Contemporary Educational Psychology 75 (2023) 102225

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Video games and attitude change: A meta-analysis

Lukáš Kolek [a,*], Ivan Ropovik [b,c], Vít Šisler [d], Herre van Oostendorp [e], Cyril Brom [a]

[a] Faculty of Mathematics and Physics, Charles University, Czech Republic
[b] Faculty of Education, Charles University, Czech Republic
[c] Faculty of Education, University of Presov, Slovakia
[d] Faculty of Arts, Charles University, Czech Republic
[e] Faculty of Science, Utrecht University, the Netherlands

ARTICLE INFO

Keywords:
Attitudes
Video games
Meta-analysis
Game mechanics
Persuasion
Attitude change

ABSTRACT

Video games are increasingly portraying many topics that we face in our everyday lives. Yet we have only limited evidence about the way narrative games affect how we think about the topics they depict; in other words, about the link between these games and attitude change. Therefore, we conducted a meta-analysis of video games' effect on attitudinal change. The findings suggest that narrative video games affect players' attitudes towards the topics depicted in games. This effect was present in studies focused on changes in both implicit ($g = 0.36$, $k = 18$) and explicit attitudes ($g = 0.24$, $k = 101$). Longer intervention duration and game mechanics such as stereotyping and meaningful feedback resulted in larger implicit attitude change. Regarding the robustness of the underlying evidence, half of the included studies were judged to be at high risk of bias. On the other hand, the impact of publication bias in this literature was found to be negligible. Altogether, this meta-analysis provides evidence that video games shape how we think about topics they represent.

1. Introduction

Tens of thousands of new video games are released every year (Grayson, 2020; Statista, 2021a; Statista, 2021b) and just over one out of every three people on the planet is playing them (Newzoo, 2020). Increasingly, the ways in which we spend our leisure time, consume art and entertainment, and interact with people are games - or experiences that closely resemble games. (Zimmerman, 2013). Video games are becoming an important form of cultural production, showing a growing diversity of genres, cultures, and worldviews. As such, games are increasingly portraying topics that we face in our everyday lives – violence, stereotypes, conflicts, history, and more – shaping these topics' "popular representations" (Chapman, 2016). Many video games expose players to real or fictitious narratives that invite certain understandings of the depicted topics as chosen by game designers (Atkins, 2003). At the same time, games provide players with agency to interact with the represented topics; thus allowing them to challenge those representations and form their own conclusions (Pötzsch and Šisler, 2019). Yet, we have only limited evidence about the way narrative games affect how we think about the topics they depict; in other words, about the link between these games and attitude change. Despite growing empirical research on attitudes, there is currently no meta-analysis of video games' effect on attitudinal change. A previously conducted narrative review (Soekarjo and van Oostendorp, 2015) focused only on the effects of serious games and included just six studies. Five of those studies found a significant attitude change after playing the game; however, the review focused on articles examining pre-identified games. It excluded commercial games, and it reviewed games which are already 8 years old.

From the perspective of attitude change research, crucial potential lies in narrative video games. We view narrative video games as those that represent a story using various types of narrative elements and techniques (Nicklin, 2022). These narrative elements and techniques frame how the story is told (Jackson et al., 2018; Nicklin, 2022; Sýkora et al., 2021).

For instance, in the platformer game *Vectronom* (Ludopium GmbH - Arte France, 2018) one jumps as a colored cube to the beat of the music through abstract geometric shapes from point A to point B, occasionally collecting smaller cubes. This game we would not consider to be a narrative game. However, the game *Cadence of Hyrule: Crypt of the Necrodancer* (Switch, 2019) is using comparable rhythmic gameplay mechanics but adds narrative elements. For example, instead of the abstract cube, it uses a human-like character with personality; instead of abstract shapes as a game environment, it uses the *Legend of Zelda* universe full of game characters with its own history; and instead of the

* Corresponding author at: Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Prague, Czech Republic.
E-mail address: kolek@ksvi.mff.cuni.cz (L. Kolek).

https://doi.org/10.1016/j.cedpsych.2023.102225

Available online 21 September 2023

game objective to get from point A to point B, players need to save the land of Hyrule. This game we would consider to be a narrative game. The way developers operationalize and represent the main characters, their actions and their impact on the game world in *Cadence of Hyrule: Crypt of the Necrodancer* are examples of how narrative elements can carry the meaning to game elements and processes. These provided meanings can potentially affect our understanding and interpretation of displayed realities in video games. Similarly, narrative video games can also deliver attitude-related information about real-life topics they might represent and potentially connect game actions' meanings with the realities we encounter in our everyday lives.

The current historically unprecedented rapid growth of the video game market is not accompanied by sufficient empirical evidence of how these video games influence our attitudes towards the topics they depict. Therefore, the first meta-analysis is needed to map the connection between video games and their effects on changes in attitudes towards the topics they portray. By assessing 3,832 studies identified by our search operators in relevant databases, we deliver such a meta-analysis reacting to this research gap.

Using the Associative-propositional evaluation model (APE model; (Gawronski and Bodenhausen, 2014; Gawronski and Bodenhausen, 2007; Gawronski and Bodenhausen, 2006), we investigate whether video games affect explicit and implicit attitude changes in the short term and over the long term. Considering further moderators of attitude change, we focus on the effects of intervention duration, persuasive game mechanics, the effect of action vs. non-action games, and comparison to various types of control groups.

2. Theoretical Background of attitude change

As defined by (Vogel and Wanke, 2016), attitude is a "summary evaluation of an object of thought. An attitude object can be anything a person discriminates or holds in mind". The crucial characteristic of attitudes is their tendency to evaluate some attitude object with some degree of "favor or disfavor" (Eagly and Chaiken, 1993). Attitudes are an essential factor when we process complex information (Sanbonmatsu and Fazio, 1990). They influence our information selection and the way we interpret obtained information (Case and Given, 2016; Pratkanis et al., 1989; Vogel and Wanke, 2016). Therefore, they play a key role in our interpretation of the world around us. The core mechanism to change someone's attitudes is through processing information related to the attitude object (Crano and Prislin, 2008). However, when we confront information that is not aligned with our attitudes, it affects how we evaluate an information source's credibility concerning that particular topic (van Strien et al., 2016). If the information provided is inconsistent with our attitudes, we consider the source less credible and vice versa (from now on, we will call this phenomenon *credibility bias*). Perceived low credibility of the information source limits its persuasive potential.

According to the APE model, we distinguish explicit and implicit attitudes. These two forms of attitudes are guided by different, but often interplaying, processes (Gawronski and Bodenhausen, 2014; Gawronski and Bodenhausen, 2007; Gawronski and Bodenhausen, 2006). Implicit attitudes are derived from associative evaluations; that is, immediate affective responses to the object. Associative evaluations are based on the object's familiarity with other concepts in our memory, so the APE model assumes the existence of a mental structure containing these mental associations in the long-term memory. This mental structure can be changed by the co-occurrence of two concepts in one's environment resulting in either strengthening the associative link between these concepts or in creating a new associative link between them. For example, exposing participants to pictures of healthy, fit people eating a plant-based diet and to pictures of obese, unhealthy people eating an animal-based diet prior to an implicit attitude measurement towards different diets will result in temporarily more favorable attitudes towards a plant-based diet (see, e.g., Banaji and Greenwald, 2013). Seeing a plant-based diet regularly associated with positive perks in one's environment will result in the creation of a more permanent associative link between them in the long-term memory. Then, we can expect positive evaluation of the plant-based diet in implicit attitude measurements even without prior exposure to any stimuli: such as the pictures already mentioned. Implicit attitudes are assessed using response time measures, as for example Implicit Association Test in which participants place words into predefined categories as quickly as possible, revealing their different associations to two concepts in the test (for more details, see (Greenwald et al., 1998). One's associative evaluations function independently in relation to what one consciously considers to be the truth.

On the other hand, explicit attitudes are based on propositional reasoning; that is, the logical conclusions derived from information related to the object in question. To change one's propositional reasoning about an object, one must be exposed to information which is not in line with one's current beliefs and/or with what one considers to be the truth. One cannot have two contradictory propositional reasonings about the same topic as this would create cognitive dissonance and then need to be resolved. In such a case, one can either reject one of the propositions or seek additional information to resolve the cognitive dissonance created and the consistency of one's beliefs (Festinger and Lindzey, 1958). Explicit attitudes are very often assessed using self-reported questionnaires. The measurement of these changes is, however, limited to one's willingness or ability to share their explicit attitudes.

2.1. Narrative video games and attitude change

As has already been said, we consider narrative video games to be those representing a story using narrative elements and techniques framing how the story is told (Jackson et al., 2018; Nicklin, 2022; Sýkora et al., 2021). These games represent the actions in the game with a corresponding premise, thus some specific setting providing meaning to those actions. Narrative elements and techniques might introduce the exposition of the story and how it is represented; for example by defining the time and place; introducing various characters, their backgrounds, motivations, and relationships; by exposing history of the world, its problems, conflicts and plots; by defining what is told or not told through the story or through other means; or by using the narrator, emotional arcs, different pacing or perspectives of the story (Fullerton, 2018; Nicklin, 2022).

For instance, a narrative game can portray an electric car (compared to a car with an internal combustion engine) as a much more efficient and faster solution within the game. This can be done, for example, by differential framing by game characters, or by the defining parameters of these cars in the depicted game world. On the other hand, the example of the non-narrative game can be the already described game *Vectronom* (GmbH and Arte France, 2018) about colored cubes jumping to the beat in the abstract environment with no story elements.

Narrative video games are often perceived as a source of entertainment, but they can also serve as a source of information. These games require players to seek and process information in the game narratives to proceed further in the game or to fulfill game objectives. Therefore, narrative video games are information systems, in which players have the opportunity to react to the depicted information, experience the results of their actions, and respond to the changes caused in the game world (Kolek, Ropovik, Sisler, van Oostendorp, & Brom, 2021; Smethurst & Craps, 2014). As such, narrative video games can deliver messages about any real-life related topics and frame them in a certain way potentially affecting our understanding and interpretation of these topics.

2.2. Persuasive game mechanics

Changes in implicit and explicit attitudes depend on different, and often interconnected, processes. Therefore, we expect that the different ways narrative video games frame information in their stories about particular real-life topics, i.e., their persuasive game mechanics, will affect explicit and implicit attitudes differently. Persuasive game mechanics can affect attitudes towards the measured phenomena regardless of the game designers' intentions. If the game contained multiple persuasive game mechanics, we selected the dominant one from those described below. If the dominant game mechanic was not clear, we chose the category Other.

2.2.1. Stereotyping

Stereotyping is a persuasive game mechanic that constructs an implicit connection between some particular group of persons depicted in the narrative game elements and some stereotype trait assigned to that group (e.g., depiction of stereotypically represented Arabs as terrorists in an action game). This connection is not necessarily essential for progress in the game, but it is ubiquitous during gameplay. Also, the mechanic does not necessarily provide complex information about this connection. As such, we assume this mechanic will affect implicit rather than explicit attitudes.

2.2.2. Meaningful feedback

As we define it, Meaningful feedback is a persuasive game mechanic that is directly related to progression in the narrative game. When players perform acts related to the measured phenomenon depicted in the narrative game, they are rewarded by a positive or negative outcome. This creates the connection between the act and its outcome value. Meaningful feedback can be implemented in various ways: including a simulation through an economical model depicting how something works; through representation of game rules and processes (Procedural Rhetoric; Bogost, 2010); through some form of a reward system linking some actions with positive rewards or to penalties; or through a combination thereof. For instance, if the narrative game simulates a child picking cotton in a field as a tedious, hard and frustrating activity with a little reward, it creates a connection between child labor and these negative feelings in the gameplay. Meaningful feedback can affect either explicit or implicit attitudes as it can provide complex information about the outcome of some actions. It can also create frequent connections between measured phenomenon and some particular concepts that have positive or negative value for players.

