Universiteit Leiden

Master Creative Intelligence & Technology

Escaping the Room, Not the Bias: Investigating the Reduction of the False Consensus Effect through Experiential and Observational Learning in an Educational Escape Room

Name: Linthe van Rooij
Student ID: s3989429
Date: 29-7-2025

1st supervisor: Maarten H. Lamers
2nd supervisor: Zane Kripe

Master's Thesis in Creative Intelligence & Technology (former Media Technology)

Leiden Institute of Advanced Computer Science
Leiden University
Einsteinweg 55
2333 CC Leiden

Escaping the Room, Not the Bias:

Investigating the Reduction of the False Consensus Effect through Experiential and Observational Learning in an Educational Escape Room

Linthe van Rooij

Master's Thesis Creative Intelligence & Technology

LIACS, Leiden University

Thesis advisors:

Maarten H. Lamers & Zane Kripe

July 2025

Abstract

The False Consensus Effect (FCE) is a social egocentric bias in which individuals believe their own position is more widely shared than those with opposing views perceive theirs to be. This leads to misperceptions, which can be addressed by increasing individuals' awareness of alternative perspectives. While this has been studied through passive forms of knowledge transmission, it has not been explored through active approaches such as experiential learning. This study investigated whether experiential and observational learning, facilitated through the execution and observation of an educational escape room, could reduce the FCE, and whether attitudes or prior engagement with the topic moderated this effect. An educational escape room was designed around the controversial topic of replacing office workers with AI. Fifty-three (recently graduated) students from Amsterdam and Leiden University participated and were assigned to either conduct the escape room, observe a recording of it, or complete only the questionnaire. GLS-based ANOVA revealed no significant effects of learning condition on the FCE. Although the FCE appeared descriptively larger in the control group compared to the other groups, this difference was not statistically significant. Interestingly, a significant FCE was found in the experiential learning condition, but not in the observational condition. Neither attitudes nor prior engagement significantly moderated the effect. Behavioural results showed high enjoyment and engagement in the learning conditions, especially in the experiential learning experience. These findings suggest that experiential learning through an educational escape room may increase engagement and reinforce the salience of participants' own attitudes, potentially intensifying, rather than reducing, the FCE. Future research could explore group-based educational escape rooms with mixed viewpoints to encourage reflection and awareness of differing perspectives.

Keywords: False Consensus Effect, Experiential learning, Observational learning, Educational Escape Room

Contents

1 Introduction

Have you ever assumed that everyone cares about AI taking over the world as much as you do? In reality, many people not only project their attitudes, beliefs and judgements onto people around them, but also overestimate how many share their positions (Marks & Miller, 1987; Ross et al., 1977). This social egocentric bias, known as the false consensus effect (FCE), appears when individuals believe their position is more widely shared with others than those with an opposing position perceive it to be (see Figure 1 for illustration). Consequently, we justify our beliefs and are more likely to abandon alternative positions (Ross et al., 1977). This can contribute to misunderstandings and conflict, as we may inaccurately assume that others share our views.

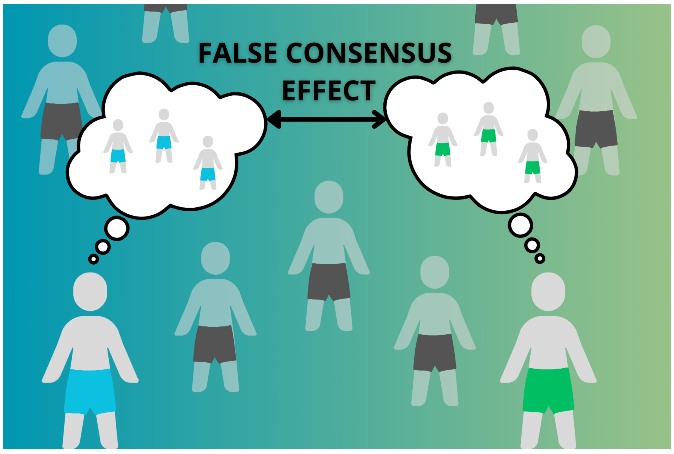


Figure 1: Illustration of the False Consensus Effect

Importantly, there is a discrepancy between the FCE and the accuracy of your perception compared to the actual prevalence of your position in society (Marks & Miller, 1987). While accuracy determines whether a belief is factually correct, it does not reveal anything about the consensus estimates surrounding your position. Thus, raising awareness about the actual distribution of beliefs is insufficient to correct systematic misperceptions (Krueger & Clement, 1994). Instead, interventions must directly target the cognitive and social mechanisms underlying the FCE to create effective change (Bauman & Geher, 2002).

Since the effect was established by Ross et al. (1977), it has been widely investigated in various research fields focusing on what contributes to this effect. Several cognitive and motivational mechanisms can account for its occurrence, such as using mental shortcuts, or heuristics, which can simplify decision-making processes (Holcomb et al., 2009; Tversky & Kahneman, 1973). More specifically, availability heuristics are used to simplify judgments based on mental information that is most easily available. These kinds of heuristics are subject to experience and stored information about events. According to the availability heuristic, people base their actions on the ease with which examples that favour a certain decision can be brought to mind. If examples of other perspectives are not strongly available, these will not be considered. Consequently, these "misperceptions" shape how we perceive and interact with those around us (Fakhri & Buchori, 2023), leading to biased decision-making with all its implications.

Moreover, seeking information that confirms your opinion seems to stimulate opinion polarisation (Li & Jager, 2023). People tend to differentiate in a society towards those who share the same opinion on a specific subject. They do this by actively avoiding information contradicting their views while seeking supportive evidence that aligns closely with their views. This becomes particularly concerning when placed in the broader context of personalised information environments. With the transformation to the digital age, the availability of "neutral" information is limited, with feeds that fuel the confirmation bias to get enough views on online material. For example, when vaccinated and unvaccinated groups absorb more information about their own opinions on the matter, they are more satisfied overall, increasing societal polarisation (Li & Jager, 2023). Bruin et al. (2020) suggests that with increased exposure to diverse other perspectives, disagreement about population estimates might be reduced, with implications for the depolarisation of society and a more balanced decision-making process.

Although existing interventions have effectively reduced the FCE (Bauman & Geher, 2002), they utilise passive forms of knowledge transmission, such as reading textual information or watching others' behaviours. The latter reflects observational learning, a traditional method of knowledge acquisition that emphasises learning through observing others' actions (Bandura & Walters, 1977). While observational learning provides people with conceptual knowledge, it lacks interactive participation. According to the constructivist approach in education, passive transmission of information is not the most optimal path to build understanding. In contrast, research shows that active participation through experiences and interactions creates better learning experiences than passive learning experiences do (Deslauriers et al., 2019; Freeman et al., 2014; Hake, 1998). Experiential learning employs this constructivist approach, emphasising engagement by combining concrete experiences, reflection on these experiences, conceptualisation and active experimentation of these concepts (Kolb, 1984). Instead of focusing merely on the information transmission used in observational learning, knowledge is constructed through direct experience, often simulating real-world settings, which enhances engagement and retention (Strickland & Kaylor, 2016). This suggests that interventions grounded in experiential learning, as a form of active learning, could provide a stronger and more engaging learning experience, contributing to the availability heuristic by making key experiences more memorable and easily retrievable, thereby encouraging lasting behaviour change, as opposed to observational learning or the absence of a learning experience.

Within the realm of Game-Based Learning, serious games have been explored as potential tools for experiential learning (Pacheco-Velázquez et al., 2023), the mitigation of cognitive biases (Barton et al., 2015), and as facilitators of behaviour change (Hammady & Arnab, 2022). Serious games are designed with learning as the primary goal while maintaining an engaging, fun, and immersive gameplay experience (Cheng et al., 2017; Grande-de-Prado et al., 2020). Their interactive nature and gamified environments foster deep cognitive engagement, which in turn promotes effective learning. Educational escape rooms are also considered a type of serious game within Game-Based Learning, where learners acquire knowledge through simulated, real-time interactive experiences centred around a specific theme (Fotaris & Mastoras, 2019). These types of escape rooms are similar to conventional escape rooms, narrative-driven and require players to solve a sequence of riddles, puzzles, and tasks under time pressure (Nicholson, 2015). They can be implemented in physical spaces, virtual environments, or even as portable setups. Educational escape rooms differ from conventional escape rooms in their target audience, which is more specifically tailored to a defined group with clear learning objectives (Grande-de-Prado et al., 2020). Their success rate tends to be higher, the puzzles are often better aligned with the content, and the time and space used is more limited. While educational escape rooms have recently gained popularity as a learning tool (Fotaris & Mastoras, 2019), it has not yet been investigated whether this form of experiential learning can serve as an effective intervention for addressing cognitive biases such as the FCE.

To address these gaps, this study examined whether knowledge acquired through an educational escape room, a form of experiential learning, leads to a greater reduction in FCE compared to knowledge gained through observational learning alone. In the next section, we will first delve further into related work regarding FCE, learning methods and escape rooms. Next, we will provide a detailed description of our testing process, followed by the results and conclusions. Finally, we discuss the implications of our research.

2 Related Work

In this section, we will discuss the underlying mechanisms that induce the FCE, explore methods that have already proven effective for reducing FCE, discuss how active learning can mitigate cognitive biases, and describe how escape rooms demonstrate potential for active learning environments.

2.1 Mechanisms underlying FCE

Several factors mediate the FCE. For example, happy emotions induce higher consensus estimates than sad emotions (Coleman, 2018), or when participants are given monetary incentives along with representative information, they tend to underestimate consensus rather than overestimate it (Engelmann & Strobel, 2000). Other contributing factors include exposure to a congruent social media feed (Luzsa & Mayr, 2019), increased self-awareness, and adopting another person's perspective (Abbate et al., 2016; Gendolla & Wicklund, 2009). Despite these mediating factors, some core cognitive and motivational mechanisms appear to play the most significant role in producing the FCE (Marks & Miller, 1987; Mullen et al., 1985; Ross et al., 1977). The underlying processes each play their role in a specific situation and operate alone or in concert. Understanding what is already known about how these effects induce the FCE is crucial for comprehending the FCE within a broader societal context. Hence, each of these processes regarding cognitive and motivational mechanisms is further explored in the following two subsections.

2.1.1 Cognitive mechanisms

The cognitive mechanisms behind the FCE suggest that this bias occurs because people rely on the availability of mental information (Bauman & Geher, 2002) and the ease of recalling this (Gershoff et al., 2008). Due to the lack of complete details on a specific issue and the ease of associating this available information with observed behaviours, people make uninformed decisions (Bauman & Geher, 2002; Fakhri & Buchori, 2023). This so-called availability heuristic posits that information is utilised based on how easily relevant instances can be recalled from memory (Tversky & Kahneman, 1973). Concerning selective exposure and engagement, when individuals are confronted minimally with opposing viewpoints, they have fewer memories of differing perspectives (Bauman & Geher, 2002). This makes it more challenging to recall alternative viewpoints, leading them to overestimate the prevalence of their beliefs and opinions, resulting in a higher FCE (Coleman, 2018).

Another main driver of the FCE reported within the cognitive domain is selective exposure (Coleman, 2018; Marks & Miller, 1987), which is the tendency to surround oneself with others who have similar views about the world, mainly to justify one's viewpoints and behaviour (Fakhri & Buchori, 2023). In addition to the FCE in general (for example, landowners allowing hunting overestimate the number of neighbours who also allow hunting compared to landowners who did not allow hunting (Casola et al., 2024)), selective exposure states that peers play a crucial role in strengthening this effect; sexually active woman estimate more sexually active woman in general than sexually non-active due to having sexually active peers as described by Bruin et al. (2020). Not only does this create a false reflection of possible perceptions, but it also reinforces the idea that their view is justified and shared by others, when in reality, it is only valid for the people they surround themselves with. Luzsa and Mayr (2021)'s research further strengthens this, suggesting that exposure to an attitudinally congruent online network is correlated with a higher FCE. When people are more willing to expose themselves to alternative viewpoints or different perspectives, it directly influences individual opinion perception, counterbalancing their assumptions about the adoption of their viewpoints in the general population, thus showing a lower FCE (Luzsa & Mayr, 2019; Wojcieszak & Price, 2009).

