patterns in student responses.

- 1) What is your gender?
 - a) Male
 - b) Female
 - c) Non-binary
 - d) Other
- 2) What is your age?
 - a) Under 18
 - b) 18 - 19
 - c) 20 - 30
 - d) 31 - 41
 - e) Over 41
- 3) What is your political affiliation?
 - a) Progressive left
 - b) Moderate left
 - c) Moderate right
 - d) Conservative right
 - e) Undecided/unsure
 - f) Other
- 4) I am a first year student at CSU.*
 - a) Yes
 - b) No
 - c) Other
- 5) How often do you use social media?
 - a) Never
 - b) Rarely

- c) Monthly
 - d) Once per week
 - e) About 3 times per week
 - f) Daily
 - g) Multiple times per day
- 6) How often do you use Twitter in particular?
- a) Never
 - b) Rarely
 - c) Monthly
 - d) Once per week
 - e) About 3 times per week
 - f) Daily
 - g) Multiple times per day
- 7) Have you had instruction about source evaluation before?
- a) Yes
 - b) No
 - c) Maybe
 - d) Other

Reviewing Tweets

Please rate the reliability of each Twitter post below on a scale from 1 to 7, with 1 meaning not at all reliable and 7 meaning very reliable.

In this study, “reliable” means the degree to which manipulation or misinformation techniques are present in the tweet. For example, a reliable tweet would have little evidence of manipulation or misinformation techniques.

PLEASE NOTE: These are all fictional posts, so please rate the reliability of each given the evidence you have in this form (not via external sources) and as if the Twitter post and its poster were real.

1. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Your Finance Guru

@GuruFinance



The Bitcoin exchange rate is being manipulated by a small group of rich bankers [#InvestigateNow](#)

3:14 PM · Jun 3, 2021

24 Retweets 15 Quote Tweets 66 Likes



2. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



US Star

@us_star_news



Brutally beaten senior desperately begs for help at hospital, is given TERRIBLE medical care only after ENORMOUS wait.

www.usstarnews.com/06042021/beaten_senior

4:50 PM · Jun 4, 2021

22 Retweets 17 Quote Tweets 34 Likes



3. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Physical fitness keeps your brain in good shape. 🏋️ 🧠

www.fitnessmag.com/05/09/2021/brain_physical_fitness

7:44 AM · May 9, 2021

12 Retweets 4 Quote Tweets 19 Likes



4. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Hey @LeoDiCaprio, it's snowing and freezing in New York. Could use some of that global warming you're always on about!

7:29 AM · Feb 4, 2021

47 Retweets 35 Quote Tweets 92 Likes



5. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Kevin Vax
@FactsVaxer22



The mainstream media has been caught in so many lies that it can't be trusted as a reliable news source
[#FakeNews](#)

8:29 PM · Jun 4, 2021

68 Retweets **56** Quote Tweets **190** Likes



6. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Jamey_D
@JDoe



Clear difference in career success between left-wing and right-wing voters #Promotion



9:44 AM · May 9, 2021

27 Retweets 19 Quote Tweets 50 Likes



7. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



American News Today

@american_news



Fed increases interest rates in an attempt to slow inflation

3:02 PM · May 11, 2021

22 Retweets 11 Quote Tweets 35 Likes



8. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Elon Musk

@Elon_Musk



Quick tip: Those in the know are all using the cryptocurrency FakeCoin. Check it out. [#yourewelcome](#)

6:00 PM · Aug 1, 2021

2 Retweets 6 Likes



9. How confident are you in your responses to the Reviewing Tweets section? (1 means not at all confident and 7 means very confident)

Thank you for completing this form and participating in the study! If you have any questions now that you have finished, please don't hesitate to contact your instructor or Mandi Goodsett at a.goodsett@csuohio.edu

Appendix B

Now that the semester is drawing to a close, this quiz is meant to explore what you have learned. If you are still unsure of your answer, don't worry about it! Just do your best with the information you received this semester.

If you have any questions about this form or your participation in the study, please contact your instructor or Mandi Goodsett at a.goodsett@csuohio.edu.

By answering the questions below and proceeding with completing the form, you consent to participate.

This instrument is based on one developed by Jon Roozenbeek, Rakoen Maertens, William McClanahan, and Sander van der Linden and licensed under a CC BY NC license.