2.2.3. Perspective-taking

Perspective-taking is a persuasive mechanic that provides players with a complex take on the measured phenomena from multiple points of view. Therefore, it introduces arguments about the topic: often from complementary or contradicting perspectives (e.g., providing players with interpretations of the Syrian conflict from the perspectives of all parties involved). Some arguments are in favor of, and some are against, players' initial explicit or implicit attitudes, but they all come from the one source. As such, they can possibly mitigate the credibility bias (see Section Theoretical Background of Attitude Change). We expect that perspective-taking will affect explicit rather than implicit attitudes.

2.2.4. Other

This category was used in cases where it was not possible to define one dominant game mechanic in the game from those previously described or in cases where different or unclear game mechanics were used.

2.2.5. Game genres

For the sake of the present meta-analysis, we contrast just two categories of game genres: action games and non-action games. The reason for this split is as follows: Action games often put players in time pressure situations requiring their imminent reactions to proceed further in the game compared to the non-action games which do not. Also, success in action games relies more on hand-eye coordination and motor skills to control the game precisely in order to react to imminent stimuli in limited periods of time. We consider games that do not fit these characteristics to be non-action games. There is potentially a difference between narrative action games and narrative non-action games in relation to attitudes. On the one hand, in narrative non-action games, players can have more time to perceive complex information about the depicted topics; and this with less distractions. On the other hand, narrative action games can create direct connections between the depicted topics in these games and some concepts of positive or negative value despite potential time pressure. The forming of these connections does not require transmission of complex information to happen. Therefore, narrative non-action games should be more suitable to affect explicit attitudes and narrative action games should be more suitable to affect implicit attitudes.

3. Hypotheses

H1. *Narrative Video Games Induce a Change in Players' Explicit (H1a) and Implicit Attitudes (H1b).*

Narrative video games can deliver messages about depicted topics and affect players' attitudes towards these topics within their persuasive game mechanics. They also allow players to interact with these depicted topics and, by doing so, let players challenge their own evaluations of the topics. There are several studies suggesting short-term and long-term effects of narrative video games on attitude evaluations (e.g., Kampf, 2015, 2016; Kolek, Ropovik, Sisler, van Oostendorp, & Brom, 2021). Therefore, we assume that, on a general level, narrative video games will significantly affect explicit and implicit attitudes.

Next, we examine the following moderating effects.

3.1. Intervention duration

H2. *Duration of Intervention is Positively Related to the Magnitude of Explicit (H2a) and Implicit (H2b) Attitude Change.*

Longer duration of an intervention results in a longer period of time during which participants are exposed to information related to the measured phenomenon within the game, i.e., how long they are exposed to the game's persuasive mechanics. On the most general level, to change someone's attitude means to provide this person with information in any form that will challenge their initial attitude. A study by (Maier and Richter, 2013) suggests that even attitude-inconsistent information will not be ignored (as assumed by the cognitive dissonance theory (Festinger and Lindzey, 1958), rather it will affect our attitudes less than attitude-consistent information. Therefore, we assume that the longer intervention duration will affect players' attitudes (both implicit and explicit) more compared to the short intervention durations.

3.2. Type of Comparator

H3. *The Magnitude of Explicit Attitude Change is Smaller in Studies Using Control Groups with Topic-related Activities than It Is in Studies Using Control Groups with Topic-unrelated Activities.*

We do expect that narrative video games have the potential to affect players' explicit and implicit attitudes as already elaborated. At the same time, there is no empirical evidence suggesting that activities unrelated to the measured phenomenon will affect explicit attitudes towards that phenomenon. Therefore, we expect that the magnitude of explicit attitude change in experimental groups in comparison to a corresponding control group will be smaller in studies using control groups with topic-related activities (e.g., books or lectures about the measured topic) than in studies using control groups with topic-unrelated activities (e.g., playing an unrelated game, watching an unrelated documentary).

H4. *Narrative Video Games Change Player's Implicit Attitudes Only if Compared to Control Groups with Activities Unrelated to the Measured Topic (H4a), but Not if Compared to Control Groups Using Activities (H4b) Related to the Measured Phenomena*.

We do expect that the main mechanism causing implicit attitude change, i.e., co-occurrence of two concepts in one's environment, can occur comparably in the game as in any related activity since it does not require any representation of complex information.

3.3. Game mechanics and their effects on attitudes

Implicit and explicit attitude changes depend on different and often interconnected processes. Therefore, we also expect that the different persuasive game mechanics affect explicit and implicit attitudes differently (see Table 1), as detailed in the following text.

H5. *Implicit Attitudes Are More Affected by Games Using Stereotyping and Meaningful Feedback as Persuasive Game Mechanics than by Other Game Mechanics*.

According to the APE model, mental structures responsible for our implicit attitudes can be changed by the frequent co-occurrence of two (or more) concepts in one's environment. We assume that this frequent occurrence of two concepts is more likely to happen when the game uses Stereotyping and Meaningful feedback as persuasive game mechanics. This is because they both can frequently link two (or more) concepts together within the game. The first one links the measured concept with some stereotypical characteristic; the second one links the measured concept with a positive or negative value assessment of players' actions within the game. Perspective-taking as a persuasive mechanic does not create such ubiquitous and clear connections between the two concepts within the gameplay.

H6. *Explicit Attitudes Are More Affected by Games Using Perspective-taking and Meaningful Feedback as Persuasive Mechanics than by Other Game Mechanics*.

Changing explicit attitudes relies on the acquisition of new information that is potentially challenging one's current beliefs. On the empirical level, several studies suggest that Perspective-taking is able to affect explicit attitudes (e.g., Kampf, 2015, 2016; Kolek, Ropovik, Sisler, van Oostendorp, & Brom, 2021; Todd & Galinsky, 2014). Within the gameplay, Perspective-taking provides players with a complex, multi-perspective take on the measured phenomena. Furthermore, it should be able to mitigate credibility bias (see Section Theoretical Background of Attitude Change). Therefore, we assume that Perspective-taking can affect players' explicit attitudes.

Meaningful feedback as a persuasive mechanic provides players particular freedom to do various, in-game actions related to the measured phenomenon. It assigns the outcomes of these actions a negative or positive value. The reasoning behind this assignment can be comprehensively elaborated within the game narrative. Therefore, we assume that the Meaningful feedback mechanic can affect players' explicit attitudes.

3.4. Game genres

H7. *Narrative Action Games Have a Larger Effect on Implicit Attitudes than on Explicit Attitudes*.

Table 1
Expected effect of persuasive game mechanics on explicit and implicit attitudes.

Persuasive mechanic	Effect on explicit attitudes	Effect on implicit attitudes
Perspective-taking	significant effect	no significant effect
Meaningful feedback	significant effect	significant effect
Stereotyping	no significant effect	significant effect

Narrative action games are more suitable for frequently demonstrating a particular link between two concepts (thus the mechanism affects implicit attitudes) than for providing topic-related complex information (thus the mechanism affects explicit attitudes). Therefore, we assume that implicit attitudes will be more affected by narrative action games than will be explicit attitudes.

H8. *Narrative Non-Action Games Have a Larger Effect on Explicit Attitudes than Do Action Games*.

Attitude challenging information is required to affect someone's beliefs and thus their explicit attitudes. Narrative, non-action video games represent a format with a higher potential to demonstrate complex information to players without time pressure limiting the exploration of their narratives. Therefore, we expect that explicit attitudes will be more affected by narrative, non-action video games than by action games.

3.5. Exploratory analyses

We also have two exploratory goals. First, we examine whether the magnitude of explicit attitude change remains the same over time between the immediate posttest and the delayed posttest or whether it changes. Predictions from theories as regards long-term attitude change are mixed. On the one hand, as described by cognitive dissonance theory (Festinger and Lindzey, 1958), newly acquired information, which is not in line with one's explicit attitudes, will result in the creation of cognitive dissonance and the subsequent motivation to solve it. This can happen either by acquisition of new information, which will resolve the dissonance and change beliefs, or by rejecting the new information. Therefore, this theory assumes that one's explicit attitudinal changes will remain the same over the long term, or that they will, to some extent, shift back to their original values. On the other hand, the "sleeper effect" theory suggests that the persuasive effect can increase over time (Priester et al., 1999) when the message is thoughtfully elaborated. Within this meta-analysis, we intend to explore the direction of the long-term trend of explicit attitude change.

Second, we examine whether there is any significant effect of age or education on attitude change. Here, we do not have any particular prediction.

4. Method

4.1. Inclusion criteria

The process of study selection is outlined in the Prisma Flowchart (Page et al., 2021; see Fig. 1). We have included in our meta-analysis only studies that met the inclusion criteria in the following four domains:

4.1.1. Intervention

We have included only narrative video games as explained in Section Narrative Video Games and Attitude Change. We have also excluded studies focused only on particular game elements not related to the game narrative (e.g., game based-learning elements). However, studies dealing with the effects of game elements possessing any potentially significant meaning for a game narrative, e.g., visual design or a dress-code for game characters, were considered as relevant for our analysis as they can be a part of persuasive game mechanics (see Section Persuasive Game Mechanics).

4.1.2. Outcome

We have included all studies dealing with attitudes, as defined in Section Theoretical Background of Attitude Change, and the changing thereof. This included all effects, impacts, changes or evolutions of attitudes in relation to the experimental intervention; i.e., playing the game. For the purpose of this meta-analysis, attitude change is

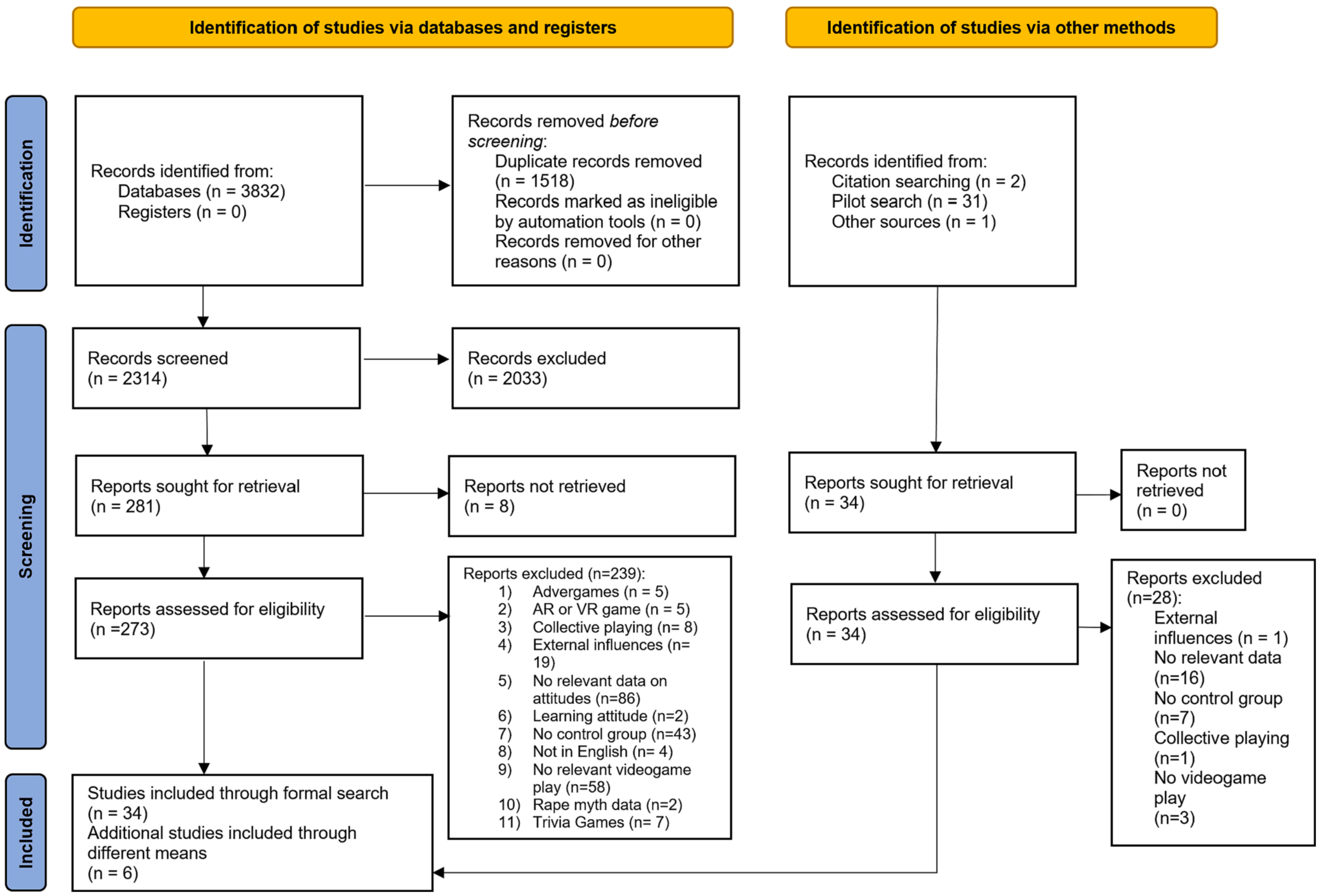

Fig. 1. Prisma Flow Diagram (Page et al., 2021).

considered the dependent variable. Furthermore, attitudes examined in the study had to be related to particular substantive topics represented in the game narrative. Therefore, we excluded all studies focusing on general attitudes towards a) games or playing games (e.g., Garneli et al., 2017; Zhu et al., 2012) or b) towards any other general activity related to the actual playing of the game; e.g., attitude towards competitiveness (Williams and Williams, 2011) or learning attitudes (e.g., Lin et al., 2011).