Other cognitive processes that influence the FCE are focus of attention and situational attribution, which are further elaborated in the review of Marks and Miller (1987).

2.1.2 Motivational mechanisms

Aside from the cognitive mechanisms playing a significant role in mediating the FCE, motivational processes also contribute to this effect. Motivational processes explain the functionality of perceived similarity (Marks & Miller, 1987). For example, it increases social support for sharing the majority's belief, justifies a specific position, or strengthens self-esteem. Perceiving a greater similarity between yourself and others is a form of social validation; people want to avoid feeling like the odd one out. They have strong desires to feel part of the majority and align their decisions in favour of a group (Gershoff et al., 2008). Inconsistency with others feels unpleasant because it threatens their fundamental need to belong (Baumeister & Leary, 1995). To avoid the negative emotions associated with this threat, they are motivated to maintain consistency with those around them (Hillman et al., 2023). Especially when people are less confident about the correctness of a particular judgment, there appears to be a greater perceived similarity between the self and others (Marks & Miller, 1987). They want to validate the rightfulness of a particular judgment or boost the perceived social support by assuming greater similarity of one's behaviour, beliefs, and judgment.

2.2 Methods reducing the FCE

The underlying mechanisms of the FCE influence how we perceive and interact with the world around us. However, when this behaviour is shaped by misrepresentations stemming from cognitive biases, it can lead to unfair decision-making and wider societal outcomes, such as increased polarisation. Therefore, we will discuss existing research introducing interventions to reduce cognitive biases that can induce the FCE.

Various approaches have been explored to reduce cognitive biases. For example, Morewedge et al. (2015) investigated whether a single training intervention, consisting of either an instructional video or a serious game, effectively reduced various cognitive biases. One of the examined biases was the projection bias, closely related to the FCE, as both involve the tendency to assume that others think and act in the same way we do. Both video and game interventions used in this study effectively reduced the projection bias, with a more significant reduction in the game intervention, compared to the control group with no intervention. They suggest that practical hands-on interventions and interactive feedback show a more substantial decrease of the tested biases. While their intervention strategies are effective for reducing these biases, earlier research has shown that education about the bias alone has proven ineffective for reducing the FCE in general (Krueger & Clement, 1994). This prompts the investigation of interventions that not only educate people about biases in general but also utilise interactive strategies, as employed in this study, to target the underlying mechanisms of the FCE.

Another study addressed one of the underlying mechanisms and successfully reduced the FCE with two intervention strategies by addressing the availability heuristic (Bauman & Geher, 2002). Therefore, we will use the reasoning of this study as a framework because it aligns closely with the purposes of our research. They did not include any educational strategies, as used in the previous studies, but achieved the reduction of the FCE by presenting participants with both sides of an issue through two different intervention strategies: a written condition and a video condition, next to a control condition with no intervention strategy (Bauman & Geher, 2002). The video condition proved particularly effective for reducing the FCE and addressing the availability heuristic, as participants observed peers engaging in a balanced debate that considered both viewpoints of a controversial statement. While both video and textual intervention strategies reduced the FCE by representing both viewpoints of a contraversial topic, the video condition reduced the FCE more in one of the topics. However, in essence, their overall approach remained observational. In contrast, prior studies by Morewedge et al. (2015) suggest that interactive and experiential learning may be a more effective strategy for addressing cognitive biases, such as the FCE.

2.3 Observational and experiential learning

Observational learning and experiential learning are both effective intervention strategies used to create awareness and acquire new knowledge. Whereas experiential learning requires more involvement and participation in general (Kolb, 1984), observational learning is achieved by observation of other people's behaviour and actions (Bandura & Walters, 1977). Here, we will discuss how existing literature utilises observational and experiential learning as effective strategies for acquiring new knowledge, understanding specific topics, and enhancing information retention.

Yoon et al. (2021) proposed that observational learning could be an effective intervention strategy for reducing the social projection bias, as discussed in the previous section. They used the same setup as the study conducted by Morewedge et al. (2015) and added another observational learning intervention (in addition to the existing instructional video and game play interventions). In this condition, participants observed other people playing the serious game. Results suggested that all intervention conditions, including the observational condition, showed a significant reduction of the social projection bias, compared to the control group with no intervention. However, the serious game play condition was still proven to be the most engaging and most effective intervention for reducing biases.

Additionally, the observational game-play condition was more effective than watching the instructional video only. While the conditions in this study have proven effective, they still only educate about the bias, rather than addressing the availability heuristics. While this study demonstrates that observational learning by watching others engage with a task can reduce the FCE, it also provides evidence that interactive, participatory methods may induce greater efficacy in reducing the FCE.

While Yoon et al. (2021)'s findings support the effectiveness of interactive learning for reducing cognitive biases, other research suggests that observational learning can be equally effective in some contexts. For instance, Bong et al. (2017) found that both active and observational participation in a simulation-based training produced comparable learning outcomes, indicating that both observational and experiential learning can facilitate knowledge acquisition but may depend on the type of learning outcome targeted.

Extending this idea, Blandin and Proteau (2000) compared how physical practice and observational learning were effective in the context of motor skill acquisition, specifically focusing on the ability to detect errors. Their results suggest that there was no significant difference in physical practice and error detection observation. In their context, similar cognitive processes are addressed in both observation and physical practice. While their study investigated skills acquisition, this study shows how both learning strategies are comparable in their underlying cognitive mechanisms.

Moreover, when comparing traditional lecture-based observational learning with active experiential learning, Specht (1985) illustrates that the mode of learning does matter in understanding management concepts. In the observation condition, participants were exposed to a lecture about management concepts, followed by a discussion. In the experiential condition, participants were required to complete an experiential learning exercise and then participate in a debate. A follow-up exam tested their knowledge retention between the two conditions. The experiential learning condition, where participants were more actively involved in knowledge acquisition, was found to be the most effective method for understanding the material, as measured by the exam questions.

Another study tested whether a role-playing form of active learning technique was an effective method to enhance knowledge about a crisis (Baglione, 2006). They did this by providing students with an exercise that simulated a realistic crisis environment in which they had to operate. This environment promoted learning and active participation; students reported that it induced more significant learning than lecturing.

Notably, Deslauriers et al. (2019) found that while active learning methods led to better actual performance, students in these conditions reported lower perceived learning and enjoyment compared to passive lecture-based learning, suggesting that perceived engagement or effectiveness does not always align with actual learning outcomes.

Overall, the literature indicates that while both learning strategies can induce desirable behavioural outcomes, experiential learning is the most promising due to its interactive nature, which enhances retention, understanding of material, and ability to mitigate other cognitive biases.

2.4 Educational escape rooms

Games are increasingly being recognised for their potential beyond pure entertainment. This has given rise to serious games, designed for serious purposes such as education and training, offering avenues for experiential learning environments. Escape rooms, particularly, are a good candidate for these learning environments with their immersive and interactive nature (Fotaris & Mastoras, 2019). Experiential learning in these contexts is illustrated by a simulated real-life experience that presents a series of challenges requiring critical thinking and problem-solving skills, thereby increasing knowledge and encouraging attitude change (Fotaris & Mastoras, 2019; Ouariachi & Wim, 2020).

Often, escape rooms are used in settings to promote teamwork and collaboration. However, as mentioned by Helbing et al. (2022), depending on the goal, either a team-based escape room or an individual escape room could be suitable. Where team-based activities could promote communication and collaboration, individual activities ensure that the individual understands every concept. Since the primary goal of the current study is not primarily focused on teamwork, this will not be taken into account. Other key advantages of using escape rooms include high levels of enjoyment and engagement, measurable learning gains, and a flexible design that can be customised to specific educational goals. Additionally, their format allows for natural intervention by instructors or researchers when needed (Fotaris & Mastoras, 2019; Kleinman & Harteveld, 2024).

The promising experiential learning environment of escape rooms is substantiated by the research of Cho et al. (2023). They highlight the importance of games in bringing an interactive and immersive experience. Escape rooms are contributing to high engagement with the topic, increasing awareness and even inducing attitude change towards misinformation. Similarly, Adams et al. (2018) provided nurse students with an interactive learning experience with an escape room designed to meet 10 educational objectives. This approach enabled students to learn through hands-on experience, facilitating critical thinking.

While a growing body of studies exists in the context of educational escape rooms, escape rooms have not been extensively investigated to mitigate cognitive biases specifically. However, a few studies investigated escape room interventions in the context of attitude change and awareness. For example, Rodriguez-Ferrer et al. (2022) demonstrated that awareness of severe mental illness can be increased by incorporating this perspective into a web-based escape room. This intervention significantly changed attitudes compared to the control group, which also performed an escape room without any awareness elements.

Moreover, Yang et al. (2023)'s findings further elaborated on the effectiveness of online game-based escape rooms to induce critical thinking and decision making in daily life, which is connected to the reasoning process and thinking through each of the challenges in the escape room to escape effectively.

Besides the effectiveness of online escape rooms in changing attitudes, physical escape rooms have also shown their potential in this regard, emphasising that the modality of escape rooms does not matter (Helbing et al., 2022). For example, Özkan Şat et al. (2025) utilised a physical escape room to investigate how escape rooms can influence students' views and attitudes regarding violence. Their results suggest that knowledge about violence was improved with escape room interventions, compared to theoretical instruction. Additionally, they indicate that attitudes were shifted and that students could identify signs of violence.

While escape rooms in education and research are still relatively new, existing studies suggest they are a promising tool for facilitating attitude change. Their immersive and interactive nature encourages active engagement and reflection, which can contribute to more meaningful cognitive processing, making them an excellent fit for an experiential learning experience. Given that attitude change is a crucial mechanism for reducing cognitive biases, escape rooms are promising for addressing the FCE. These findings highlight the potential of escape rooms as a valuable and innovative addition to the field of bias reduction.

3 Research Statement

This study builds on existing literature of the FCE, availability heuristics, experiential learning and escape rooms in education. It addresses the lack of research exploring how active learning environments, such as educational escape rooms, can serve as interventions to reduce the FCE. Thus, this study examined whether participating in an escape room (experiential learning) and observing others solve it (observational learning) reduces the FCE by addressing the availability heuristic, and compared their relative efficacy. Furthermore, it investigates whether topic attitudes and prior engagement with the topic moderate this relation. From this, the following hypotheses are proposed:

- Hypothesis 1: Experiential learning in an educational escape room reduces the FCE.
- Hypothesis 2: Observational learning through watching escape room gameplay reduces the FCE.
- Hypothesis 3: Experiential learning reduces the FCE more than observational learning.
- Hypothesis 4: Topic attitude moderates the relationship between learning type (experiential vs. observational) and the FCE.
- Hypothesis 5: Prior topic engagement moderates the relationship between learning type (experiential vs. observational) and the FCE.

4 Experimental Methods

In this section, we describe the experimental design, procedures, and measurements used in the study. We also provide a detailed explanation of the selected FCE statement and offer a comprehensive overview of the escape room setup, including its theme, narrative, puzzles, the design of the observational learning condition, and the approach to data analysis.

4.1 Experimental design

This study included two experimental learning conditions—experiential learning and observational learning—as well as a control condition without any learning manipulation. In this control condition, participants only completed a questionnaire. For the experiential learning condition, a custom educational escape room was designed and built specifically for this study. When the escape room was fully functional, a video recording was made of the entire solution process, including detailed explanations of the puzzles and the underlying thought patterns. This video formed the observational learning condition, following the video-based approach used in Bauman and Geher (2002). The control group was included to assess whether the experimental conditions resulted in a reduction in the FCE compared to when no learning technique was applied. Ethical approval for this study was obtained from the Ethics Committee for Mathematics and Natural Sciences of Leiden University.