1. Please enter the special code that you received before taking this quiz:

Reviewing Tweets

Please rate the reliability of each Twitter post below on a scale from 1 to 7, with 1 meaning not at all reliable and 7 meaning very reliable.

1. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Daily News

@dailynewsmag



BREAKING: Insurance companies are now using your phone to track your fast food consumption. #BeAware

9:50 PM · Jun 3, 2021

56 Retweets **44** Quote Tweets **72** Likes



2. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



National Inquiry Magazine

@nationalinquiry



NEWS ALERT: Baby formula linked to horrific outbreak of new, terrifying disease among helpless infants. Parents despair

www.nationalinquirymag.com/06052021/baby_formula_disease

2:50 PM · Jun 5, 2021

56 Retweets **44** Quote Tweets **72** Likes



3. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



I am a Genius Dude

@Genius_Dude2



Another shark loan for developing countries
[@WorldBank](#) [#WorldOfExtortion](#) [#HumanBanking](#)

www.geniusdudeblog.com/09152021/world_bank_loan_shark

2:43 PM · Apr 6, 2021

15 Retweets 2 Quote Tweets 25 Likes



4. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



HBÖ Channel

@hbo_channel



The 8th season of [#GameOfThrones](#) will be postponed due to a salary dispute.

6:22 AM · Jun 4, 2021

100 Retweets 65 Quote Tweets 29 Likes



5. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Science News Today

@science_news



Rare images of a trio of supermassive black holes in the act of coming together have been released.

4:50 PM · May 12, 2021

18 Retweets 1 Quote Tweet 24 Likes



6. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



Express News Weekly

@XpressNews



Report reveals scientists' underlying agendas and preferences, casting doubt on their findings.

[#QuestionScience](#)

3:29 PM · Jun 4, 2021

100 Retweets 88 Quote Tweets 312 Likes



7. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)



EFG News

@EFG_News



US President Donald Trump wants to build a wall between the United States and Mexico



9:27 AM · Jun 20, 2017

102 Retweets 67 Quote Tweets 140 Likes



8. How reliable is this tweet? (1 means not at all reliable and 7 means very reliable.)

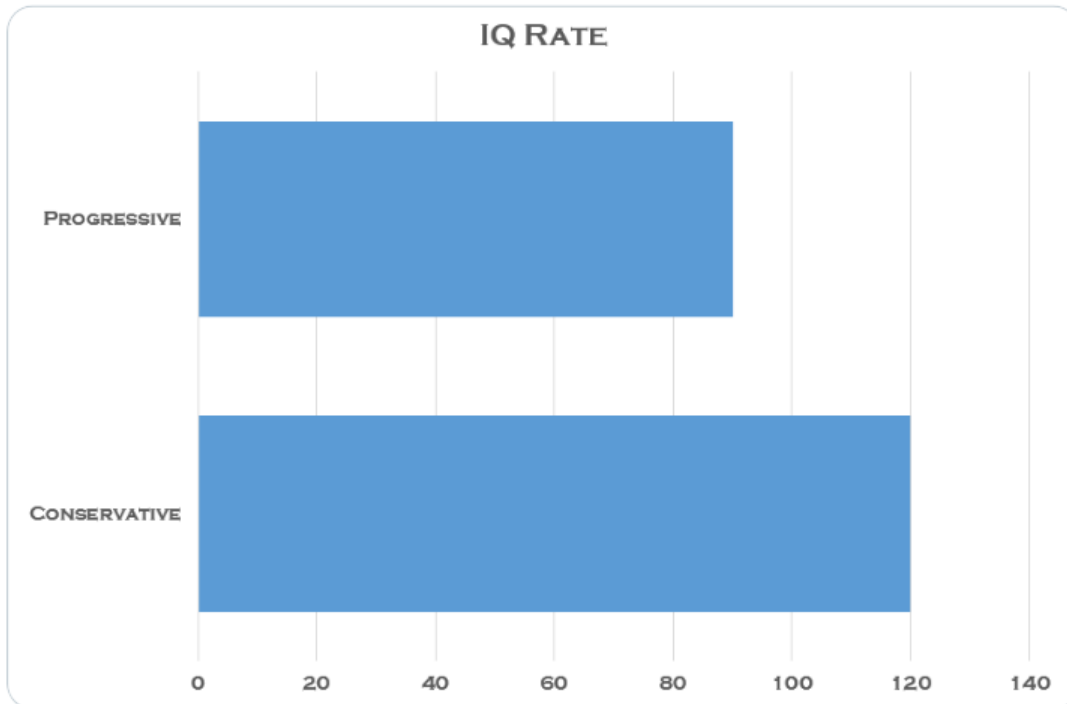


Missy Fitness World

@MissyFitness



The myth of equal IQ between left-wing and rightwing people exposed [#TruthMatters](#)