4.1.3. Study design

Studies in our research sample had to collect quantitative data and allocate subjects to the interventions and some sort of control group(s). At least one experimental group in the study had to experience intervention through a narrative video game. Beyond this general condition, the intervention by the narrative video game should have been the only intended element affecting a player's attitudes towards the topic. Based on this reasoning, we excluded studies with video game interventions that were preceded or accompanied, for example, by a seminar, workshop or collective debate about the topic (Hornung et al., 2000; Strawhacker et al., 2018) and also studies with external elements within the study design, e.g., driving a car in a video game while telephoning in a real life (e.g., Downs, 2014) that would purposefully affect players. Along the same line of reasoning, we also excluded so-called "exergames", e.g., games combining playing with physical activity as a form of exercise. Second, we also excluded studies examining the effectivity of "advergames"; that is, games promoting a product or a brand. Those games' persuasiveness is rooted in their interest in favoring a particular commercial product or brand, which is a qualitatively different factor compared to other games in our study.

4.1.4. Data availability

We also included three practical limitations. First, the paper had to be in English. Second, it had to contain the relevant data about the examined groups or be available upon request (all authors were contacted at least twice). Third, the paper's full text had to be accessible (paywalled papers included).

4.2. Search strategy

To optimize for the best tradeoff between the search's recall and precision, we used an adaptation of the relative recall technique (Sampson et al., 2006). The strategy was to carry out a pilot search with maximum sensitivity, arrive at a legacy set of reference studies, and iteratively adjust the search string to return all or nearly all of the identified legacy set in each database while optimizing for specificity. This iterative development was carried out only during the pilot testing.

First, in September 2018, we pilot-searched Scopus, Web of Science Core Collection and Google Scholar databases using the search string "attitude* AND game*". This broad pilot search was intentionally aimed at maximizing sensitivity, while sacrificing search specificity. We screened the first 1,500 most relevant studies in Web of Science, the first 1,000 most relevant studies in Google Scholar because that was the maximum number that Google Scholar allowed to screen for a single search at the time, and the first 500 most relevant studies in Scopus. Screening in Scopus was discontinued after 500 studies because of saturation, a quickly dropping ratio of included-to-screened studies. We also examined references in the identified records to find information about other relevant studies. We ended up identifying 26 possibly eligible studies. Second, we iteratively adjusted the pilot search string to make it more specific in order to recall all, or nearly all, of the 26 identified legacy studies present in each database. We did this while limiting the maximum number of hits per database to 1,500. Databases used to search the literature are listed in Table S1 in the Appendix A. The final search took place on August 18, 2020 and the final form of the search string was determined prior to this search. Search capabilities differ slightly across databases, so we had to use a distinct translation of our search string for each database. Here is a search operator used for the Scopus database:

TITLE-ABS-KEY ((attitude* OR stereotype*) AND (change OR effect OR significant*[1] OR impact) AND (game*) AND (experiment* OR empirical* OR intervention)).

4.3. Selection of studies

Using various translations of the above-given search string across the 8 databases (Table S1 in the Appendix A), we have identified 3,832 studies. For an additional eligibility check, we have included 34 possibly eligible studies from the pilot search not recalled during the systematic database check, citation searching, and other resources; i.e., one study in review (See Prisma flowchart in Fig. 1). We randomly selected approx. 13 % of the papers ($k = 41$) in the eligibility screening phase to check the inter-rater agreement for the final inclusion/exclusion decisions. Cohen's κ for include/exclude decisions was at 0.875, corresponding to a 97.6 % agreement.

4.4. Data extraction

The screening process was based on reading the study title, abstract and keywords. If these did not clearly indicate whether the study was suitable for meta-analysis based on the inclusion criteria, the full text of the study was assessed. If there was uncertainty about the inclusion of a study, the other authors were consulted about its inclusion or exclusion. The entire screening process was carried out by the first author based on the inclusion criteria pre-specified jointly with the other authors. The first author, who carried out screening, is a PhD researcher experienced in video game development and game psychology. References to all articles from the screening process are available in the supplementary materials, allowing other researchers to replicate the process or use a different methodology for analysis. Likewise, given that we openly share the underlying data and code, anyone is free to explore our coding decisions, re-run the analysis, and test the empirical robustness of our interpretations.

Coding quality was checked by utilizing a second coder who coded a random 30 % of all studies included. Both coders had background in psychology. The coding scheme including the descriptions of all variables was pre-defined prior to the coding process and jointly developed by all authors involved. The development and iteration of the coding manual was guided by the experience and data from the pilot search of 3000 articles in the first phase of the coding process. Data in each coded category were aggregated into larger units where necessary, for example, in the education category, the first level of primary school and the second level of primary school were aggregated into primary school only. The first author of the study, who was also one of the coders, introduced the second coder to the coding manual. At the same time, they jointly applied the coding manual to several selected studies to subsequently compare their results and discuss ambiguities that arose before they began to independently code the data from their selected studies. Coding disagreements were discussed and, if needed, resolved by consulting one of the other authors. The aim was not just to catch coding errors, but also to look for potential problems in the coding scheme. One particular variable was redefined after this process – Persuasive mechanics. Originally, it consisted of 6 values – Stereotyping, Perspective-taking, Economical model, Procedural rhetorics, Reward system and Others. However, the analysis revealed that the categories Economical model, Procedural rhetorics and Reward system overlapped and difference between them was not reliably recognizable. Therefore, we have decided to unite them into one category Meaningful feedback: due to the many similarities between them. Inter-rater reliability for metric variables ranged from Cohen's $\kappa = 0.67$ (for Intervention Duration) to $\kappa = 1$, with a mean Cohen's κ at 0.92. For categorical (mostly binary) variables, we computed the percentage agreement, which ranged from 76 % (whether the effect was focal) to 100 %, with a mean percentage agreement of 94 %. Complete report of inter-rater agreement calculations can be found in the Appendix C (online only).

In case of missing data needed for the computation of effect sizes, we have contacted article authors at least twice by email, leaving at least a 14-day period between attempts.

4.5. Moderators

4.5.1. Type of attitudes

We have distinguished between a) implicit attitudes, i.e., those assessed using response time measures (e.g., Implicit Association Tests); b) explicit attitudes, i.e., those assessed using self-reported questionnaires involving Semantic differentials or Likert Scales.

4.5.2. Intervention duration

We have collected data about the means of experimental intervention durations. In cases where authors stated this value as an interval, we have used its mean value.

4.5.3. Persuasive mechanics

We have divided the studies into four categories: a) Perspective-taking; b) Meaningful feedback; c) Stereotyping (see Section Persuasive Game Mechanics); and d) a Non-defined category for studies with unclear or multiple persuasive mechanics. The games were categorized based on gameplays of experimental games or, in case of their unavailability, we have analyzed their description from the study.

4.5.4. Posttest delay

We have recorded data about the number of days between the intervention and the posttest collection of data from participants. In cases where authors stated this value as an interval, we have used its midpoint.

4.5.5. Game genres

The games used in experimental interventions were coded as being divided into the following two categories: a) action games; b) non-action games (see Section Persuasive Game Mechanics for the theoretical background).

4.5.6. Control groups

We coded control group types as follows: a) Activity unrelated, which involves a lecture/presentation/reading on a topic unrelated to the one in the experimental group; b) Game with a different mechanic, which is a game featuring an unrelated topic and using different mechanics than the game in the experimental group; c) Game with a similar mechanic, which is a game on an unrelated topic but which uses similar mechanics to the game in the experimental group; still, some of their aspects differ like avatars, mission, etc.; d) Activity- related, which involves a lecture/presentation/reading on a topic related to the one in the experimental group; e) No activity apart from study measurements and f) Combination of various activities, such as games, videos or reading.

[1] In the pilot testing of the search string, the keyword "significant" helped to increase the recall of quantitative studies. This was the initial reason for including this keyword, but we also carried out the search with this keyword left out. This led to the omission of two studies, but there was no noteworthy change in the results ($\Delta g = 0.0051$). Details about excluding the keyword "significant" can be found in the Appendix B).

4.5.7. Age/Education

We have collected data on participants' mean ages and also data on education level. We have distinguished the following categories: a) elementary school students; b) secondary school students; c) university students; d) other, which included everyone else outside the first three categories.

4.5.8. Other moderators

Moderators that were coded but not used for any hypothesis are described in Appendix D (online only).

4.6. Effect size computation

We used primarily group posttest means, *SD*s (or *SE*s) and *N*s to compute Hedges' *g*, a standardized mean difference effect type corrected for small sample bias (Hedges & Olkin, 1985). In case group descriptives were not available, we converted the effect sizes from reported test statistics or other types of effect sizes. The computation and conversion of all effect sizes were carried out in code, using formulas laid out in (Borenstein et al., 2009). To counteract a possibly biasing effect from undisclosed subject exclusions, we checked whether the sum of group *N*s approximately matched the total sample size (N +/-2). If it did, we used the respective group *N*s. If it did not, we tried to compute group *N*s based on the reported degrees of freedom, assuming a balanced design. If only the total sample size was reported, we also assumed a balanced design. We excluded effects for which the essential data was not reported and could not be recovered from the authors.

4.7. Analysis

Effect sets including more than 10 effect sizes were considered informative and were synthesized employing multilevel random-effects models with Satterthwaite's small-sample adjustment. We included all theoretically relevant effects for each study. To account for dependencies among the effects, we employed robust variance estimation (RVE) with the CHE working model (Correlated and hierarchical effects; Pustejovsky and Tipton, 2020). These models account for both types of dependencies among the effects simultaneously – nesting of effects within studies and clustering due to estimation of effects based on the same participants. As data on the sampling correlations among the effects is frequently unavailable, a constant sampling correlation of .5 was assumed. As a sensitivity analysis, we relaxed this assumption by varying the sampling correlation from 0 to .6 in increments of .2. To test for equality of effect sizes across the levels of the moderators studied, we used the robust HTZ-type Wald test (Pustejovsky and Tipton, 2020).

Apart from the effect size estimates, we examined the absolute and relative heterogeneity using τ and I^2, respectively. To estimate the range of true effects to be expected in similar future studies, we calculated 95 % prediction intervals.

Prior to our analyses, we carried out an in-depth diagnosis of the random-effects meta-analytic model. Specifically, we screened for influential outliers using the Baujat plot and influence diagnostics indices. Outliers exerting an excessive influence on the meta-analytic model (if any) were only excluded in a sensitivity analysis.