4.2 Procedures

Participants were randomly assigned to one of the three conditions without being informed of their assigned condition, following a between-subjects experimental design. This setup allowed for a robust examination of the intervention's effect, avoiding the risk of anchoring positions. The learning conditions took place in person, while the control condition was distributed online via a link and throughout Leiden University with a QR code. Both learning conditions took place in an isolated room with minimal distractions and without the experimenter's physical presence, to enhance participant engagement and immersion and to avoid any unintentional non-verbal cues. The room had transparent walls, allowing the experimenter to monitor participants' progress during the session, while being absent from the room. Before the experiment began, participants read the information provided on the information sheet and signed an informed consent form.

Participants in the experiential learning condition received additional information after signing the informed consent form about the escape room's procedures, theme, and objectives. For those unfamiliar with escape rooms, it was explained as a sequence of puzzles leading from one to the next. The participants were instructed to solve the escape room as quickly as possible, but were informed that no time limit was enforced to ensure completion of the learning experience. Instructions included advice on organising materials, writing on the laminated paper and using hint envelopes, which could be opened after a set time or if truly stuck, to support successful completion. After the practical verbal instructions, they read an introductory text about the theme and objective. Once questions were answered, a visible timer was started, marking the beginning of the escape room experience. Details on the design and puzzles of the escape room are in section 4.5. After finishing, participants briefly reflected verbally on their experience before completing the questionnaire

In contrast, participants in the observational learning and control conditions received no additional instructions beyond the information sheet. They proceeded directly to the experimental tasks after signing the informed consent form. For the observational condition, participants watched a 30-minute recording of the escape room, after which they proceeded directly to the questionnaire. For the control condition, participants started the questionnaire immediately after signing the informed consent form.

To ensure the completion of the escape room, the researcher intervened when necessary by providing practical instructions to guide participants when they required additional support. This was only necessary in some cases and happened no more than twice per participant. Minor adjustments were made to the escape room's flow between experiments to enhance clarity. These changes enhanced the experience, making instructions and puzzle progression easier to follow without altering the core experience.

Upon completing the questionnaire, participants were fully debriefed, including an explanation of the study's purpose and a disclosure of the condition to which they had been assigned. They were then thanked for their participation and, in the learning conditions, offered a small token of appreciation in the form of sweets or coffee.

The experiment of the experiential and observational conditions took approximately 45 minutes to 1 hour to complete. This included a short briefing, the escape room activity, and the post-experiment questionnaire. The escape room lasted on average 44 minutes, followed by a 10-minute survey.

4.3 Measurements

No additional measurements were taken during the escape room or while watching the recording; all data were collected through a post-experiment questionnaire developed in Qualtrics. This included reflective questions about the escape room's materials, tasks, learning objectives (arguments), theme, and narrative. Participants also answered qualitative questions on the escape room design and reported their level of engagement during the learning condition. To assess the FCE, participants indicated their endorsement of a statement ("are you in favour of this statement") and estimated the percentage of peers who were in favour of this statement. Attitudes and prior engagement were measured on an 11-point Likert scale. The FCE was tested on three controversial issues: (1) replacing office workers with AI (linked to the learning experience), (2) introducing a meat tax, and (3) drug legalisation. The latter two served as control topics to assess the general presence of the FCE. Finally, demographic data and indicative questions on prior knowledge of the FCE and experience with AI were collected.

4.4 FCE statement: Replacing office workers with AI

Similar to the study by Bauman and Geher (2002), this study was designed based on existing literature on addressing the availability heuristic by presenting information from different viewpoints around a controversial statement. The statement used in the current study to illustrate the FCE was formulated as "Replacing office-workers with AI, assuming AI is good enough for the job?". This statement was selected after consultation with an AI expert working at the *Volkskrant* to discuss current AI topics that are considered controversial among the public. This was followed by an informal inquiry among university students to assess the perceived level of controversy surrounding the statement. The goal was to ensure a relatively balanced distribution of pro and con opinions, which is particularly important for accurately measuring the FCE.

After selecting the statement, the underlying argumentation for both sides was drawn from a combination of talking with the AI expert, querying ChatGPT, watching talk shows, the researcher's personal experience and intuition, and journal articles (Trivedi et al., 2023). The final list of arguments was narrowed down to three main areas, which served as the primary pillars for the escape room, while supporting the overall theme and narrative. These areas focused on: (1) the cost-effectiveness of AI versus human teams, (2) assessing creativity and efficiency among AI and human employees, and (3) potential client and employee satisfaction rates as side effects if AI were to be implemented.

4.5 Educational escape room design

4.5.1 General design

The purpose of the escape room was to create a playful and interactive learning environment with different arguments related to the main statement. It was designed to be performed individually to ensure exposure to all the present arguments. Similar to other escape rooms, the escape room was designed as a puzzle-sequencing quest around a specific theme and narrative. Unlike the majority of conventional escape rooms, this escape room was designed with a portable setup, with all the props fitting on a single table (Figure 2)

Figure 2: Design of the educational escape room.

4.5.2 Theme and narrative

The overarching theme of the escape room was an office environment. The underlying aligned narrative placed participants in the role of a company CEO facing a financial crisis. As they progressed through a series of puzzles, participants gathered information from various layers of the company to make an informed decision on whether to proceed with an AI transition as a possible solution to overcome financial difficulties. The general structure of the narrative, including its main objective, was provided to participants in a briefing before the event, accompanied by an "AI Transition sheet" with structured placeholders to organise their main findings. This briefing also instructed them to contact an imaginary co-CEO via the provided telephone when instructed (see bottom right of Figure 2), and especially when they had gathered all the necessary information to make an informed decision about the AI transition in the end. The full briefing can be found in Appendix A. The phone was constructed using an Arduino microcontroller with an MP3-playback component, connected to the phone's built-in speaker, microphone, and keypad (see Appendix A for a technical overview). This allowed for user interaction and enabled the playback of pre-recorded audio content (see Appendix B). The telephone played a central role in advancing the narrative, supporting the reflective aspect of the educational experience, and enhancing immersion throughout the escape room.

4.5.3 Puzzles and hints

The escape room included a combination of searching for materials, solving the puzzles and completing specific tasks. The escape room was path-based, divided into three layers that all contributed to a larger meta-puzzle (see Figure 3 for a schematic overview, as adapted from (Nicholson, 2015)). Each puzzle and layer contributed to the development of the storyline. Hints were aimed at practicalities during the experience, addressing objects, calculations and measurements, not biasing the reflective part of the design.

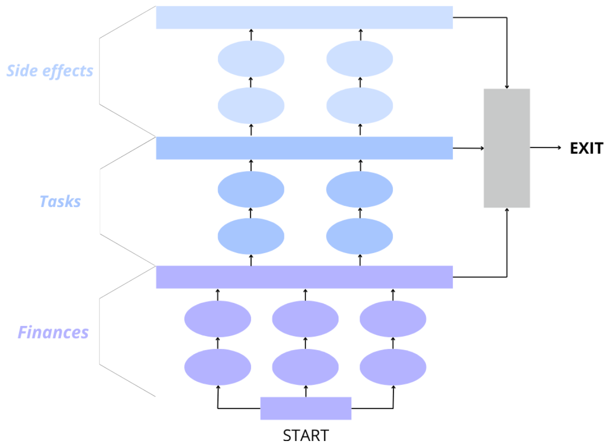


Figure 3: Schematic overview of path-based puzzle organisation.

The cost-effectiveness of AI versus human teams The first layer was defined as the finances layer, with all the materials for the puzzles placed in an office-like zipped folder (see bottom left of Figure 2). A note on top of the folder indicated that participants needed to use the telephone first, providing them with further instructions on the main objective of this layer. This included comparing the financial expenses between AI and employees for three different departments of the company: marketing, sales and support. Each department included its own puzzle, presented in the form of charts and accompanied by physical elements. These puzzles can be categorised into different types:

- Marketing: A physical puzzle involving the assembly of pieces and measurement using a ruler.
- Sales: A pattern identification task using transparent overlays that blacked out certain numbers to reveal only the relevant ones.
- Support: A magnet-based puzzle where participants used a magnet on various coins to determine which symbol would stick, corresponding to a symbol on the Y-axis of the chart, which was connected to a number.

There was no specific order to solve the three different puzzles, which also allowed them to be solved in parallel. Each of the puzzles resulted in a number, which was collected in an overview of costs. The outcome of every department's puzzle was essential for contributing to the final result of the first layer. Through simple mathematics (calculating differences and adding up), a final calculation provided a code for the briefcase lock. When the correct code was entered, the next layer was unlocked. The lock itself prevented the participants from accessing this layer directly or in parallel to the finances layer.

Assessing creativity and efficiency among AI and human employees The second layer was defined as the tasks layer. Here, all the materials were placed in the briefcase (see top right of Figure 2). A note instructed participants to use the telephone again, offering a reflection on the previous layer and providing instructions for the next one. The primary objective in this layer was to assess the suitability of two specific marketing tasks for assignment to either AI or human employees. The tasks and puzzles can be categorised into different types:

- Creative content generation: A puzzle which required comparing the number of comments on two Instagram posts. One post featuring a human-designed website and the other an AI-designed version. This was done on an iPad, which required a code to unlock. The code was hidden inside a shredder and had to be deciphered by matching colours to numbers.
- Efficiency and optimisation: A puzzle to compare the total distances of two routes, one created by a human and the other by AI. The routes were drawn on transparent sheets and had to be aligned with coloured paths on a map. Each colour corresponded to a number, which was added up and compared to obtain the most efficient route.

In this layer, the puzzles also did not require a specific order between the two tasks. Each of the puzzles resulted in a number which, combined, formed a code for accessing a small bag with a lock, which represented the final layer. Again, the lock on this bag prevented the participants from accessing it directly after the first layer.

Potential customer and employee satisfaction rates if AI were to be implemented The last layer was defined as the side-effects layer. All the materials were provided in the little bag (see top left of Figure 2). Again, a note instructed participants to use the telephone, offering a reflection on the previous layer and providing instructions for the current one. The primary objective was to assess the satisfaction rates of clients and employees resulting from the AI transition. The tasks and puzzles can be categorised into different types:

- Client satisfaction: A puzzle in deciphering a client review using two different ciphers to extract the intended message and satisfaction rate.

- Employee satisfaction: A puzzle with a red decoder to overlay on a team image when changing (partly) to AI, to reveal the change in satisfaction rate.

The puzzles can be solved in parallel again, without requiring a specific order. This layer provided the final pieces of information necessary for the decision, ultimately leading to the conclusion of the escape room.

Deciding about the AI transition As instructed in the briefing, participants contacted the provided telephone number once they completed all three layers and gathered all necessary information on their AI transition sheet. This conversation was the final meta puzzle, combining the previous layers into a cohesive whole. In this conversation, participants were asked to provide all the information by either pressing numbers or speaking it aloud, depending on the questions asked via the telephone audio (see Appendix B). The last question of the escape room was, "Based on all your findings, are we going through with the transition?" where participants had the choice to either proceed with the transition or not. The answer to this question only affected the response on the telephone; however, both choices led to a successful completion of the escape room, as illustrated by a victory sound.

4.5.4 Experiential learning objectives

The overall learning objectives included raising awareness of the different underlying arguments within each layer of the escape room. The main arguments were tied to the storyline of the escape room but could be extracted as follows: (1) overall, AI saves money, but it depends on the team within a company, if it is more cost-efficient than keeping human employees; (2) AI is very good at performing automatisation tasks efficiently, such as optimisation; (3) AI could also perform creative tasks, but humans are outperforming AI when doing creative work; (4) employee and client satisfaction could differ, when a company introduces AI. Additionally, a sub-argument also present was (5) AI misses human empathy. In the questionnaire, one question related to these learning objectives was included to assess whether participants were aware of these arguments after completing the escape room.

According to the experiential learning cycle of Kolb (1984), going through each stage of this learning cycle develops understanding, as it describes how experiences are transformed into learning. This includes the concrete experience (undergoing the experience), reflection (assessing what you have experienced), conceptualisation (making sense of the experience) and active experimentation (testing the conclusions). In this escape room, each layer embodied the concrete experience of the narrative. The phone calls were crucial for the reflective part of each layer. Conceptualisation occurred on the AI-transition sheet, where the conclusions of each layer were recorded. Active experimentation was employed in the final phone call, in which all the subconclusions were tested and integrated to inform the final decision on whether to proceed with the AI transition.