9:12 PM · Mar 8, 2021

17 Retweets 4 Quote Tweets 44 Likes



9. How confident are you in your responses to the Reviewing Tweets section?

Thank you for completing this form and participating in the study! If you have any questions now that you have finished, please don't hesitate to contact your instructor or Mandi Goodsett at a.goodsett@csuohio.edu.

Appendix C

Chaos Creator Game Lesson Plan

Introductory Information

- Target audience: undergraduate students from variety of majors in a first-year writing section
- Accompanying assignment: research prospectus
- Setting: varies
- Time: 50 minutes
- To bring: slides -> https://csuohio-my.sharepoint.com/:p/g/personal/6000463_csuhio_edu/Ec1NARA0HhJLu4A2BIhryvKBmqMLOICAB9sNMOXjknqA6Q?e=GtphqB

Learning Outcomes

Students will be able to ...

- Identify strategies used to spread misinformation online in order to avoid engaging with false or misleading messages.
- Assess the reliability of claims made by various sources in order to avoid engaging with false or misleading messages.
- Approach online claims with skepticism, while approaching information institutions (journalism, science, higher education) with measured trust.

Outline

- 1) Introduction and Welcome (10 min.)
 - a) Pre-learning quiz – 8 minutes
 - b) <https://forms.gle/gD46xm9hM6FJHEXKA>
- 2) Class Content
- 3) **Introduction to Misinformation** (10 min.)
 - a) Definitions and examples
 - b) Why do people share misinformation? What do you think?
 - c) Questions?
- 4) **Chaos Creator Game** (20 – 25 min)
 - a) Explain the rules
 - b) Break into teams
 - c) 6 challenges
 - d) Count points
 - e) Debriefing as a group
- 5) Review and Wrap-up (10 min.)
 - a) Post-learning quiz – 8 minutes
 - b) <https://forms.gle/3rSjHqoXssJEAjtSA>

Appendix D

CRAP Test Lesson Plan

Introductory Information

- Target audience: undergraduate students from variety of majors in a first-year writing section
- Accompanying assignment: research prospectus
- Setting: varies
- Time: 50 minutes
- To bring: slides -> https://csuohio-my.sharepoint.com/:p:/g/personal/6000463_csuohio_edu/EU8qgsbsnSIHnEqUoxqlfRUBjwGO58mt1uHP5hdowUMIHQ?e=CnQRqr

Learning Outcomes

Students will be able to ...

- Identify strategies used to spread misinformation online in order to avoid engaging with false or misleading messages.
- Assess the reliability of claims made by various sources in order to avoid engaging with false or misleading messages.
- Approach online claims with skepticism, while approaching information institutions (journalism, science, higher education) with measured trust.

Outline

- 1) Introduction and Welcome (10 min.)
 - a) Pre-learning quiz – 8 minutes
 - b) <https://forms.gle/qD46xm9hM6FJHEXKA>
- 2) Class Content
- 3) **Introduction to Evaluating Sources** (15 min.)
 - a) Evidence, authority, consensus
 - b) Popular vs Scholarly sources
 - i) Practice w/ 1-2 examples
 - ii) Is it popular or scholarly? How can you tell?
 - iii) Are popular sources bad? Pros/cons of popular and scholarly
- 4) **CRAP Test** (20 min)
 - a) Walk through the elements of the CRAP test
 - b) Practice time – evaluate two sources using the CRAP test
 - c) Can work by yourself or in a small group of 2-3 people
- 5) Review and Wrap-up (10 min.)
 - a) Post-learning quiz – 8 minutes
 - b) <https://forms.gle/3rSjHqoXssJEAjtSA>