In a sensitivity analysis, we also checked whether excluding studies with a high overall risk of bias (utilizing algorithmic-based judgment) and effects based on mathematically inconsistent means or SDs did have a meaningful influence on the meta-analytic inferences.

All models were fitted using restricted maximum-likelihood estimation using R packages metafor, version 2.5 (Viechtbauer, 2010) and clubSandwich, version 0.4.2. (Pustejovsky, 2020). The data analysis was carried out in R also using the following packages: esc (Lüdecke, 2017), tidyverse (Wickham, 2019), lme4 (Bates et al., 2015), dmetar (Harrer et al., 2019), and psych (Revelle, 2018).

4.8. Adjustment for publication bias

As null or negative results are less likely to get published, available studies represent a biased sample of the conducted (and all conceivable) studies. Under the influence of publication bias, the meta-analytic effect size estimates tend to be inflated to an unknown and possibly substantial degree and have an excessive false-positive rate (Carter et al., 2019; Hong and Reed, 2021; Ioannidis, 2008).

Although bias adjustment methods assume a more realistic selection process, they may fail to recover the "true" magnitude of the studied effects under a number of realistic conditions. The estimates should thus rather be seen as approximations (see Ropovik et al., 2021). If the adjusted estimates from selection models markedly diverged from the crude meta-analytic estimates, then we primarily used bias-corrected estimates to guide our substantive inferences.

4.8.1. Selection models

As the primary bias-adjustment approach, we applied a permutation-based implementation of the step-function selection model (see McShane et al., 2016). Selection models are a statistically principled, highly flexible family of models that directly map the functional form of the biasing selection process. In short, a 3-parameter selection model includes the following parameters: population effect size, heterogeneity, and the likelihood that a non-significant vs. significant result gets published. The model then uses maximum likelihood to estimate the three parameter values under which the observed data are most likely (McShane et al., 2016). By default, we applied the 4-parameter selection model (it also estimates the probability of the effect being in the opposite direction). If there was too little data (at least one of the *p*-value intervals contained less than 4 focal *p*-values), the estimation procedure automatically reverted to the 3-parameter selection model. All selection models (including the one-parameter selection models *p*-uniform* and *p*-curve) subset only the results that were deemed to be the study's focal effects (reported in the abstract).

As selection models suited for the analysis of multi-level data are yet to be developed, the dependencies among the effects were handled using a permutation-based approach. We randomly drew only a single focal outcome from each study, estimated the model repeatedly in 5,000 iterations, and averaged over this set of iterations by taking the model with the median estimate. This procedure sidesteps the use of arbitrary and potentially biasing decision rules for choosing independent effects.

To examine the variability in adjusted effect size estimates under different assumptions about the selection process, we also computed a series of three (Vevea and Woods, 2005) step function models with a priori defined selection weights. We used a fine-grained 10-step function to model different levels of severity of bias: moderate, severe, and extreme.

4.8.2. Exploratory bias-adjustment methods

For exploratory purposes, we also supplemented the primary selection modeling approach with the following secondary methods. Namely, we used the multi-level, RVE-based implementation of the PET-PEESE method (Stanley and Doucouliagos, 2014), Weighted Average of the Adequately Powered studies (WAAP-WLS estimator; Stanley et al., 2017) and *p*-uniform*(van Aert and van Assen, 2018). The details about the implementation of these methods with the results of the latter two methods are provided in the Appendix E).

4.9. Quality of evidence assessment

To appraise the quality and integrity of the evidence, we have carried out the following procedures.

4.9.1. Risk of bias

First, we have assessed the risk of bias using the Revised Cochrane risk of bias tool for randomized trials (RoB 2; Sterne et al., 2019). The risk of bias was assessed in five domains: namely, bias arising from the randomization process; bias due to deviations from intended interventions; bias due to missing outcome data; bias in measurement of the outcome; and bias in selection of the reported result. The judgments about bias in these domains were made using an algorithmic approach based on signaling items – a decision-tree guided evaluation aiming to elicit assessor's judgment using several questions to identify potential biases. These questions, mapped onto specific risk domains, serve as 'signals' of potential bias. A decision tree pre-specified in the RoB 2 documentation (Sterne et al., 2019) directs the assessor's path, integrating responses to these signaling questions and producing preliminary bias risk judgments. While the process fosters systematic and objective assessment, final risk of bias determinations also require the integration of assessors' informed, context-specific judgments. When justified, the assessor could override the suggested risk of bias judgments, but this could be done only conservatively, i.e., in the direction of downgrading the judgment.

4.9.2. Numerical inconsistencies in reported means and SDs

Second, using GRIM (Brown and Heathers, 2017) and GRIMMER (Anaya, 2016) tests, we tried to identify effects based on means or standard deviations that are mathematically inconsistent with the reported sample sizes. Checking for such inconsistencies is possible if the outcome was a discrete variable (e.g., Likert-type individual items or scales). In that case, means and *SDs* follow a fixed granular pattern for each combination of *N* and the number of items (Anaya, 2016; Brown and Heathers, 2017).

4.9.3. Numerical inconsistencies in reported p-values

Next, we screened all included studies for inconsistencies in reported *p*-values. This machine-based screening was carried out using the statcheck package (Epskamp and Nuijten, 2018). The method works as follows: (1) pdf files are converted to plain text, (2) which gets scanned for statistical results reported in APA style, (3) test statistics and degrees of freedom are extracted to recompute the *p*-value, (4) which is compared to the reported *p*-value. Having extracted that data, we computed in which proportion of cases the *p*-values were inconsistent with the reported test statistics and how many of those cases led to an inferential decision error.

4.9.4. Assessment of evidential value

Using the *p*-curve method, we tested whether selective reporting can be ruled out as the sole explanation of the observed findings (Simonsohn et al., 2014). If there is evidential value in the given literature, a right-skewed distribution of *p*-values can be observed regardless of power. It follows that a set of direct replications is expected to yield a non-zero effect. On the other hand, a left-skewed distribution of *p*-curves may indicate a substantial prevalence of questionable research practices in the literature.

In the present meta-analysis, *p*-values were recomputed from the reported descriptive statistics. The dependencies between the *p*-values were handled using a permutation-based procedure, repeatedly drawing only a single focal effect from each study (with 200 iterations), estimating the *p*-curve, and averaging over the set by selecting the model with the median *z*-score for the right-skew of full p-distribution.

4.9.5. Median statistical power in the literature

Lastly, we also computed the average statistical power to detect various smallest effect sizes of interest (0.20, 0.50, and 0.70). In the Appendix F, we also report median power to detect the bias-corrected estimates.

4.9.6. Transparency and openness

We report how we selected studies for this meta-analysis and the process is outlined in the Prisma Flowchart (Page et al., 2021); see Fig. 1), including all data exclusions, all manipulations, and all measures in the meta-analysis. All data, analysis code, and research materials are available at Open Science Framework on the following link https://osf.io/4aeqt/. Data were analyzed using R packages metafor, version 2.5 (Viechtbauer, 2010) and clubSandwich, version 0.4.2. Data analysis, used R packages and procedures are documented in the Methods section. This study's design and its analysis were not pre-registered.

5. Results

Sixty-seven studies from 40 papers matched the inclusion criteria. Out of those, 58 independent-sample studies (reported in 35 papers) provided sufficient information to recompute 119 effect sizes. In total, the included set of effects summarized data from 14,272 unique participants, with a median *N* across the included effects of 127. The vast majority of the included effects (96 %) originated from randomized studies (see Appendix F for the full analytic outputs and Appendix G for the list of studies and the effect sizes; both online only).

Prior to the analyses, we screened the full meta-analytic set of 119 effects for outliers. Based on the examination of the Baujat plot and influence diagnostics indices, none of the included effect sizes exerted undue influence on the meta-analytic model.

In what follows, we first carry out a comprehensive synthesis of the entire literature that provides evidence on video games' overall effect on attitude change. Second, we address the substantive questions posed by this review. Third, we look at the methodological moderators to identify design-related and meta-study factors that may affect the size of the detected effects in this literature. Fourth, we conduct a detailed appraisal of the quality of empirical evidence; check for the presence of reporting inconsistencies; and indications of *p*-hacking at the level of the literature. Lastly, we carry out several sensitivity analyses to examine the robustness of our results to arbitrary methodological decisions.

5.1. Narrative video Games' overall effect on attitude change

The set of effects reported in the literature that concerned narrative video games' effect on attitude change exhibited a high degree of heterogeneity, $Q(118) = 483$, $p < .001$. The standard deviation of true effects was $\tau = .40$, while $I^2 = 93$ % of the total variance across the observed effect estimates was of a systematic nature (86 % due to between- and 7 % due to within-cluster heterogeneity). Accordingly, the 95 % prediction interval was wide; with the true effect in a newly published study being expected to fall between −0.55 and 1.05.

The Random-effects RVE-based model estimated a mean effect size of $g = 0.25$, 95 % CI (0.14, 0.37), $p < .001$, which is a small effect. Publication bias-adjusted effect estimated by the permutation-based 3-parameter selection model (3PSM) was likewise significant and of similar magnitude as the unadjusted estimate, $g = 0.33$, 95 % CI (0.14, 0.53), $p = .001$. To further examine the variability in bias-adjusted estimates under different assumptions about the selection process, we also computed a series of (Vevea and Woods, 2005) step function models with a priori defined selection weights. The effect size estimates for the assumed moderate, severe, and extreme selection of results for publication were 0.15, 0.04, and −0.07. The markedly higher estimate for the maximum-likelihood-based 3PSM thus indicates that the selection by significance in the given literature is less severe than the selection process assumed by the "moderate" model. In fact, only 14 % of all effects included from primary studies were significant and Kendall's correlation between the effect sizes and their standard errors was only $r = .04$ – suggesting only slight asymmetry in the chances of non-significant and significant effects to be published (as can also be seen from the funnel plot in Fig. 2).

As a last, exploratory approach to bias correction, we also applied the multi-level RVE-based implementation of the PET-PEESE method. Assuming a hypothetical study with an infinitely large sample size, the method did not indicate the presence of an effect and returned a bias-adjusted effect size estimate that was effectively zero, $g_{PET} = -0.03$,

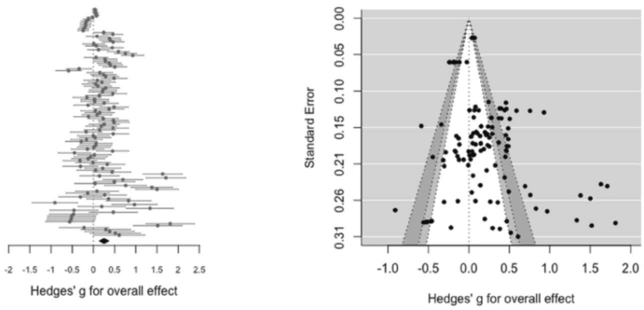

Fig. 2. Forest Plot and Funnel Plot for Video Games' Overall Effect on Attitude Change. In the Forest Plot, Effects are Sorted by Ascending SE.

95 % CI (-0.34, 0.27), p =.84. The estimate was, however, rather weakly informative (judging by a relatively large CI width) and overlapped with the 3PSM estimate and thus was not significantly different.

Although this exploratory result adds a layer of uncertainty by pointing to the rather suboptimal amount of information in the data, the primary analyses (naive meta-analytic model and 3PSM) indicate the presence of a small, but robust, general effect of narrative video games on attitude change.

H1. *The Effect of Narrative Video Games on Change in Explicit and Implicit Attitudes.*

A far larger proportion of the included studies examined the effect of video games on explicit attitude change (k = 101, 10 % significant) than on implicit attitude change (k = 18, 39 % significant). We did not detect substantial heterogeneity in the implicit attitude effects (neither absolute, nor relative), while the heterogeneity of explicit attitude effects was substantial (see Table 2).