4.5.5 Pilot tests

After completing the design of the educational escape room, it underwent three pilot tests and evaluations, following iterative design principles (Eukel & Morrell, 2021). The primary objective of the pilot tests was to assess the clarity, difficulty, and timing of the escape room, as well as its alignment with the intended experiential learning cycle and the comprehensibility of the questionnaire. Based on participant feedback and observed behaviour, refinements were made to the instructions, placement of materials, and puzzle difficulty to reduce ambiguity and better align the overall design with the intended time limit. Moreover, adjustments to the flow and narrative were implemented to enhance the learning components of the experiential learning cycle and to enhance overall immersion.

4.6 Observational learning design

The observational condition included watching a recording of the previously mentioned escape room to simulate the video condition of Bauman and Geher (2002). The recording was directed at a table, focusing on the gameplay while the researcher was performing in the escape room, and narrating the process of solving the presented puzzles and associated thoughts. Afterwards, the video was edited to shorten it to 30 minutes, smooth out the storyline, and support the narrative with corresponding visual information. The visual information was present to facilitate following the narrative and puzzles while conveying the argumentation as effectively as possible. The final recording was evaluated with one pilot participant and refined for clarity, understandability and flow accordingly.

4.7 Data analysis

The data analysis consisted of several parts, each designed to test specific hypotheses, and was executed in R using output from Qualtrics. First, a two-way ANOVA was used to assess the interaction effect of endorsement and condition on peer estimates. Assumptions of normality were evaluated using the Shapiro-Wilk test and Q-Q plots, while equal variances were checked with Levene's test. Significant ANOVA results were followed up with post hoc tests and pairwise comparisons.

Second, linear regression analyses examined moderation effects by testing the interaction of condition and endorsement on peer estimates, alongside attitudes or engagement. Regression assumptions—linearity, independence of errors, homoscedasticity, and normality of residuals—were evaluated via Q-Q plots, residual inspection, influential outlier assessment, Variance Inflation Factors, and the Breusch–Pagan test.

Third, additional analyses assessed the general FCE across two other controversial issues from the questionnaire and tested whether prior knowledge of the FCE or familiarity with AI influenced the main comparisons. A two-way ANOVA, with Levene's test for variance equality, evaluated endorsement differences across conditions on peer estimates for the statements on meat tax and drug legalisation. Prior knowledge of the FCE was included as a covariate, and its influence on the main interaction was tested via an additional ANOVA.

Finally, qualitative analyses explored the learning effects of the interventions, the balance of arguments presented, and participants' engagement and enjoyment with the different learning experiences.

5 Participants

Fifty-three (recently graduated) students from Amsterdam and Leiden University participated in the study, which was conducted in June. The students were recruited either through the researcher's network or by open invitation on-site at the universities. The main criteria for participation were to be a student or a recent graduate (within 2 years) and to be fluent in English, as all instructions and recordings were provided in English. This was assessed via self-report. Prior escape room experience was not used as a criterion to keep the activity accessible to a broad target audience beyond those familiar with puzzle-solving. Nonetheless, most participants did report some familiarity with escape rooms or similar puzzle-based activities.

The participants (21 males, 29 females, and three who preferred not to disclose their gender) were aged between 18 and 35 years. The majority of the students were active in the field of Natural Sciences (47%), followed by Applied Sciences (19%) and a combination of Social Sciences, Humanities, and Engineering (34%). Gender and age were approximately evenly distributed across the three conditions (see Table 1). Group sizes were evenly distributed across experimental conditions (15 per group). Applied Sciences were slightly overrepresented in the control condition and Natural Sciences in the experiential condition (differences >5 participants or 23%). One participant did not complete the escape room and, due to insufficient progress to engage in the experiential learning experience or begin the survey, was assigned to the observational condition instead.

Participants were excluded if their endorsement and attitude scores were inconsistent (e.g., opposing a statement but showing favourable attitudes). One participant in the AI replacement analysis was excluded for this reason, as well as one participant in the drug legalisation analysis.

Demographics	Overall N = 53[1]	Experiential N = 15[1]	Observational N = 15[1]	Control N = 23[1]
Gender				
Female	29 (55%)	8 (53%)	10 (67%)	11 (48%)
Male	21 (40%)	7 (47%)	5 (33%)	9 (39%)
Prefer not to say	3 (5.7%)	0 (0%)	0 (0%)	3 (13%)
Age				
18-24 years old	30 (57%)	8 (53%)	8 (53%)	14 (61%)
25-34 years old	23 (43%)	7 (47%)	7 (47%)	9 (39%)
Student				
Recent graduate	5 (9.4%)	1 (6.7%)	1 (6.7%)	3 (13%)
University student	48 (91%)	14 (93%)	14 (93%)	20 (87%)
Field				
Applied sciences	10 (19%)	0 (0%)	2 (13%)	8 (35%)
Engeneering	6 (11%)	1 (6.7%)	2 (13%)	3 (13%)
Humanities (including Law)	4 (7.5%)	0 (0%)	3 (20%)	1 (4.3%)
Natural Sciences and Medicine	25 (47%)	12 (80%)	5 (33%)	8 (35%)
Other	3 (5.7%)	0 (0%)	2 (13%)	1 (4.3%)
Social Sciences (including Economics)	5 (9.4%)	2 (13%)	1 (6.7%)	2 (8.7%)

[1] n (%)

Table 1: Descriptive demographic information on gender, age, occupation and field, split by condition

6 Results

This research was conducted to investigate whether experiential and observational learning techniques are effective interventions for reducing the FCE and to test whether attitudinal and prior engagement effects moderate this influence. Before presenting the inferential results, the following subsection will first present the descriptive analysis of the data.

6.1 Data

Table 2 presents the number of participants in each condition and their corresponding endorsement responses. The distribution of endorsement values was uneven; most participants opposed AI replacement, while most supported drug legalisation and the introduction of a meat tax. Notably, in the drug legalisation experimental conditions, only one participant endorsed the "in favour" position in each group.

	Overall N = 53[1]	Experiential N = 15[1]	Observational N = 15[1]	Control N = 23[1]
Endorsement AI Replacement				
No	36 (68%)	9 (60%)	11 (73%)	16 (70%)
Yes	17 (32%)	6 (40%)	4 (27%)	7 (30%)
Endorsement Drug Legalisation				
No	10 (19%)	1 (6.7%)	1 (6.7%)	8 (35%)
Yes	43 (81%)	14 (93%)	14 (93%)	15 (65%)
Endorsement Meat Tax				
No	22 (42%)	4 (27%)	6 (40%)	12 (52%)
Yes	31 (58%)	11 (73%)	9 (60%)	11 (48%)

[1] n (%)

Table 2: Descriptive information absolute endorsement values split by condition for all three issues (AI replacement, drug legalisation and meat tax)

Figure 4 and Table 3 display the distribution and the descriptive statistics of peer estimates, split by endorsement value, across the different conditions. Descriptively, participants who endorsed the statement tended to report higher peer agreement than those who did not, across all conditions. These comparisons are further examined in Sections 6.2 and 6.3. Differences in SD as presented in Table 3 and the sizes of the error bars in 4 suggest that the variability in peer agreement scores varied across groups and endorsement. The overall data followed a normal distribution, and the individual group distributions were also approximately normal, as detailed in Table 3.

Peer estimates - AI replacement								
Endorsement	N	Mean	SD	SE	Shapiro-Wilk	P-value Shapiro-Wilk	Min	Max
Experiential								
No	9	21.000	14.457	4.819	0.908	0.299	4	50
Yes	6	35.000	4.472	1.826	0.853	0.167	30	40
Observational								
No	11	42.727	18.314	5.522	0.962	0.801	15	70
Yes	4	57.750	9.570	4.785	0.804	0.110	49	67
Control								
No	15	25.933	20.831	5.378	0.927	0.247	0	64
Yes	7	48.571	24.453	9.242	0.879	0.223	25	85

Table 3: Descriptive statistics of peer estimates of AI replacement, split by condition and endorsement values.

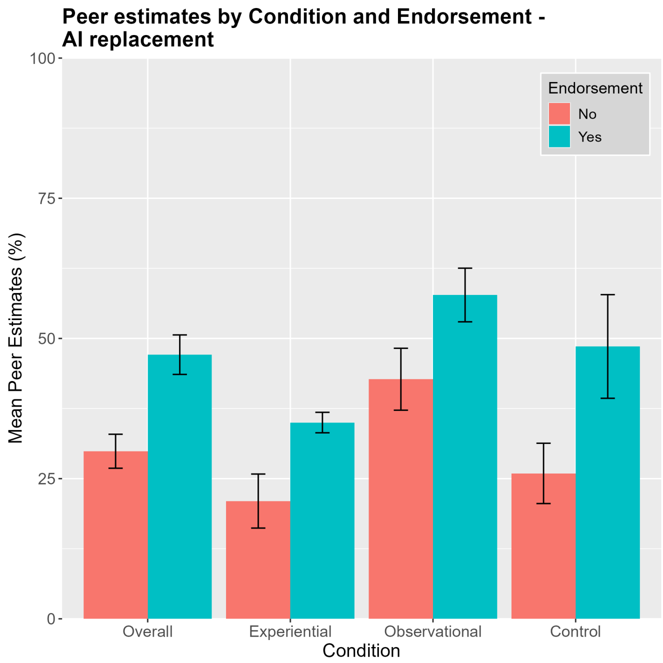


Figure 4: Distribution of descriptive data showing mean peer estimates of AI replacement with error margins across the conditions and split by endorsement value.

The descriptive statistics for the other two FCE statements, drug legalisation and meat tax, are presented in Table 4 and 5, respectively. Again, for both statements, the mean value of peer estimates tends to be higher for endorsers than non-endorsers. Table 4 illustrates the descriptive statistics of drug legalisation generalised over conditions, because of the earlier-mentioned minimal observations in the non-endorsers experimental conditions. Furthermore, in the drug legalisation condition, one outlier was identified in the endorser group using the interquartile range criteria (15% perceived peer agreement). However, this data point was retained because individual opinions can vary across the entire range of peer estimates.

Peer estimates - drug legalisation								
Endorsement	N	Mean	SD	SE	Shapiro-Wilk	P-value Shapiro-Wilk	Min	Max
No	9	40.556	21.841	7.280	0.925	0.432	5	68
Yes	43	67.093	16.492	2.515	0.946	0.041	15	100

Table 4: Descriptive statistics of peer estimates about drug legalisation split by endorsement values

The distribution of peers' estimates for the meat tax statement was generally normal (Table 5, p = .592). However, the non-endorsement group in the experiential condition showed deviation from normality (p < .05), attributable to an outlier observed in the residual plot.

Peer estimates - meat tax								
Endorsement	N	Mean	SD	SE	Shapiro-Wilk	P-value Shapiro-Wilk	Min	Max
Experiential								
No	4	28.500	31.204	15.602	0.731	0.025	8	75
Yes	11	45.000	23.664	7.135	0.942	0.546	10	80
Observational								
No	6	33.167	13.273	5.419	0.932	0.593	15	50
Yes	9	47.444	20.329	6.776	0.923	0.419	25	84
Control								
No	12	40.000	24.636	7.112	0.927	0.352	0	70
Yes	11	48.182	17.005	5.127	0.898	0.172	20	75

Table 5: Descriptive statistics of peer estimates about meat tax, split by condition and endorsement values

Furthermore, the correlation plots of peer estimates of AI replacement with moderator effects show a non-significant negative correlation with engagement, r(51) = -.21, p = .129, as displayed in Figure 6. To explore how prior topic engagement related to peer estimates across conditions, Pearson correlations were computed within each group. A moderate, negative correlation was found in the experiential condition (r = –0.49, p = .065), suggesting that participants who reported greater prior engagement with the AI topic tended to estimate lower peer estimates. A similar trend was observed in the observational condition (r = –.41, p = .134), although it was not statistically significant. The control condition showed a very weak and non-significant correlation (r = –.12, p = .591).