Both sets of effects yielded small-to-medium-sized average effects. Likewise, our primary bias-adjustment method (3PSM) indicated that, even after accounting for publication bias, the effect size estimate did not diverge from the unadjusted estimates. The pattern of estimates for the series of Veeva & Woods selection models was also lower than the naïve estimates similar to the overall effect results. This suggests low severity of publication bias in both subsets. As the secondary, exploratory bias-adjustment method, PET-PEESE, detected a significant effect only for implicit attitude change (but not for explicit attitude change) of a practically identical magnitude as for the unadjusted effect. Overall, the literature we studied provides empirical evidence for a modest efficacy of video games for change of both explicit as well as implicit attitudes (based on the secondary PET-PEESE analysis, the effect on implicit attitudes seems more empirically robust though). H1 was thus supported. For more detailed results and plots, please see the full analytic output in the Appendix F.

To compare the mean effects bound to explicit vs. implicit attitude change, we tested a meta-regression model controlling for several design-related factors that may have been prognostic with respect to the effect sizes (i.e., might vary between these sets of effects). We adjusted the comparison for overall risk of bias, published status, mean age of participants, and whether the intervention was administered in a lab. We did not find a difference between the effects related to explicit vs implicit attitude change, Wald's-type test $F(1, 3.13) = 0.45$, p =.55. Nor was there an effect with the covariates left out.

5.2. *Substantive moderators*

H2. *The Relation between the Duration of Intervention and Attitude Change.*

We did not detect a relationship between the duration of intervention and attitude change in the overall set of effects: meaning lack of support for H2 (Table 3). When broken down, however, effect sizes related to explicit attitudes (H2a) did not prove to be associated with the duration of the intervention (p =.19), while effects related to implicit attitudes (H2b) did (p =.003). When we subset just the effects based on delayed posttests, the magnitude of explicit attitude change also did not change significantly over time (from immediate posttest to the delayed posttest).

H3 & H4. *Characteristics of the Comparator Group and Attitude Change.*

We also first assumed that the magnitude of explicit attitude change (H3) is smaller in studies using control groups with topic-related activities than in studies using control groups with topic-unrelated activities. Although the difference between the respective subgroups was in the expected direction, g = 0.29 for subgroup using unrelated activities and

Table 2
Meta-analysis Results for Video Games' Effect on Explicit and Implicit Attitudes.

	k	g [95 % CI]	SE	τ	I^2	3PSM estimate	3PSM p-value	V&W estimate
Explicit	101	0.24 [0.11, 0.37]	0.07	.43	94 %	0.32 (0.09, 0.55)	.006	0.12 (-0.01, 0.26)
Implicit	18	0.36 [0.24, 0.48]	0.05	.11	31 %	0.37 (0.20, 0.54)	<.001	0.31 (0.16, 0.46)

Note. Values in brackets represent 95 % CI. V&W = Veeva & Woods step function model assuming moderate selection.

Table 3
Substantive Moderators.

Hypothesis	Moderators	Groups of effects	K	Effect size (95 % CI)	Statistical test
H2	Duration of intervention		115	B = 0.05 (-0.03, 0.14)	t = 1.31, p =.19
H3	Control group type & explicit attitude change	Topic-related activities	30	g = 0.13 (-0.01, 0.28)	F(1, 3.67) = 3.05, p =.16
		Topic-unrelated activities	89	g = 0.29 (0.14, 0.44)	
H5†	Persuasive mechanics in implicit attitudes	Stereotyping & Meaningful feedback	14	g = 0.37 (0.24, 0.51)	F(1, 6.25) = 10.7, p =.02*
		Perspective-taking	2	g = 0.17 (-0.17, 0.52)	
H6†	Persuasive mechanics in explicit attitudes	Perspective-taking & Meaningful feedback	92	g = 0.23 (0.08, 0.38)	F(1, 3.59) = 2.36, p =.21
		Stereotyping	8	g = 0.42 (0.10, 0.73)	
H7 & H8	Type of game	Action games –> Implicit attitudes	13	g = 0.32 (0.21, 0.42)	F(1, 3.18) = 1.38, p =.32
		Action games –> Explicit attitudes	31	g = 0.23 (0.03, 0.42)	
		Non-action games –> Explicit attitudes	69	g = 0.26 (0.10, 0.42)	F(1, 14.30) = 0.18, p =.68
		Action games –> Explicit attitudes	31	g = 0.21 (0.02, 0.39)	
Exploratory	Age		91	B = -0.01 (-0.03, −0.00)	t = -2.06, p =.04*
Exploratory	Education	Primary, Secondary, Tertiary, as ordinal	95	B = -0.11 (-0.00, 0.23)	t = 1.91, p =.06
Exploratory	Gender		105	B = 0.60 (-0.14, 1.33)	t = 1.59, p =.11

Note: † = Satterthwaite small-sample correction used to compute the test and Cis. * = Significant at the α < .05 level.

$g = 0.13$ for related activities subgroup, it was not significant (Table 3). This indicates a lack of support for H3.

At the same time, we expected that video games would change player's implicit attitudes (H4) only if compared to control groups with unrelated activities towards the measured topic, but not if compared to control groups using related activities towards the measured phenomena.

This prediction (as stated in H4) could not be tested as none of the implicit attitude effects were based on a design with controls doing related activities.

H5 & H6. *Persuasive Mechanics and Attitude Change.*

Next, we tested whether implicit attitudes are affected more by games using Stereotyping and Meaningful feedback as persuasive mechanics than by Perspective-taking (H5). Games using Stereotyping and Meaningful feedback were associated with significantly larger effect sizes, $g = 0.37$ compared to games employing Perspective-taking, $g = 0.17$. The difference was statistically significant, thus corroborating H5.

On the other hand, we assumed that explicit attitudes are affected more by games using Perspective-taking and Meaningful feedback as persuasive mechanics than by other game mechanics, i.e., Stereotyping [H6]. While the pattern of mean effect sizes for these subgroups was in the opposite direction, the difference was not significant. H6 was thus not supported by our data (see Table 3).

H7 & H8. *Action Games' Effect on Attitude Change.*

Concerning the type of game, we expected that action games have a larger effect on implicit attitudes than on explicit attitudes (H7). Although the pattern of meta-analytic estimates was in the expected direction, $g = 0.32$ for implicit and $g = 0.23$ for explicit, the difference was not statistically significant.

From a different perspective, we also hypothesized that non-action games have a larger effect on explicit attitudes than do action games [H8]. Here, both subgroups were quite similar in terms of average effect size, with $g = 0.26$ for non-action games and $g = 0.21$ for action games. This difference was therefore not significant. Current data provided evidence neither for H7, nor for H8 (see Table 3).

5.2.1. Basic characteristics of the sample and attitude change

As an exploratory analysis, we also looked at the relationship between attitude change and (1) mean age of the participants, (2) education level of participats and (3) the sample's gender composition (see Table 3). First, we found some feeble evidence of a negative relationship between the magnitude of attitude change and age. The impact of video games very slightly diminished with increasing age. Second, when age was operationalized as a discrete variable, breaking down the effects by the education level of the participants that the respective studies have targeted (i.e., primary, secondary, and tertiary education students), the moderation effect was even weaker and formally non-significant. Third, the effect of video games seemed invariant with respect to the sample's gender composition (percentage of females).

5.3. Methodological and Meta-Study moderators

In brief, we also assessed the moderating role of several design-related and meta-scientific factors.

First, we did not detect a difference between effects from non-laboratory studies ($k = 23$, $g = 0.15$) compared to effects coming from in-lab studies ($k = 88$, $g = 0.28$), $F(1, 5.68) = 1.05$, $p = .35$.

Second, studies restricting the sampling frame using demographic factors ($k = 97$) that may play a role in attitude change outcomes (thus decreasing the sampling variability) found larger effect sizes, $g = 0.31$, than studies not applying a restrictive sampling scheme ($k = 20$), $g = -0.04$, $F(1, 9.64) = 13.50$, $p = .005$. Since the majority of studies used some kind of range restriction, effect sizes found in this meta-analysis will likely prove to be smaller in more general samples. That said, an F-test of the equality of variances could not reject the hypothesis that the population variances for restricted and unrestricted samples were identical, $F(19, 97) = 0.90$, $p = .58$. Thus, the effect of range restriction may well be negligible.

Third, although effects produced by commercial games ($k = 51$, $g = 0.35$) were larger in our sample of effects, compared to non-commercial games ($k = 65$, $g = 0.18$), the difference (or the precision of the estimates) was not large enough to be significant, $F(1, 51.7) = 1.89$, $p = .18$.

Lastly, using covariate-adjusted models, we examined whether (a) the precision of the study designs has been improving over the years, (b) whether more informative (lower *SE*) studies tend to attract more citations, or (c) the same studies tend to get published in higher-impact journals, (d) whether studies reporting larger effect sizes tend to get more attraction, and lastly (e) whether there is a decline effect where studies showing more extreme (possibly opposite) results appear early in the research line rather than later as data accumulates (Ioannidis, 2008). We found empirical support only for (c) the positive relationship between the study's precision and the journal's impact. Details and results of these analyses can be found in the Appendix F.

5.4. Assessment of the quality of evidence underlying the overall effect

Most concerns regarding the risk of bias in the included set of studies were due to bias arising from the randomization process (49 % of studies being at low risk of bias) and due to bias in the selection of the reported results (only 20 % being at low risk). Overall, only 29 % of studies were rated as low risk, 15 % raised some concerns, while 56 % were at high risk of bias (i.e., being at high risk of bias in at least one domain or raising some concerns in multiple domains). See Fig. 3 for more detailed results on each domain. The picture regarding the relationship between the effect sizes and studies' overall risk of bias was not entirely clear. On average, studies with low risk ($k = 14$) of bias reported an effect of $g = 0.10$, those having some concerns $g = 0.41$ ($k = 30$), and studies with high risk of bias yielded a mean effect size of $g = 0.17$ ($k = 75$). That said, studies differing in their overall risk of bias did not differ significantly in their effect sizes, albeit marginally so, $F(2, 18.2) = 3.41$, $p = 0.06$.

Second, 71 % of the included effects targeted outcomes measured on a discrete scale. To examine the presence of reporting inconsistencies in the literature, we checked whether means and standard deviations underlying these effects were mathematically consistent with the reported sample sizes. Here, we found that 22 % of these effects were flagged as being based on at least one mathematically impossible mean (18 %) or *SD* (4 %) – that is, mean or *SD* that cannot be arrived at given the reported *N*.

Third, we carried out machine-based, full-text screening (Epskamp and Nuijten, 2018) to extract all test statistics and the associated p-values reported in all 40 papers included (not just the ones providing sufficient effect size data). These test statistics were properly reported in APA style in 48 % of the papers. Out of 469 extracted results, 10.5 % were flagged as errors. In 8 % of those flagged as errors, the error led to the opposite conclusion regarding the presence of the effect. Overall, reporting errors related to inconsistency between the reported test statistics and the p-value were distributed relatively evenly across the literature: with at least one such error being present in 58 % of all papers included.

Fourth, we examined the indications of the presence of p-hacking using a permutation-based p-curve analysis. The distribution of p-values was right-skewed, indicating the presence of evidential value, $z_{half} = -2.74$, $p_{half} = .003$. As the p-curve only includes independent significant effects, the median model was based on just 7 effect sizes. That said, we did not detect any pattern consistent with a large prevalence of p-hacking in this literature. This also held for both subgroups, implicit attitudes as well as explicit attitudes subsets, where both p-value distributions were associated with $p_{half} < .005$. The median p-curve for the overall effect can be seen in Fig. 4.