However, topic attitude and peer estimates of AI replacement were found to be positively correlated, r(50) = .48, p < .001, as shown in Figure 5. Within each condition, Pearson correlations were calculated between attitudes and peer estimates. Observational and control conditions showed significant positive correlations (r = 0.65, p = 0.001; r = 0.53, p < 0.05, respectively), indicating that higher attitudes towards AI are associated with higher peer estimates. The experiential condition showed a smaller, non-significant correlation (r = 0.35, p = 0.208), suggesting a weaker association.

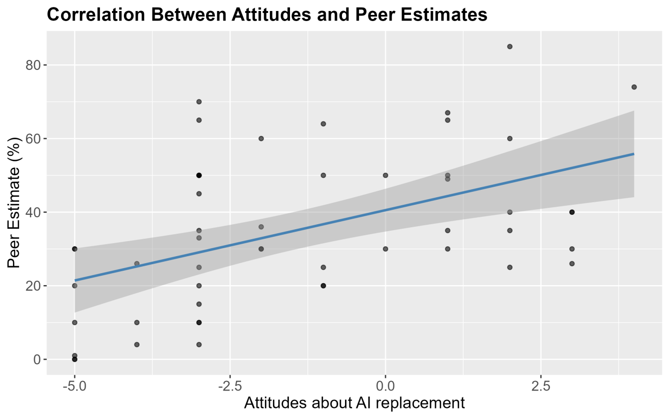


Figure 5: Distribution of attitude scores of AI replacement, measured by an 11-point Likert scale and peer estimates

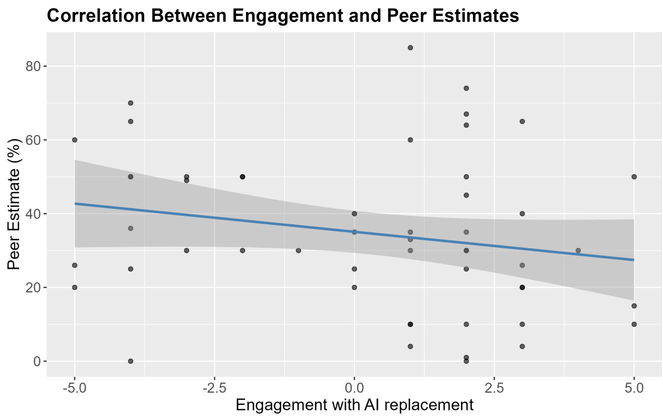


Figure 6: Distribution of prior engagement scores of AI replacement, measured by an 11-point Likert scale and peer estimates

6.2 Hypotheses 1 and 2: Experiential learning in an educational escape room and observational learning through watching escape room gameplay reduces the FCE

It was predicted that the FCE would be reduced in both the observational and experiential learning conditions, as compared to the control condition. Levene's test indicated that the assumption of equal variances was violated, $F(5, 46) = 2.45$, $p = .0476$, as also indicated by the residual plot in Figure 7. To account for this heteroscedasticity, a two-way ANOVA was conducted using Generalised Least Squares (GLS) with adjusted weighted variances across levels of condition and endorsement. This approach provided a fair and robust method, resulting in more accurate estimation while allowing for planned post-hoc pairwise comparisons.

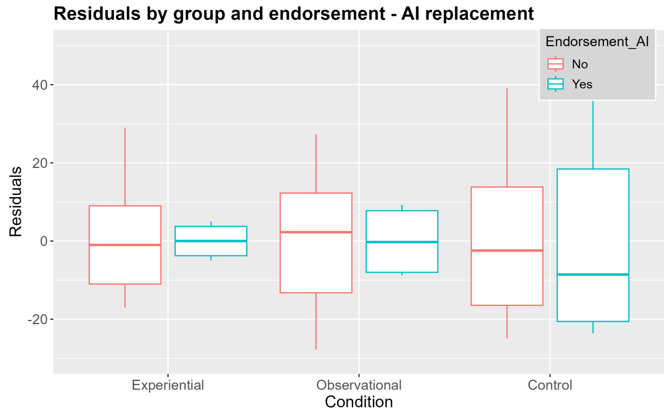


Figure 7: Residuals of peer estimates of endorsement between conditions

The ANOVA analysis revealed significant main effects of condition, $F(2, 46) = 10.70$, $p < 0.001$, and endorsement, $F(1, 46) = 15.56$, $p < 0.001$. However, the interaction between condition and endorsement was not significant, $F(2, 46) = 0.27$, $p = .7667$, indicating that the magnitude of the FCE did not significantly differ across conditions. To follow up on the significant main effect of condition, post hoc pairwise comparisons of estimated marginal means (EMMs) were conducted to assess differences in peer agreement estimates across the three learning conditions. The difference between the experiential and control conditions was not significant, $b = -9.25$, $SE = 5.94$, $t(14.8) = -1.56$, $p = .1402$. The difference between the observational and control conditions showed a marginal effect, with participants in the observational condition reporting higher estimates than those in the control, $b = 12.99$, $SE = 6.48$, $t(18.2) = 2.01$, $p = .0600$. Although the Condition × Endorsement interaction was not significant, planned pairwise comparisons were conducted to examine peer estimates of endorsement values within each condition, assessing the FCE and are shown in Figure 8. No correction was applied due to the planned comparison structure of the analysis. Although the control condition showed the largest mean difference between endorsement groups ($M = -22.6$, $SE = 10.7$), this effect did not reach statistical significance, $t(10.2) = -2.12$, $p = .60$ (Table 6). In contrast, the experiential condition reported a lower mean difference (–14.0), which was statistically significant, $t(10.1) = -2.72$, $p < .05$, indicating the presence of the FCE. In the observational condition, the difference in peer estimates between endorsement groups was non-significant, t(10.7) = –2.06, p = .065, with a mean difference of –15 (SE = 7.31), indicating the absence of the FCE. Although there are differences between endorsement groups, which, given the directionality of the difference, indicates the existence of FCE, within each condition (as shown by the planned comparison), the size of those differences does not differ significantly across conditions.

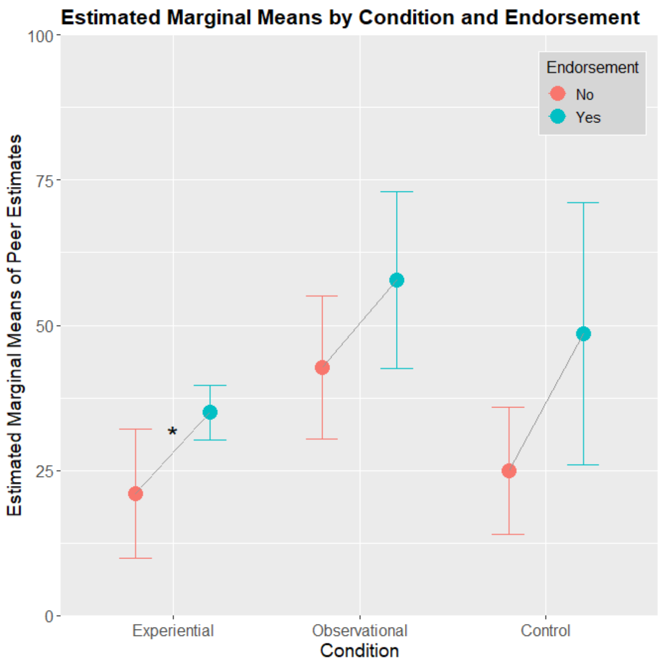


Figure 8: Results of pairwise comparison of peer estimates by endorsement across conditions. * $p < .05$

6.3 Hypothesis 3: Experiential learning reduces the FCE more than observational learning

It was expected that experiential learning would reduce the FCE more than observational learning. Given the non-significant interaction reported above between endorsement and condition, it indicates that the magnitude of the FCE did not differ statistically across the three learning conditions. The earlier-mentioned planned pairwise comparisons within each condition revealed a significant FCE in the experiential learning group (p = .0214), but only a marginal effect in the observational learning group (p = .0651). This appears to indicate that observational learning reduced the FCE more than experiential learning, contrary to the hypothesis. However, because the overall interaction was not significant, this pattern should be interpreted with caution.

Following the pairwise comparisons of the significant main effect condition, participants in the experiential condition estimated significantly lower peer agreement than those in the observational condition, b = –22.24, SE = 4.47, t(18.9) = –4.97, $p < .001$. While this supports the idea that learning conditions influence peer estimates overall, this does not provide direct statistical support for the hypothesised greater reduction in the FCE in one condition compared to another.

6.4 Hypothesis 4: Topic attitude moderates the relationship between learning type (experiential vs. observational) and the FCE

To test whether topic attitude moderated the relationship between learning condition and the FCE, a linear model including all three-way interactions among condition, endorsement, and attitude was estimated. Model assumptions were met: the studentized Breusch-Pagan test indicated no heteroskedasticity (BP = 10.00, df = 10, p = 0.441); VIFs indicated no multicollinearity (all values were approximately 1); and no influential outliers were detected. Visual checks confirmed linearity and normality of residuals.

The interaction term of experiential learning reveals a non-significant effect (t(41) = 0.11, p = 0.911). Due to a lack of variability in topic attitudes among participants in the observational/endorsers group (SD = 0), the three-way interaction term for observational learning could not be established. As a result, Hypothesis 4 could not be formally tested with the current data.

As a result, we examined two-way interactions separately with two simpler models that tested (1) the moderation of attitudes toward AI between condition and peer estimates (with endorsement as a control variable), and (2) the moderation of attitudes toward AI between AI endorsement and peer estimates (with condition as a control variable). Assumptions were tested and met. Since neither of these two-way interactions was significant, attitude toward AI was included in the final model as a control variable. The final model examined the interaction between condition and AI endorsement on peer estimates, while controlling for attitudes toward AI. The model was statistically significant, F(6, 45) = 5.44, $p < .001$, with an adjusted $R^2 = .34$. Attitudes toward AI significantly predicted peer estimates, t(45) = 2.83, p = 0.007, confirming the observed correlation between attitudes and peer estimates.

6.5 Hypothesis 5: Prior topic engagement moderates the relationship between learning and the FCE

To test whether prior topic engagement moderated the relationship between learning condition and the FCE, a linear model including all three-way interactions among condition, endorsement, and engagement was estimated. Model assumptions were met: the studentized Breusch-Pagan test indicated no heteroskedasticity (BP = 13.95, df = 11, p = .236), VIFs indicated no multicollinearity (all approximately 1), and no influential outliers were detected. Visual checks confirmed linearity and normality of residuals.

The overall model was statistically significant, F(11, 41) = 3.03, p = .005, with an adjusted R^2 = .30. However, the main effect of engagement was not significant, t(41) = -1.08, p = .286, in line with the previous observed correlation effects, nor were the three-way interaction involving engagement (ps > .75).

Additionally, a correlation model was constructed to investigate whether engagement with the AI topic is associated with overall perceived peer agreement. For this, the non-endorsers' peer estimates were inverted (i.e., subtracted from 100) to reflect the perceived peer agreement with the participant's own endorsement. The model was statistically significant, r(51) = .31, p < .05, and the results are displayed in Figure 9.

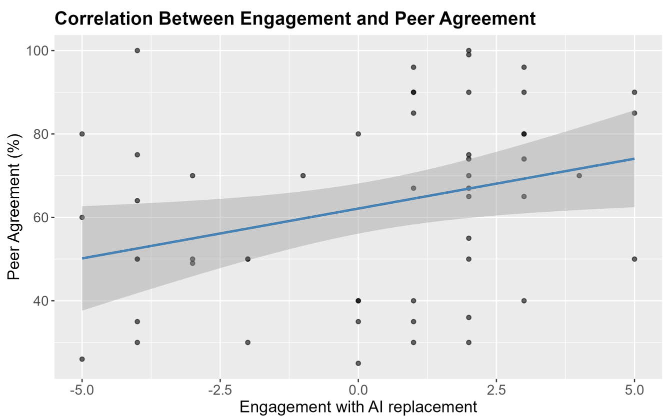


Figure 9: Distribution of prior engagement scores of AI replacement, measured by an 11-point Likert scale and peer agreement.