Lastly, the overall literature was adequately powered (.98) to detect

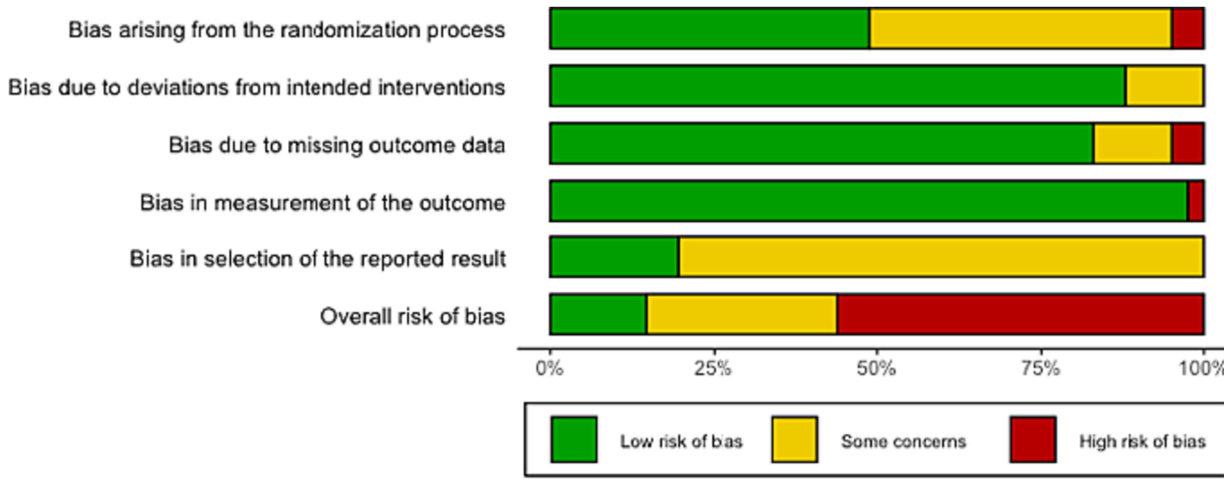

Fig. 3. Risk of Bias Chart.

effect sizes of medium magnitude ($d = 0.50$) on average. On the other hand, the median statistical power to detect a small hypothetical effect size ($d = 0.20$) was relatively low; at only .36, on average.

6. Discussion

6.1. General effect of narrative video games on attitudes

On the most general level, available evidence indicates that narrative video games do affect attitudes towards the depicted topics: both explicit and implicit. As we have already mentioned, video games today are not only a popular leisure time media in our societies; they are increasingly being used for personal storytelling, news reporting, cultural narratives, and political propaganda (Zimmerman 2013). The fact that games seem to have a marked impact on attitudes is crucial for debates about the way they represent our world. For instance, narrative video games' representations of marginalized groups, history or gender stereotypes are, according to these findings, transcending the medium itself and affect our daily lives and our interpretation of the world.

Beyond this general finding, we have also examined multiple moderators of this global effect: like intervention duration, different game design models embedded in video games, and other methodological aspects of research designs utilized in this field.

6.2. Intervention duration

Evidence related to Hypothesis 2 was rather mixed. The intervention duration did not prove to have the presumed effect on explicit attitudes (H2a). Nevertheless, the longer intervention duration resulted in larger effect sizes on implicit attitudes (H2b) as presumed.

This means that the longer one plays the narrative video game, the larger the impact it has on their implicit attitudes but not necessarily on their explicit attitudes. These findings support the idea theoretically posited by the APE model: that implicit attitude change is caused by the frequent co-occurrences of the measured concept with another concept of positive or negative value. As exposure time seems to be related to the magnitude of the effect on associations about conveyed topics, this implies that narrative video games with longer durations may have a more significant effect on players' attitudes.

Possibly, as regards the lack of support for H2a, the effect on explicit attitude change may be tied to the role of participants' relationships to the measured topic. Unlike implicit attitudes, explicit attitudes are based on multiple, consistent, logical conclusions which reflect what one considers to be truth. As such, the participants' relationships to the topic is potentially of similar or greater importance than the intervention duration. This has been suggested by a few studies. For example, the results of studies by Kampf (2016) and Alhabash and Wise (2015) indicate that belonging to a particular national group is crucial for the effect of narrative video games on attitude change towards the Israeli-Palestinian conflict or the parties involved.

The finding that the length of the intervention might not be essential for changing explicit attitudes is also supported in the study by (Pena et al., 2018). They collected data in the middle of the experiment and at its end; participants' explicit attitudes towards the topic did not change

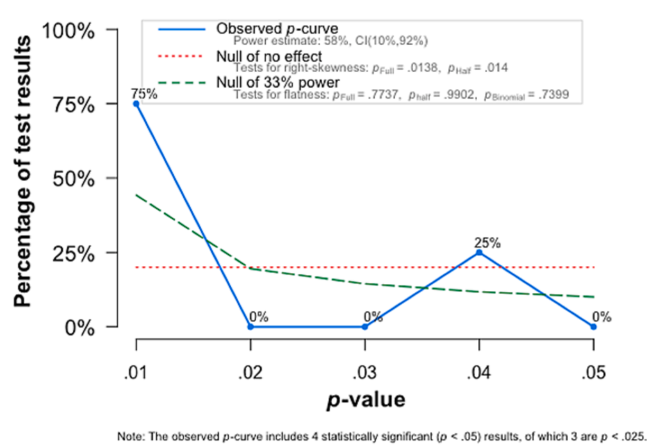

Fig. 4. *p*-curve or the overall effect.

between those two data collection points even though the duration of the intervention was doubled. However, more data is needed on that issue.

6.3. Persuasive mechanics

Our results indicate that particular persuasive game mechanics (Stereotyping and Meaningful feedback) have a larger effect on implicit attitudes than others (H5). On the other hand, we have found no support for the idea that Perspective-taking and Meaningful feedback have larger effect on explicit attitudes compared to the Stereotyping mechanic (H6).

These results suggest that particular game design patterns linking depicted topics with some characteristics may differ in their effects on implicit attitudes. However, our findings regarding explicit attitudes do not support our original prediction. There are two possible explanations. First, similar to intervention duration, explicit attitudes may be more prone to being affected by participants' relationships to the measured concept (Alhabash & Wise, 2015; Kampf, 2016). Second, the chosen category Meaningful feedback is relatively broad. However, that is speculation and further research needs to be done to examine it.

On a general level, our findings about the persuasive game mechanics are the first such complex data brought to the debate about which particular game elements are responsible for attitude change. Focusing on game persuasive mechanics, we have identified which game design patterns are associated with changes in implicit attitudes. However, there was a lack of such clear signals in the domain of explicit attitudes. This particular area of research is still in its beginnings, but our findings are not only relevant for mapping the effect of narrative games on society, they also possess unique value for game designers and experts in education. They can help the latter groups effectively develop games for change. Nevertheless, more studies are needed to fully understand the particular game mechanisms that are affecting our attitudes.

6.4. Action vs Non-Action games

Our analysis did not prove different effect of narrative for action (H7) and non-action games (H8) on implicit or explicit attitudes, so neither hypothesis was supported. These findings suggest that, whether or not the game put players in time pressure situations with a need for fast reactions and a focus on hand-eye coordination, it has no noticeable effect on players' attitudes. Therefore, the distinction between action and non-action games does not seem to be key in research on attitude change.

6.5. Comparator groups

Available evidence does not indicate that magnitude of explicit attitude change is greater for control groups using topic-unrelated activities than for those using topic-related activities (H3). We assume that the reason games are more effective in attitude change than these more traditional formats is that narrative video games require players to interact with the depicted topics in game narratives in order to proceed further in the gameplay; the games often offer players opportunities to react to the depicted topics.

We have not identified any study allowing us to evaluate the hypothesis related to control groups and implicit attitudes (H4). Therefore, we were not able to test it.

6.6. Exploratory Goals: Gender and age effects

In our exploratory analysis, we have not identified any effect of gender on the magnitude of attitude change. However, the data suggests that the potential of video games to affect attitudes slightly decreases with age. Again, these outcomes are of an exploratory nature and should be approached with caution. A study by (Wang and Chen, 2006) suggests that working memory could have a mediating influence on the effect of age on attitude change. Specifically, that attitude change among older adults ($M_{age} = 74.97$) relies more on argument quantity than argument quality compared to younger adults ($M_{age} = 20.03$). The authors suggest this is caused by the limits of working memory at a higher age. However, the weighted mean age across the studies included in the present meta-analysis was relatively low ($M_{age} = 21.13$), thus the suggested limits of working memory are unlikely to be a reason for the difference. Also, several studies disprove the general effect of age on susceptibility to change in attitude (e.g., Krosnick and Alwin, 1989; Tyler and Schuller, 1991). Accordingly, we see two other variables possibly responsible for this age effect. First, video games as a format may be a less accessible or trustworthy format for older generations. Second, the relationship of older players to the depicted topics in games might be different. All these interpretations require more data for further clarification.

6.7. Summary of implications

In educational practice, our findings can be used in many positive ways. For instance, videogames can promote better and more complex understanding of our past. They can also help in conflict areas by promoting less polarized views on the crucial issues. Also, even in informal learning, messages and mechanics embedded in video games can help to achieve positive changes in our society, for example by supporting positive views on protecting our environment and the subsequent change of behavior. In the context of how widespread video games are, they should not be underestimated in how easily they can speak to the key political or cultural debates we have about the critical issues of the day. However, it should be borne in mind that they can equally be misused for political propaganda.

The longer we spend with the video the videogame the larger effect they might have on our implicit attitudes. Thus, large commercial video games, in which players often spend dozens of hours playing them in informal settings can be a very influential medium in terms of what themes they portray and how they portray them.

Our findings do not suggest a relationship between time spent playing the game and the magnitude of explicit attitude change. Thus, from a practical point of view, video games that promote positive attitudes are suitable for effective implementation in school settings, despite the limited amount of time offered in the classroom. At the same time, it should be said that not all games can effectively change attitudes. Certain assumptions can be made based on the game mechanics used (see the Sections on Persuasive Mechanics and Action vs. Non-Action Games in the discussion), but given the limited data, it is still useful to consult the evaluation of the educational benefits of individual games. In addition, our findings (see the Section Comparison Groups) suggest that video games may be more effective than other formats in changing explicit attitudes. This fact alone might be sufficient to encourage their use in educational contexts.

6.8. Future directions

In general, there was sufficient information in the literature to support the main hypothesis. However, the field is still relatively fragmented and more studies are needed to understand fully the effect of video games on players and society. For instance, very few studies collected data on the long-term effects of video games. Also, more studies are focused on explicit attitudes than are on implicit attitudes. Furthermore, research on particular game elements responsible for attitude change is only in its beginnings, so there are no other studies to compare our results with since our analysis of the persuasion mechanics is the first of its kind. Plus, the experiments in the studies included were predominantly focused on the effects of relatively short interventions compared to often multiple-hour-long gameplays offered by most popular video game titles. In relation to this, as our data suggests, the effect of intervention duration on implicit attitudes is worth monitoring. Upcoming studies should explore the effects of longer exposure to video games (for ecological validity reasons). Also, this study has approached narrative games as a group and did not differentiate between different types of narrative approaches in the analyzed video games. Future research should explore classification of studies based on the differences in the structure and centrality of game narratives to gameplay. Finally, much more focus, especially when studying explicit attitudes, should be given to players' stances on the topics assessed. This should be mapped using attitude measurements and studying the role of individual characteristics in general. For instance, our exploratory analysis suggests that older participants might be less susceptible to attitude change than younger participants.

6.9. Limitations

There are several limitations to this study that are worth mentioning. First, despite the relatively elaborate and labor intensive academic database search for all relevant papers, our final research sample consists of only 35 papers meeting our criteria. Second, any comparisons of the effects of video games on explicit and implicit attitudes should be interpreted with caution, as there were substantially more studies focused on changing explicit attitudes ($k = 101$) than on changing implicit attitudes ($k = 18$). Third, our meta-analysis did not focus in detail on players' characteristics: especially their relationships to depicted topics. That is because the vast majority of studies do not report this data. Fourthly, our meta-analysis expects that the majority of players in the analyzed studies play the intervention games as intended by developers or researchers having on average a comparable experience of the game. In our opinion, none of the limitations undermines the paper's key findings.

7. Conclusion

This meta-analysis is the first contribution of its kind to the body of knowledge in the field of narrative video games and their effects on attitudes. The findings suggest that narrative video games are able to affect players' attitudes towards the topics depicted in these games. This effect is present in studies focused on both implicit and explicit changes in attitude.