6.6 Supplementary inferential analyses

To assess whether the pattern of FCE is specific to AI or reflects a general FCE, we also analysed responses to two additional controversial issues: meat tax and drug legalisation. Due to limited data in some condition subgroups (n = 1, see Table 2) for drug legalisation, we did not include an interaction term for the condition variable in the model.

Similarly, the interaction term for the condition variable was also removed for meat tax, since assumption checks indicated violations of normality of the interaction residuals for the non-endorsers in the experiential condition (p < 0.05, see also Table 5 for reference). Instead, we tested for a general FCE effect across the full sample for both statements. Levene test reported the presence of equal variances, F(1, 50) = 1.50, p = .227 for drug legalisation and F(1, 51) = 0.89, p = .349 for meat tax. An additional Shapiro-Wilk test was performed to assess normality across endorsement values without the interaction term, revealing a normal distribution for both endorsers (p = .452) and non-endorsers (p = .288).

Results of ANOVA showed a significant difference between endorsement values, F(1, 50) = 17.195, p < 0.001 for drug legalisation, reflecting FCE with endorsers showing a higher mean (67.1 | SD = 16.5) than non-endorsers (40.6 | SD = 21.8). Additionally, results of ANOVA showed a marginally significant difference between endorsement values for meat tax, F(1, 51) = 3.35, p = .073, while endorsers showed a higher mean (46.8 | SD = 19.9) than non-endorsers (36.0 | SD = 22.8), see also Appendix D.

Additionally, we compared the model including the main predictors and their interaction with a model that also controlled for participants' knowledge of FCE by adding knowledge as a covariate. The likelihood ratio test showed that the main model without the covariate knowledge was a better fit, $X^2(2) = 12.92$, p < .01, indicating that prior knowledge of the FCE did not explain additional variance in the outcome. The same was done for familiarity with AI. The likelihood ratio test showed that the main model without the covariate familiarity was a better fit, $X^2(2) = 14.79$, p < .01, indicating that familiarity with AI did not explain additional variance in the outcome.

6.7 Qualitative analysis

To explore participants' perceptions of the learning experiences in more depth, we analysed the reported results of the escape room experience.

6.7.1 Engagement and enjoyment

Of all the participants who experienced the escape room, the 15 participants who executed the escape room themselves (through experiential learning) reported high engagement. When analysing these participants' open-ended questions about what they enjoyed the most, they reported various aspects of the escape room, such as the overall office design and structure, the interactive elements with the telephone, the puzzles themselves, especially the decoding puzzle, the broad range of different challenges, and the level of difficulty. One participant reported "*Figuring out how to solve the puzzles, it was also interesting seeing the different areas that AI/Human employees did better (like creativity of optimisation)*".

In contrast, among the 15 participants who watched the recording of the escape room (observational learning), eight reported engagement, while seven reported no engagement with the material. The participants who were not engaged reported that they did not like the topic of AI, escape rooms, or puzzle solving in general, that the video was too long, the lack of depth in the discussion, or that they had doubts about whether the numbers were based on reality. However, the engaged participants reported that the content was explicit and well-explained, fun to watch, they were curious about the next steps, and that the storyline was well-built and easy to follow. One participant reported "*yes, it was paced well so you could have time to predict the solution but not have to wait very long. the tasks were also fun.*".

When analysing the results from the open-ended overall feedback question, the majority of participants reported that they had a lot of fun while completing the escape room. Many described the experience as "really cool." One participant even stated, "I think this may have made me slightly more in favour of AI replacing white collar jobs," suggesting that the experience was not only enjoyable but also succeeded in conveying the intended learning component.

Additionally, participants reflected verbally on the escape room experience after completing the questionnaire. The remarks were overall very positive. Participants in the experiential condition remarked "*It was a really good one*", "*I was totally immersed*", "*I completely lost track of time*". Participants in the observational condition reported mixed comments. Some noted that the video was quite long, tedious, and too detailed, and that it felt like a lecture. However, participants also mentioned "*I actually really liked it, felt like I escaped a bit myself*" and "*I was surprised that watching an escape room execution could be this entertaining*".

6.7.2 Balanced perspective

To determine whether the escape room presented a balanced perspective on the arguments for and against the AI transition, both the participants' self-reported decisions to proceed with the transition and their responses to an open-ended question about balance were analysed. Of the participants, nine reported that they would not go through with the AI transition, while six indicated that they would. Interestingly, when comparing the decision to escape the room (whether or not going through with the AI transition) with the overall endorsement of the statement, three endorsers reported not going through with the transition, while three non-endorsers reported going through with it. These findings indicate that participants' explicit endorsement of the statement did not always align with their in-game decisions. This reinforces the idea that behaviour, especially in complex or immersive environments, is shaped not only by personal attitudes but also by context.

Furthermore, the open-ended results showed that the majority of participants in the experiential condition reported a perceived balanced perspective. Examples of reported balanced perspectives are "*Yes. It shows that there are both pros and cons to replacing office workers*", *'Yes. Both sides were represented and it was up to me to collect the info and compare'*, and *'Yes, it had you figure out by yourself what the advantages and disadvantages of AI are (if the data provided is accurate of course)'*. In contrast, some reports noted that the perspective was not balanced "*It didn't seem to have much to do with reality. I just treated the numbers as numbers*", and *'To be honest, I was focusing more on solving the puzzles than looking at the overall theme'*. In the observational condition, participants reported an overall balanced perspective. However, some participants who did not endorse the statement noted that there was a slight skewness towards AI replacement "*Yes, it was quite a balanced view. However, I did feel a slight preference towards AI*".

6.7.3 Learning components

This escape room was primarily designed to raise awareness of different perspectives on the statement "Replacing office workers with AI". To evaluate whether the intended learning outcomes as stated in section 4.5.4 were achieved, the responses containing selected arguments were further analysed.

Eight participants reported awareness of all the present arguments. The majority of participants (16) reported awareness of three out of four main arguments. Five participants reported awareness of two arguments, and only one reported awareness of one argument. Furthermore, 10 participants also reported awareness of the subargument related to AI missing human empathy.

7 Discussion

This study aimed to investigate whether experiential learning and observational learning techniques, facilitated through the execution and observation of escape rooms, reduced the FCE, compared their relative efficacy, and examined whether topic attitude and prior topic engagement moderate this relationship.

7.1 No substantial reduced FCE through learning techniques

The significant main effect for endorsement confirmed the presence of the general FCE in the statement "Replacing office workers with AI" across the whole sample. This was also the case for the drug legalisation control statement, but not for the meat tax statement (see Appendix D for a descriptive overview). This absence could indicate that factors such as topic relevance, participant engagement, or the nature of the statement itself influence whether the FCE occurs. Still, it also confirms that the nature of the AI statement is comparable to the nature of the drug legalisation statement, proving that the controversial nature of the AI statement is justified. This is further confirmed by the reported qualitative results, which indicate that the escape room provided a well-balanced perspective on this controversy.

However, the interaction between endorsement of AI replacement and learning experience was not statistically significant. In other words, the magnitude of the FCE does not differ across learning conditions. This suggests that the type of learning experience did not significantly affect the overall FCE. Although the absolute means of the pairwise comparisons between endorsement values in the control condition appear to be larger than in the learning conditions, this difference did not reach significance.

According to the findings, the results did not align with the previously stated hypotheses 1-3. It was expected that both observational and experiential learning would reduce the FCE, compared to no learning and that experiential learning would reduce the FCE significantly more than observational learning. The results showed that the FCE persisted in the experiential learning group, while there was no FCE present in the observational and control groups. This suggests that rather than reducing the tendency to overestimate peer agreement, the experiential learning intervention may have heightened participants' engagement with the topic, making their own attitudes more salient and readily available in memory. This may indicate that the intervention increased personal involvement, which in turn amplified the FCE.

Interestingly, while participants in the experiential condition reported high engagement and enjoyment, this did not translate into a reduction of the FCE. This aligns with findings from Deslauriers et al. (2019), who showed that higher engagement or enjoyment in active learning environments does not always correlate with better awareness or actual learning. Thus, participants' positive experiences may have made their own attitudes more salient without promoting reflection on alternative viewpoints, potentially reinforcing the FCE rather than reducing it.

The absence of a significant FCE in both the control and observational conditions suggests that the FCE may have been reduced in the observational condition; however, since the control condition also did not exhibit an FCE, the baseline level of the effect was already low or absent. The absence of a significant FCE in the observational condition is consistent with previous research, including studies on passive learning techniques (Bauman & Geher, 2002). However, since the FCE could not be established in the control condition of the current study, it remains unclear whether the reduced effect in the observational condition reflects an actual intervention effect.

There appears to be a greater difference in peer estimates of endorsement values within the control condition compared to the two learning conditions. How-

AI replacement							
	Endorsement				*FCE*		
	For		**Against**				**p-value**
Condition	Mean	N	Mean	N	diff	SE	p
Experiential learning	35.0	6	21.0	9	14.0	5.15	.0214*
Observational learning	57.8	4	42.7	11	15.0	7.31	.065
Control	48.6	7	25.9	15	22.6	10.70	.06

Table 6: FCE, as measured by the difference in endorsement across conditions for AI replacement. *$p < .05$

ever, the variability within the endorsement-condition groups was too large to reach significant differences between the means of the groups. Additionally, the endorsers, who are in favour of replacing office workers with AI, don't seem to overestimate the peer agreement (compared to the lower-rated peer estimates from those who are not in favour). An examination of the means suggests that the endorsers are converging toward the average peer's estimate. At the same time, non-endorsers still perceive a greater level of peer agreement (reflected in the lower peer estimates). This is not in line with the literature, stating that exposure converges both groups towards the average (Marks & Miller, 1987).

The lack of a substantial difference in interaction could also be attributed to the small sample size of the subgroups, which was particularly low in the observational, non-endorsers group. Since this design did not account for endorsement positions beforehand to obtain balanced groups, the size of the groups was dependent on the participants' responses. The reason not to account for this was that it would have resulted in non-random sampling and potentially biased responses. To account for balanced groups, the statement was carefully chosen due to its controversial nature, evaluated by an AI expert and inquiries, and by targeting a wider audience. However, targeting a wider audience could also have influenced the distribution of the FCE. As shown in Table 1, the control condition included a larger number of students from applied sciences. In contrast, the experiential condition had a higher proportion of students from natural sciences.

The results are inconsistent with those of Morewedge et al. (2015) and Yoon et al. (2021), who demonstrated that introducing video and game-based interventions can effectively reduce projection bias. One possible explanation for this discrepancy is the difference in implementation: their study included a clear pre-post design, allowing for a direct measurement of change in bias, whereas the current study relied on a between-subjects comparison without individual baseline measures. Consistent with their results is that interactive games are more engaging than observation of games.

7.2 Prior engagement and attitudes do not affect the relation of learning and FCE

It was also expected that prior engagement with the topic and attitudes would moderate the relation between learning and FCE. However, while all models were significant overall, no interaction effects were found significant. In other words, no moderation effects were established for either prior topic engagement or attitudes. The significant effect of attitude and its correlation indicate that attitude still has an impact on overall peer estimates, suggesting that a more positive attitude towards AI is associated with higher overall peer estimates. In other words, people with more positive attitudes toward AI tend to believe that their peers are also more supportive of AI, which is consistent with the overall findings on endorsement values, e.g., the FCE.

On the other hand, the lack of a significant effect of engagement suggests that prior engagement with the topic does not appear to influence how much people think others support AI (peer estimates). However, the additional positive correlation results between peer agreement and engagement were significant, indicating that a higher engagement with a topic results in a higher estimation of peer agreement. However, these results are not further supported by the interaction model's results, which include the moderation effect on peer estimates. These findings suggest that engagement influences perceived peer agreement, consistent with the availability heuristic, explaining that greater engagement with a topic makes related information more easily retrievable from memory (Marks & Miller, 1987). However, since the model did not display any significant results, we cannot conclude that engagement affects the relationship between learning and FCE.

7.3 Limitations and future research

This research also posits its limitations. The small sample size of the study, as mentioned earlier, could have contributed to inconsistencies in the results. Especially compared to other research, this sample size is significantly lower than what would have been ideal; however, due to the time constraints of a master's thesis project, it was not feasible to obtain a sufficiently large sample size.

Another drawback is that this study did not control for people's preferences in learning techniques. While the escape room experience was designed to be accessible to a broad audience, some participants found the difficulty of the escape room challenging, as well as not paying attention to the narrative because they were so focused on finding the right puzzle solutions. Additionally, some participants in the observational condition expressed a lack of interest in the subject and its presentation, or stated that they would have preferred to conduct the actual escape room, which may indicate that not everyone fully incorporated the presented information.

Moreover, successfully executing the escape room in general requires a puzzle mindset, which already excludes a part of the participants who struggle to solve it, as illustrated by one participant who was unable to complete the escape room and therefore participated in the observational learning condition instead. Although this was controlled for through guidance from the experimenter and the provision of hints, escape rooms still require a basic understanding of puzzles. As a result, the experiential learning condition may have unintentionally favoured participants with a natural affinity for problem-solving, potentially limiting the generalisability of the findings to individuals less inclined toward such tasks.

Another limitation is the lack of in-depth questions exploring the reasoning behind participants' endorsement of the AI statement. In the current design, it is crucial that the reported endorsement values accurately reflect participants' actual attitudes toward the statement. While only one participant reported incongruent information between their stated attitude and their endorsement, this part appears to be generally reliable. However, the absence of follow-up questions could therefore not explain possible variability in the data, as participants could have endorsed or rejected the statement based on different interpretations. For instance, one participant might reject the statement due to a general opposition to AI, while another might endorse it despite having concerns about certain aspects. These nuances in reasoning were not captured, which could have affected the consistency of the endorsement data and corresponding peer estimates, especially in the control group.

Future research could benefit by conducting this study with a larger sample to reduce variability within the subgroups. Addressing the limitation of asking for elaboration on endorsement could be another improvement, as well as controlling for learning style preferences. To account for the tendency to reinforce one's own attitudes, the escape room could be executed in a group of students, rather than creating an individual experience. For example, this could involve creating balanced groups of endorsers and non-endorsees participating together in an escape room. This would allow for a shared experience, making them aware of different perspectives and introducing reflective points where they must discuss the presented arguments.

8 Conclusion

Based on the availability heuristic as the underlying cognitive mechanism explaining the occurrence of the FCE, this research aimed at reducing the FCE by providing participants with alternative viewpoints around a controversial topic. This was achieved by comparing experiential and observational learning techniques within an educational escape room context, alongside a control group that did not receive any learning intervention.

The results did not support the prior hypotheses. Although the FCE appeared descriptively larger in the control group than in both learning conditions, suggesting a possible reduction of the FCE through learning interventions, this difference was not statistically significant. Additionally, participants' attitudes and prior engagement with the topic did not moderate this relationship. In other words, neither learning condition was proven to significantly reduce the FCE compared to the control, and the effect did not depend on how strongly participants engaged with the topic beforehand or held attitudes about the topic.

However, the study offers valuable insights. The presence of the FCE in the experiential learning condition suggests that, rather than reducing it, active learning may even amplify the FCE by increasing participants' focus on their own perspectives. The positive correlations between attitudes, engagement, and perceived consensus illustrate how personal beliefs and prior involvement influence perceptions of social agreement. Taken together with the high engagement and enjoyment in the escape room, this research contributes to a better understanding of addressing the FCE and highlights the value of playful learning techniques.

9 Acknowledgements

I would like to thank the following people who made it possible to carry out my thesis project. First of all, I would like to thank my supervisors, Maarten Lamers and Zane Kripe. Both individually and together, you supported me in every part. Your encouragement, enthusiasm, advice, and network helped me in conceiving, conducting, and writing up the research.

Special thanks to those who contributed to the creation of the escape room. Laurens Verhagen offered valuable advice on current controversial AI statements. Anastasia Thambwe shared her knowledge about escape rooms and helped improve the concept. Thijs van Zeijts contributed to creating the puzzles and pilot tested the recording of the escape room. I would also like to thank two other pilot testers, Joyce den Hertog and Jeroen Tiebout. Observing your performance and hearing your feedback took the escape room and its recording to the next level.

A huge thank you to my critics, Emma Boom and Dewi Becu, for reading my thesis, attending my presentation, and providing feedback.

Finally, I would like to thank all the participants for their time and effort.

References

Abbate, C. S., Boca, S., & Gendolla, G. H. (2016). Self-awareness, perspective-taking, and egocentrism. *Self and Identity*, *15*, 371–380. https://doi.org/10.1080/15298868.2015.1134638

Adams, V., Burger, S., Crawford, K., & Setter, R. (2018). Can you escape? creating an escape room to facilitate active learning. *Journal for Nurses in Professional Development*, *34*, E1–E5. https://doi.org/10.1097/NND.0000000000000433

Baglione, S. L. (2006). Role-playing a public relations crisis: Preparing business students for career challenges. *Journal of Promotion Management*, *12*, 47–61. https://doi.org/10.1300/J057V12N03_05

Bandura, A., & Walters, R. H. (1977). *Social learning theory* (Vol. 1). Prentice hall Englewood Cliffs, NJ.

Barton, M., Symborski, C., Quinn, M., Morewedge, C. K., Kassam, K. S., & Korris, J. H. (2015). The use of theory in designing a serious game for the reduction of cognitive biases. *DiGRA 2015 - Proceedings of the 2015 DiGRA International Conference: Diversity of Play*. https://doi.org/10.26503/TODIGRA.V2I3.53

Bauman, K. P., & Geher, G. (2002). We think you agree: The detrimental impact of the false consensus effect on behavior. *Current Psychology*, *21*, 293–318. https://doi.org/10.1007/S12144-002-1020-0/METRICS

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Blandin, Y., & Proteau, L. (2000). On the cognitive basis of observational learning: Development of mechanisms for the detection and correction of errors. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *53*, 846–867. https://doi.org/10.1080/713755917

Bong, C. L., Lee, S., Ng, A. S. B., Allen, J. C., Lim, E. H. L., & Vidyarthi, A. (2017). The effects of active (hot-seat) versus observer roles during simulation-based training on stress levels and non-technical performance: A randomized trial. *Advances in Simulation*, *2*, 1–13. https://doi.org/10.1186/s41077-017-0040-7

Bruin, W. B. D., Galesic, M., Parker, A. M., & Vardavas, R. (2020). The role of social circle perceptions in ”false consensus” about population statistics: Evidence from a national flu survey. *Brief Report Medical Decision Making*, *40*, 235–241. https://doi.org/10.1177/0272989X20904960

Casola, W. R., Desrochers, H. M., & Peterson, M. N. (2024). Evaluating the false consensus effect in private landowner perceptions of hunting. *Human Dimensions of Wildlife*, *29*, 441–445. https://doi.org/10.1080/10871209.2023.2247411

Cheng, M. T., Lin, Y. W., She, H. C., & Kuo, P. C. (2017). Is immersion of any value? whether, and to what extent, game immersion experience during serious gaming affects science learning. *British Journal of Educational Technology*, *48*, 246–263. https://doi.org/10.1111/BJET.12386

Cho, Y., Coward, C., Lackner, J., Windleharth, T. W., & Lee, J. H. (2023). The use of an escape room as an immersive learning environment for building resilience to misinformation. *Journal of Librarianship and Information Science*. https://doi.org/10.1177/09610006231208027

Coleman, M. D. (2018). Emotion and the false consensus effect. *Current Psychology*, *37*, 58–64. https://doi.org/10.1007/s12144-016-9489-0

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 19251–19257. https://doi.org/10.1073/pnas.1821936116

Engelmann, D., & Strobel, M. (2000). The false consensus effect disappears if representative information and monetary incentives are given. *Experimental Economics*, *3*, 241–260.

Eukel, H., & Morrell, B. (2021). Ensuring educational escape-room success: The process of designing, piloting, evaluating, redesigning, and re-evaluating educational escape rooms. *Simulation and Gaming*, *52*, 18–23. https://doi.org/10.1177/1046878120953453

Fakhri, N., & Buchori, S. (2023). False consensus effect on college students in indonesia. *Jurnal Psikologi Talenta Mahasiswa*, *2*. https://www.researchgate.net/publication/371984884

Fotaris, P., & Mastoras, T. (2019). Escape rooms for learning: A systematic review. *Proceedings of the European Conference on Games-based Learning*, 235–243. https://doi.org/10.34190/GBL.19.179

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 8410–8415. https://doi.org/10.1073/PNAS.1319030111

Gendolla, G. H., & Wicklund, R. A. (2009). Self-focused attention, perspective-taking, and false consensus. *Social Psychology*, *40*, 66–72. https://doi.org/10.1027/1864-9335.40.2.66

Gershoff, A. D., Mukherjee, A., & Mukhopadhyay, A. (2008). What’s not to like? preference asymmetry in the false consensus effect. *Journal of Consumer Research*, *35*, 119–125. https://doi.org/10.1086/524416

Grande-de-Prado, M., García-Martín, S., Baelo, R., & Abella-García, V. (2020). Edu-escape rooms. *Encyclopedia*, *1*, 12–19. https://doi.org/10.3390/ENCYCLOPEDIA1010004

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*, 64–74. https://doi.org/10.1119/1.18809

Hammady, R., & Arnab, S. (2022). Serious gaming for behaviour change: A systematic review. *Information (Switzerland)*, *13*, 142. https://doi.org/10.3390/INFO13030142/S1

Helbing, R. R., Lapka, S., Richdale, K., & Hatfield, C. L. (2022). In-person and online escape rooms for individual and team-based learning in health professions library instruction. *Journal of the Medical Library Association : JMLA*, *110*, 507. https://doi.org/10.5195/JMLA.2022.1463

Hillman, J. G., Fowlie, D. I., & MacDonald, T. K. (2023). Social verification theory: A new way to conceptualize validation, dissonance, and belonging. *Personality and Social Psychology Review*, *27*, 309–331. https://doi.org/10.1177/10888683221138384

Holcomb, T. R., Ireland, R. D., Holmes, R. M., & Hitt, M. A. (2009). Architecture of entrepreneurial learning: Exploring the link among heuristics, knowledge, and action. *Entrepreneurship: Theory and Practice*, *33*, 167–192. https://doi.org/10.1111/j.1540-6520.2008.00285.x

Kleinman, E., & Harteveld, C. (2024). The untapped potential of escape rooms as gamified research environments. *Companion Proceedings of the 2024 Annual Symposium on Computer-Human Interaction in Play*, 276–278. https://doi.org/10.1145/3665463.3678865

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice hall Englewood Cliffs, NJ.

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*, 596–610. https://doi.org/10.1037/0022-3514.67.4.596

Li, T., & Jager, W. (2023). How availability heuristic, confirmation bias and fear may drive societal polarisation: An opinion dynamics simulation of the case of covid-19 vaccination. *Journal of Artificial Societies and Social Simulation*, *26*, 2. https://doi.org/10.18564/JASSS.5135

Luzsa, R., & Mayr, S. (2019). Links between users' online social network homogeneity, ambiguity tolerance, and estimated public support for own opinions. *Cyberpsychology, Behavior, and Social Networking*, *22*, 325–329. https://doi.org/10.1089/CYBER.2018.0550

Luzsa, R., & Mayr, S. (2021). False consensus in the echo chamber: Exposure to favorably biased social media news feeds leads to increased perception of public support for own opinions. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *15*, 1. https://doi.org/10.5817/CP2021-1-3

Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin*, *102*. https://doi.org/10.1037/0033-2909.102.1.72

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences*, *2*, 129–140. https://doi.org/10.1177/2372732215600886

Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., & Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, *21*, 262–283. https://doi.org/10.1016/0022-1031(85)90020-4

Nicholson, S. (2015). Peeking behind the locked door: A survey of escape room facilities [White paper available at http://scottnicholson.com/pubs/erfacwhite.pdf].