Furthermore, our data suggests several moderating effects. Most notably, it seems that longer intervention duration and persuasive game mechanics such as Stereotyping and Meaningful feedback result in larger implicit attitude change. Also, our exploratory analysis suggests that participants' attitudes are less affected by video games as they reach a higher age. However, this effect is rather small.

Narrative video games are widespread phenomena in our societies and culture. Our meta-analysis provides evidence that they are not only part of our world, but that they are also shaping how we think about it. There are still significant gaps in fully understanding this process, and more studies are needed to provide a better picture of how to design video games promoting positive attitudinal change.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data and the code supporting the article are shared via OSF Data Repository referenced in the article.

Acknowledgment

For this work, CB, LK and VS were supported by the PRIMUS/HUM/03 project at Charles University. IR and VS were supported by NPO Systemic Risk Institute LX22NPO5101 funded by European Union – Next Generation EU. IR was further supported by PRIMUS/20/HUM/009 and APVV-18-0140 grants. VS was further supported by KREASCZ.02.1.01/0.0/0.0/16_019/0000734. We would like to thank Markéta Matějová and Veronika Petrova for their help in checking the data.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cedpsych.2023.102225.

References[2]

Alblas, E. E., Folkvord, F., Anschütz, D. J., van't Riet, J., Granic, I., Ketelaar, P., & Buijzen, M. (2018). Investigating the impact of a health game on implicit attitudes towards food and food choice behaviour of young adults. *Appetite, 128*, 294–302. https://doi.org/10.1016/j.appet.2018.05.141

Alhabash, S., & Wise, K. (2015). Playing their game: changing stereotypes of Palestinians and Israelis through videogame play. *New Media & Society, 17*(8), 1358–1376. https://doi.org/10.1177/1461444814525010

Anaya, J. (2016). The GRIMMER test: A method for testing the validity of reported measures of variability. *In PeerJ*. https://doi.org/10.7287/peerj.preprints.2400

Atkins, B. (2003). More than a game: The computer game as fictional form. *Manchester University Press*. https://doi.org/10.7228/manchester/9780719063640.001.0001

Banaji, M. R., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*. New York: Delacorte Press.

Barthel, M. L. (2013). President for a day: Video games as youth civic education. *InformationCommunication & Society, 16*(1), 28–42. https://doi.org/10.1016/j.chb.2014.11.062

Barwood, D., Smith, S., Miller, M., Boston, J., Masek, M., & Devine, A. (2020). Transformational Game Trial in Nutrition Education. Australian. *Journal of Teacher Education, 45*(4). https://doi.org/10.14221/ajte.2020v45n4.2

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Behm-Morawitz, E., Hoffswell, J., & Chen, S. W. (2016). The virtual threat effect: A test of competing explanations for the effects of racial stereotyping in video games on players' cognitions. *Cyberpsychology, Behavior, and Social Networking, 19*(5), 308–313. https://doi.org/10.1002/ab.20329

Bluemke, M., Friedrich, M., & Zumbach, J. (2010). The influence of violent and nonviolent computer games on implicit measures of aggressiveness. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 36*(1), 1–13. https://doi.org/10.1002/ab.20329

Bogost, I. (2010). *Persuasive games: The expressive power of videogames*. MIT Press.

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of research synthesis and meta-analysis* (pp. 221–235).

Brady, S. S., & Matthews, K. A. (2006). Effects of media violence on health-related outcomes among young men. *Archives of Pediatrics & Adolescent Medicine, 160*(4), 341–347. https://doi.org/10.1001/archpedi.160.4.341

Brown, N. J. L., & Heathers, J. A. J. (2017). The GRIM test: A simple technique detects numerous anomalies in the reporting of results in psychology. *Social Psychological and Personality Science, 8*(4), 363–369. https://doi.org/10.1177/1948550616673876

Cangas, A. J., Navarro, N., Parra, J., Ojeda, J. J., Cangas, D., Piedra, J. A., et al. (2017). Stigma-Stop: A serious game against the stigma toward mental health in educational settings. *Frontiers in psychology, 8*, 1385. https://doi.org/10.3389/fpsyg.2017.01385

Carnagey, N. L. (2006). Is it competitiveness or violent content? The effects of violent sports video games on aggression. *Iowa State University*. 10.31274/rtd-180813-13158.

[2] The following studies were included in the meta-analysis: Alblas et al. (2018), Barthel (2013), Barwood et al. (2020), Behm-Morawitz et al. (2016), Bluemke et al. (2010), Brady and Matthews (2006), Cangas et al. (2017), Carnagey (2006), Chen et al. (2020), Fiellin et al. (2017), Greer et al. (2017), Hammer (2014), Hung et al. (2009), Jacobs (2018), Kampf (2015), Kampf (2015), Kampf (2016), Kampf (2017), Kolek et al. (2021), Květon and Jelínek (2019), Parisod et al. (2018), Peña and Hernández Pérez (2019), Peña et al. (2018), Potter et al. (2021), Potter et al. (2019), Price et al. (2016), Quick et al. (2013), Ruggiero (2015), Saleem (2008), Saleem and Anderson (2013), Sharma et al. (2015), Soekarjo and Oostendorp (2015), Teng et al. (2011), Waddington et al. (2013), Yang et al. (2014).

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for Bias in Psychology: A Comparison of Meta-Analytic Methods. *Advances in Methods and Practices in Psychological Science, 2*(2), 115–144. https://doi.org/10.1177/2515245919847196

Case, D. O., & Given, L. M. (2016). *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior* (4th edition). Bingley: Emerald Group Publishing Limited.

Chapman, A. (2016). Digital games as history: How videogames represent the past and offer access to historical practice. *Routledge*. https://doi.org/10.4324/9781315732060

Chen, T., Stewart, M., Bai, Z., Chen, E., Dabbish, L., & Hammer, J. (2020). Hacked Time: Design and Evaluation of a Self-Efficacy Based Cybersecurity Game. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (pp. 1737–1749). https://doi.org/10.1145/3357236.3395522

Crano, W. D., & Prislin, R. (Eds.). (2008). *Frontiers of social psychology. Attitudes and attitude change.* Psychology Press.

Downs, E. (2014). Driving home the message: Using a video game simulator to steer attitudes away from distracted driving. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 6*(1), 50–63. https://doi.org/10.4018/ijgcms.2014010104

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich college publishers.

Epskamp, S., & Nuijten, M. B. (2018). statcheck: Extract statistics from articles and recompute p-values (1.3.1) [R package]. Retrieved from https://cran.r-project.org/web/packages/statcheck/index.html.

Festinger, L. (1958). The motivating effect of cognitive dissonance. In G. Lindzey (Ed.), *Assessment of human motives* (pp. 65–86). New York: Holt, Rinehart and Winston.

Fiellin, L. E., Hieftje, K. D., Pendergrass, T. M., Kyriakides, T. C., Duncan, L. R., Dziura, J. D., et al. (2017). Video Game Intervention for Sexual Risk Reduction in Minority Adolescents: Randomized Controlled Trial. *Journal of Medical Internet Research, 19*, e314.

Fullerton, T. (2018). *Game design workshop: A playcentric approach to creating innovative games.* CRC Press.

Garneli, V., Giannakos, M., & Chorianopoulos, K. (2017). Serious games as a malleable learning medium: The effects of narrative, gameplay, and making on students' performance and attitudes. *British Journal of Educational Technology, 48*(3), 842–859. https://doi.org/10.1111/bjet.12455

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2007). Unraveling the processes underlying evaluation: Attitudes from the perspective of the APE model. *Social Cognition, 25*(5), 687–717. https://doi.org/10.1521/soco.2007.25.5.687

Gawronski, B., & Bodenhausen, G. V. (2014). Implicit and Explicit Evaluation: A Brief Review of the Associative-Propositional Evaluation Model. *Social and Personality Psychology Compass, 8*, 448–462. https://doi.org/10.1111/spc3.12124

Grayson, N. (2020, January 8). Annual Steam Release Numbers Used to Climb Every Year, But in 2019, They Held Steady. Kotaku. Retrieved 2020, March 2, from https://kotaku.com/annual-steam-release-numbers-used-to-climb-every-year-1840878072.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of personality and social psychology, 74*(6), 1464. https://doi.org/10.1037/0022-3514.74.6.1464

Greer, M., Lin, L., & Atkinson, R. K. (2017). Using a computer game to teach school-aged children about asthma. *Interactive Learning Environments, 25*(4), 431–438. https://doi.org/10.1080/10494820.2015.1135469

Hammer, J. (2014). Playing Prejudice: The Impact of Game-Play on Attributions of Gender and Racial Bias. *Columbia University*. https://doi.org/10.7916/D8BP00ZK

Harrer, M., Cuijpers, P., Furukawa, T.A, & Ebert, D. D. (2019). Doing meta-analysis in R: A hands-on guide. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/.

Hong, S., & Reed, W. R. (2021). Using Monte Carlo experiments to select meta-analytic estimators. *Research Synthesis Methods, 12*(2), 192–215. https://doi.org/10.1002/jrsm.1467

Hornung, R. L., Lennon, P. A., Garrett, J. M., DeVellis, R. F., Weinberg, P. D., & Strecher, V. J. (2000). Interactive computer technology for skin cancer prevention targeting children. *American Journal of Preventive Medicine, 18*(1), 69–76. https://doi.org/10.1016/S0749-3797(99)00115-4

Hung, C., Chiu, C. H., Chen, Y., Su, M., & Chen, H. (2009). Effectiveness of Game-based Learning of a National Health e-Learning Network for nutrition education in elementary school. In *2009 11th International Conference on e-Health Networking, Applications and Services (Healthcom)* (pp. 184–186). https://doi.org/10.1109/HEALTH.2009.5406187

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology, 19*(5), 640–648. https://doi.org/10.1097/EDE.0b013e31818131e7

Jackson, L. C., O'Mara, J., Moss, J., & Jackson, A. C. (2018). A critical review of the effectiveness of narrative-driven digital educational games. *International Journal of Game-Based Learning (IJGBL), 8*(4), 32–49.

Jacobs, R. S. (2018). Play to win over: Effects of persuasive games. *Psychology of Popular Media Culture, 7*(3), 231–240. https://doi.org/10.1037/ppm0000124

Kampf, R. (2015). Computerized Simulations of the Israeli-Palestinian Conflict, Knowledge Acquisition and Attitude Change: PeaceMaker vs Global Conflicts. *Proceedings of the European Conference on Games Based Learning, 2015*(January), 269–275.

Kampf, R. (2016). Long-Term Effects of Computerized Simulations in Protracted Conflicts: The Case of Global Conflicts. In *Persuasive Technology* (pp. 242–247). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-31510-2_21.

Kampf, R. (2016). The Context Is the Message?! Learning about the Israeli-Palestinian Conflict through Computerized Simulations in Different National Contexts. *Universal Journal of Educational Research, 4*(8), 1787–1796. https://doi.org/10.13189/ujer.2016.040807

Kampf, R. (2017). Role-Taking and Attitude Change in a Computerized Simulation of the Israeli-Palestinian Conflict: Direct Party vs. Distant Party. In *Proceedings of the 12th Chais Conference for the Study of Innovation and Learning Technologies: Learning in the Technological Era* (pp. 45–51). The Open University of Israel.

Kolek, L., Ropovik, I., Sisler, V., van Oostendorp, H., & Brom, C.. Video Games and Attitude Change: A Meta-analysis. Open Science Framework. https://osf.io/4aeqt/.

Kolek, L., Šisler, V., Martinková, P., & Brom, C. (2021). Can video games change attitudes towards history? Results from a laboratory experiment measuring short-and long-term effects. *Journal of Computer Assisted Learning, 37*(5), 1348–1369. https://doi.org/10.1111/jcal.12575

Krosnick, J. A., & Alwin, D. F. (1989). Aging and susceptibility to attitude change. *Journal of Personality and Social Psychology, 57*(3), 416–425. https://doi.org/10.1037/0022-3514.57.3.416

Květon, P., & Jelínek, M. (2019). Frustration and violence in mobile video games: An experimental evaluation of their effect on implicit aggression. *Swiss Journal of Psychology, 79*(2), 63–70. https://doi.org/10.1024/1421-0185/a000236

Lin, M. C., Tutwiler, M. S., & Chang, C. Y. (2011). Exploring the relationship between virtual learning environment preference, use, and learning outcomes in 10th grade earth science students. *Learning, Media and Technology, 36*(4), 399–417. https://doi.org/10.1080/17439884.2011.629660

Lüdecke, D. (2017). esc: Effect Size Computation for Meta Analysis. R package version 0.4.0. Retrieved from https://CRAN.R-project.org/package=esc.

Ludopium GmbH - Arte France. (2018). *Vectronom.* Retrieved from https://vectronom.arte.tv/.

Maier, J., & Richter, T. (2013). Text-belief consistency effects in the comprehension of multiple texts with conflicting information. *Cognition and Instruction, 31*(2), 151–175. https://doi.org/10.1080/07370008.2013.769997

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis: An evaluation of selection methods and some cautionary notes: An evaluation of selection methods and some cautionary notes. *Perspectives on Psychological Science, 11*(5), 730–749. https://doi.org/10.1177/1745691616662243

Newzoo (2020). Global Games Market Report 2020. Retrieved from https://newzoo.com/products/reports/global-games-market-report/.

Nicklin, H. (2022). *Writing for Games: Theory and Practice* (1st ed.). CRC Press. https://doi.org/10.1201/9781003182832.

Nintendo Switch. (2019). *Cadence of Hyrule: Crypt of the NecroDancer Featuring The Legend of Zelda.* Retrieved from https://www.nintendo.com/store/products/cadence-of-hyrule-crypt-of-the-necrodancer-featuring-the-legend-of-zelda-switch/.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical research ed.), 372*, Article n71. https://doi.org/10.1136/bmj.n71

Parisod, H., Pakarinen, A., Axelin, A., Löyttyniemi, E., Smed, J., & Salanterä, S. (2018). Feasibility of mobile health game "Fume" in supporting tobacco-related health literacy among early adolescents: A three-armed cluster randomized design. *International journal of medical informatics, 113*, 26–37. https://doi.org/10.1016/j.ijmedinf.2018.02.013

Peña, J., & Hernández Pérez, J. F. (2019). Game perspective-taking effects on willingness to help immigrants: A replication study with a Spanish sample. *New Media & Society, 22*(6), 944–958. https://doi.org/10.1177/1461444819874472

Pena, J., Hernández Pérez, J. F., Khan, S., & Cano Gómez, Á. P. (2018). Game Perspective-Taking Effects on Players' Behavioral Intention, Attitudes, Subjective Norms, and Self-Efficacy to Help Immigrants: The Case of "Papers, Please". *Cyberpsychology, Behavior, and Social Networking, 21*(11), 687–693. https://doi.org/10.1089/cyber.2018.0030

Potter, S. J., Demers, J. M., Flanagan, M., Seidman, M., & Moschella, E. A. (2021). Can video games help prevent violence? An evaluation of games promoting bystander intervention to combat sexual violence on college campuses. *Psychology of Violence, 11*(2), 199–208. https://doi.org/10.1037/vio0000365

Potter, S. J., Flanagan, M., Seidman, M., Hodges, H., & Stapleton, J. G. (2019). Developing and Piloting Videogames to Increase College and University Students' Awareness and Efficacy of the Bystander Role in Incidents of Sexual Violence. *Games for health journal, 8*(1), 24–34. https://doi.org/10.1089/g4h.2017.0172

Pötzsch, H., & Šisler, V. (2019). Playing cultural memory: Framing history in call of duty: Black ops and Czechoslovakia 38–89: Assassination. *Games and Culture, 14*(1), 3–25. https://doi.org/10.1177/1555412016638603

Pratkanis, A. R. (1989). The cognitive representation of attitudes. In R. S. Pratkanis, J. Breckler, & A. G. Greenwald (Eds.), *The Third Ohio State University, Vol. on Attitudes and Persuasion. Attitude Structure and Function* (pp. 71–98). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Priester, J., Wegener, D., Petty, R., & Fabrigar, L. (1999). Examining the psychological process underlying the sleeper effect: The elaboration likelihood model explanation. *Media Psychology, 1*(1), 27–48. https://doi.org/10.1207/s1532785xmep0101_3

Price, C. A., Gean, K., Christensen, C. G., Beheshti, E., Pernot, B., Segovia, G., et al. (2016). Casual Games and Casual Learning About Human Biological Systems. *Journal of Science Education and Technology, 25*, 111–126. https://doi.org/10.1007/s10956-015-9580-6

Pustejovsky, J. (2020). clubSandwich: Cluster-Robust (Sandwich) Variance Estimators with Small-Sample Corrections. R package version 0.5.3 [Computer software]. Retrieved from https://github.com/jepusto/clubSandwich.

Pustejovsky, J. E., & Tipton, E. (2020). Meta-analysis with Robust Variance Estimation: Expanding the range of working models. *In MetaArXiv*. https://osf.io/mq9hj/.

Quick, V., Corda, K. W., Chamberlin, B. A., Schaffner, D. W., & Byrd-Bredbenner, C. (2013). Ninja Kitchen to the Rescue: Evaluation of a Food Safety Education Game for Middle School Youth. *British Food Journal, 115*(5), 686–699. https://doi.org/10.1108/00070701311331481

Revelle, W. (2018). psych: Procedures for personality and psychological research, Northwestern University, Evanston, IL, USA. Retrieved from https://cran.r-project.org/web/packages/psych/index.html psych Version = 1.8.4.

Ropovik, I., Adamkovic, M., & Greger, D. (2021). Neglect of publication bias compromises meta-analyses of educational research. *PLoS One1, 16*(6), e0252415.

Ruggiero, D. (2015). The effect of a persuasive social impact game on affective learning and attitude. *Computers in Human Behavior, 45*, 213–221. https://doi.org/10.1016/j.chb.2014.11.062

Saleem, M. (2008). Effects of stereotypic video game portrayals on implicit and explicit attitudes. *Iowa State University*. 10.31274/rtd-180813-16673.

Saleem, M., & Anderson, C. A. (2013). Arabs as terrorists: Effects of stereotypes within violent contexts on attitudes, perceptions, and affect. *Psychology of Violence, 3*(1), 84. https://doi.org/10.1037/a0030038

Sampson, M., Zhang, L., Morrison, A., et al. (2006). An alternative to the hand searching gold standard: Validating methodological search filters using relative recall. *BMC Medical Research Methodology, 6*, 33. https://doi.org/10.1186/1471-2288-6-33

Sanbonmatsu, D. M., & Fazio, R. H. (1990). The role of attitudes in memory-based decision making. *Journal of Personality and Social Psychology, 59*, 614–622. https://doi.org/10.1037/0022-3514.59.4.614

Sharma, S. V., Shegog, R., Chow, J., Finley, C., Pomeroy, M., Smith, C., et al. (2015). Effects of the Quest to Lava Mountain Computer Game on Dietary and Physical Activity Behaviors of Elementary School Children: A Pilot Group-Randomized Controlled Trial. *Journal of the Academy of Nutrition and Dietetics, 115*(8), 1260–1271. https://doi.org/10.1016/j.jand.2015.02.022

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General, 143*(2), 534–547. https://doi.org/10.1037/a0033242

Smethurst, T., & Craps, S. (2014). Playing with Trauma: Interreactivity, Empathy, and Complicity in The Walking Dead Video Game. *Games and Culture, 10*(3), 269–290. https://doi.org/10.1177/1555412014559306

Soekarjo, M., & van Oostendorp, H. V. (2015). Measuring Effectiveness of Persuasive Games Using an Informative Control Condition. *International Journal of Serious Games, 2*(2), 37–56. https://doi.org/10.17083/ijsg.v2i2.74

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

Stanley, T. D., Doucouliagos, H., & Ioannidis, J. P. A. (2017). Finding the power to reduce publication bias: Finding the power to reduce publication bias. *Statistics in Medicine, 36*(10), 1580–1598. https://doi.org/10.1002/sim.7228

Statista (2021a, February 1). Number of games released on Steam worldwide from 2004 to 2020. Retrieved 2021, May 24, from https://www.statista.com/statistics/552623/number-games-released-steam/.

Statista (2021b, February 4). Number of apps released worldwide between 3rd quarter 2019 and 2nd quarter 2020, by category, Retrieved 2021, May 24, from https://www.statista.com/statistics/373529/new-app-releases-category/.

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., et al. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ (Clinical Research Ed.), 366*, Article l4898. https://doi.org/10.1136/bmj.l4898

Strawhacker, A., Bers, M., Verish, C., Sullivan, A., & Shaer, O. (2018). Enhancing children's interest and knowledge in bioengineering through an interactive videogame. *Journal of Information Technology Education: Innovations. Practice, 17*(1), 55–81. https://doi.org/10.28945/3976

Sýkora, T., Stárková, T., & Brom, C. (2021). Can narrative cutscenes improve home learning from a math game? An experimental study with children. *British Journal of Educational Technology, 52*(1), 42–56.

Teng, S. K. Z., Chong, G. Y. M., Siew, A. S. C., & Skoric, M. M. (2011). Grand Theft Auto IV comes to Singapore: Effects of repeated exposure to violent video games on aggression. *Cyberpsychology, Behavior, and Social Networking, 14*(10), 597–602. https://doi.org/10.1089/cyber.2010.0115

Todd, A. R., & Galinsky, A. D. (2014). Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass, 8*, 374–387. https://doi.org/10.1111/spc3.12116

Tyler, T. R., & Schuller, R. A. (1991). Aging and attitude change. *Journal of Personality and Social Psychology, 61*(5), 689–697. https://doi.org/10.1037/0022-3514.61.5.689

van Aert, R. C. M., & van Assen, M. A. L. M. (2018). Correcting for publication bias in a meta-analysis with the P-uniform* method. *In MetaArXiv*. 10.31222/osf.io/zqjr9.

van Strien, J. L. H., Kammerer, Y., Brand-Gruwel, S., & Boshuizen, H. P. A. (2016). How attitude strength biases information processing and evaluation on the web. *Computers in Human Behavior, 60*, 245–252. https://doi.org/10.1016/j.chb.2016.02.057

Vevea, J. L., & Woods, C. M. (2005). Publication bias in research synthesis: Sensitivity analysis using a priori weight functions. *Psychological Methods, 10*(4), 428–443. https://doi.org/10.1037/1082-989x.10.4.428

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1–48. https://doi.org/10.18637/jss.v036.i03

Vogel, T., & Wanke, M. (2016). *Attitudes and Attitude Change* (2nd ed). New York: Routledge.

Waddington, D. I., Thomas, T., Venkatesh, V., Davidson, A. L., & Alexander, K. (2013). Education from inside the bunker: Examining the effect of Defcon, a nuclear warfare simulation game, on nuclear attitudes and critical reflection. *Loading..., 7*(12).

Wang, M., & Chen, Y. (2006). Age differences in attitude change: Influences of cognitive resources and motivation on responses to argument quantity. *Psychology and Aging, 21*(3), 581–589. https://doi.org/10.1037/0882-7974.21.3.581

Wickham, et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Williams, A. J., & Williams, R. H. (2011). Multiple identification theory: Attitude and behavior change in a simulated international conflict. *Simulation & Gaming, 42*(6), 733–747. https://doi.org/10.1177/1046878111429228

Yang, G. S., Gibson, B., Lueke, A. K., Huesmann, L. R., & Bushman, B. J. (2014). Effects of avatar race in violent video games on racial attitudes and aggression. *Social Psychological and Personality Science, 5*(6), 698–704. https://doi.org/10.1177/1948550614528008

Zhu, D., Thomson, C. L., & Hsu, Y. (2012). Using the technology acceptance model to evaluate user attitude and intention of use for online games. *Total Quality Management & Business Excellence, 23*(7–8), 965–980. https://doi.org/10.1080/14783363.2012.704269