Ouariachi, T., & Wim, E. J. (2020). Escape rooms as tools for climate change education: An exploration of initiatives. *Environmental Education Research*, *26*, 1193–1206. https://doi.org/10.1080/13504622.2020.1753659

Özkan Şat, S., Aktaş, N., & Akbaş, P. (2025). The impact of escape room game and role-playing on nursing students' violence against women assessment skills and attitudes toward violence against women: A randomised controlled trial. *Nurse Education in Practice*, *84*. https://doi.org/10.1016/j.nepr.2025.104346

Pacheco-Velázquez, E., Salinas-Navarro, D. E., & Ramírez-Montoya, M. S. (2023). Serious games and experiential learning: Options for engineering education. *International Journal of Serious Games*, *10*, 3–21. https://doi.org/10.17083/ijsg.v10i3.593

Rodriguez-Ferrer, J. M., Manzano-León, A., Cangas, A. J., & Aguilar-Parra, J. M. (2022). A web-based escape room to raise awareness about severe mental illness among university students: Randomized controlled trial. *JMIR Serious Games 2022;10(2):e34222 https://games.jmir.org/2022/2/e34222*, *10*, e34222. https://doi.org/10.2196/34222

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, *13*, 279–301. https://doi.org/10.1016/0022-1031(77)90049-X

Specht, P. H. (1985). Experiential learning-based vs. lecture-based discussion: The impact of degree of participation and student characteristics on comprehension and retention. *The Journal of Business Education*, *60*, 283–287. https://www.tandfonline.com/doi/10.1080/00219444.1985.10534983

Strickland, H. P., & Kaylor, S. K. (2016). Bringing your a-game: Educational gaming for student success. *Nurse education today*, *40*, 101–103. https://doi.org/10.1016/j.nedt.2016.02.014

Trivedi, A., Kaur, E. K., Barnwal, P., & Choudhary, C. (2023). Should ai technologies replace the human jobs? *2nd International Conference for Innovation in Technology (INOCON)*, 1–6. https://doi.org/10.1109/INOCON57975.2023.10101202.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/https://doi.org/10.1016/0010-0285(73)90033-9

Wojcieszak, M., & Price, V. (2009). What underlies the false consensus effect? how personal opinion and disagreement affect perception of public opinion. *International Journal of Public Opinion Research*, *21*. https://doi.org/10.1093/ijpor/edp001

Yang, C. L., Chang, C. Y., & Jen, H. J. (2023). Facilitating undergraduate students' problem-solving and critical thinking competence via online escape room learning. *Nurse Education in Practice*, *73*, 103828. https://doi.org/10.1016/J.NEPR.2023.103828

Yoon, H., Scopelliti, I., & Morewedge, C. K. (2021). Decision making can be improved through observational learning. *Organizational Behavior and Human Decision Processes*, *162*, 155–188. https://doi.org/10.1016/j.obhdp.2020.10.011

A Appendix A: Introduction text escape room

You are the CEO of NOVAtech, a tech company that delivers green software solutions to other businesses. Lately, the company has been facing serious financial difficulties, and you're now in the middle of a financial crisis.

Together with your co-CEO, Lyrie, you're exploring whether replacing employees with AI could be a way to prevent bankruptcy.

Over the next 30 minutes, you'll gather and review key information to help make this decision. Use the provided AI transition sheet to organize your findings as you go. You'll focus on three main areas:

1. Finances: Understand the current financial situation and potential savings from using AI.
2. Task Analysis: Look at the specific tasks employees perform and which ones could be automated by AI.
3. Side Effects: Consider the broader impacts of transitioning to AI, such as employee and client satisfaction.

In every "area", you have 2 hints available. You can read the hints if you get stuck or need some help in that area after every 4 minutes.

Once all the information is gathered, you'll discuss it with Lyrie and decide on the best course of action for NOVAtech.

You can reach her after gathering all information on these 3 points on the "AI Transition" sheet by calling * 1234# You can also reach out to her in the meantime. Good luck!

B Appendix B: Audio texts escape room

AUDIO 1 Hi, this is Lyrie, your fellow CEO at NOVAtech. To get a clear picture of the financial impact of the AI transition, I need you to compare the yearly costs of three departments: Sales, Marketing, and Customer Support, for both AI and human teams. Here's what to do: First, Identify the yearly cost of each department, separately for AI and employees. Then, Calculate the difference between the two for each department. Lastly, add up those differences, that total will give us an estimate of potential savings. Everything you need is in your folder. Thanks, I'll check in once you've submitted the final number.

AUDIO 2 Hi! Lyrie here again, thanks for the financial breakdown! It's interesting to see how replacing the Sales team with AI turns out to be much more expensive than doing the same for Customer Support. I'm guessing that's because support work, especially the chat - tends to be more repetitive and structured, while sales relies heavily on human interaction? But maybe you have different insights and opinions about this? I also noticed that marketing costs were surprisingly balanced between AI and human teams. So for the next step, let's focus on performance. I'd like you to evaluate two specific marketing tasks: creative content generation and optimization. For each, determine whether AI or humans deliver better results. Talk soon!

AUDIO 3 Quick takeaway from the marketing tasks: Human-made content still leads in creativity and emotional impact, but interestingly, the AI-built website performed surprisingly good on Instagram. How would you interpret this? And no surprise, AI clearly outperformed humans in billboard placement—faster and data-driven—but I'm unsure how it handles real-time changes like roadblocks. Finally, let's assess these side effects during the transition: employee and client satisfaction rates. After that, you'll have everything needed for a balanced AI transition report. Please complete the report with all ratings, finances, and tasks details, and call me once completed!

AUDIO 4 Hi, Lyrie speaking! I was expecting your call! I hear you have the results ready for our final decision. Start by pressing the combined total number of stars given by both clients and employees during this AI transition.

AUDIO 5 Thanks! I guess there is always room for improvement. But let's summarize your other findings! Based on your financial findings which team would be most likely to be replaceable by AI? You can say it out loud.

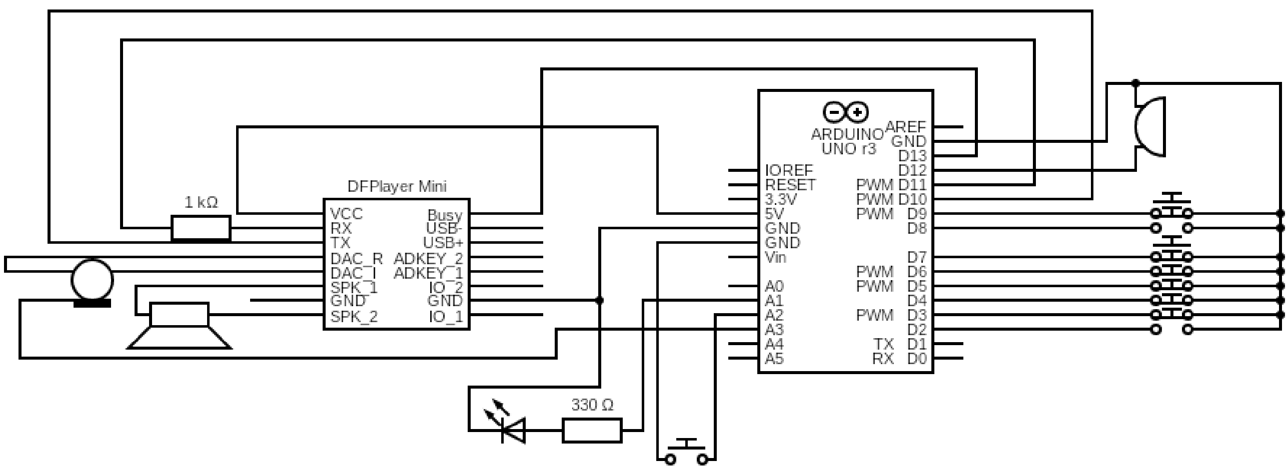
AUDIO 6 Interesting… And when we look more closely at the tasks, which one is better suited for AI?

AUDIO 7 Okay, good! So, based on all your findings, are we going through with the transition? Press 8 for yes and 9 for no.

AUDIO DECISION 7a Thank you for your thoughts! Let's go through with the transition! Thank you for your hard work in supporting this decision! You completed our quest in whether or not replacing human employees with AI is the best decision for our company. You can hang up now and go back to the researcher! Thanks again!

AUDIO DECISION 7b Thank you for your thoughts! We are not going through with the transition after all. Let's see how we can find alternatives to fix our financial difficulties, without the help of AI! Thank you for your hard work in supporting this decision! You completed our quest in whether or not replacing human employees with AI is the best decision for our company. You can hang up now and go back to the researcher! Thanks again!

C Appendix C: Schematic overview of the Arduino telephone diagram



D Appendix D: Descriptive distribution of peer estimates, split by endorsement value for both the meat tax and drug legalisation statements.

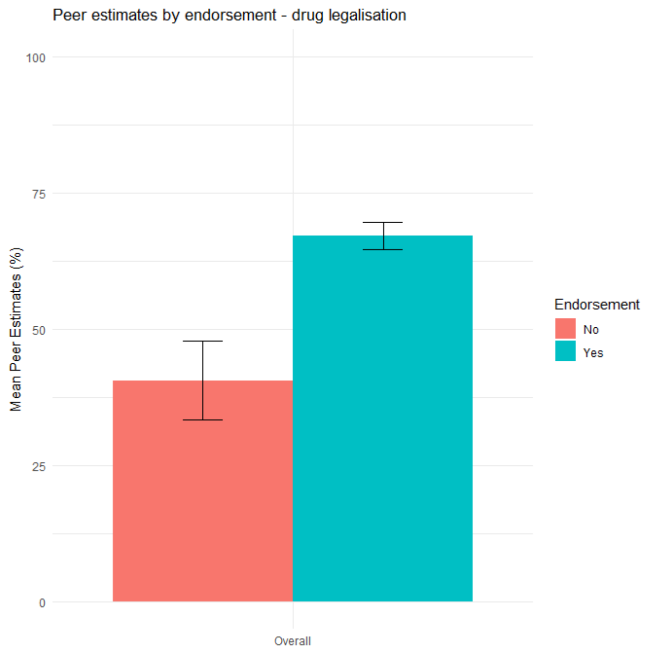


Figure 10: Distribution of peer estimates by endorsement value for drug legalisation

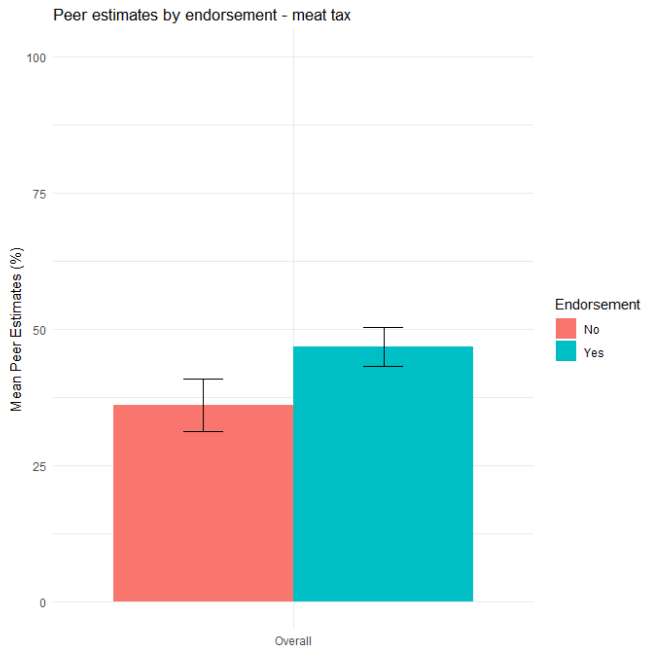


Figure 11: Distribution of peer estimates by endorsement value for meat tax