Can Video Games Change Attitudes towards History? Results from a Laboratory Experiment Measuring Short- and Long-term Effects

Authors:

Lukáš Kolek (corresponding author) [a]

Vít Šisler [a]

Patrícia Martinková [bc]

Cyril Brom [d]

Affiliations

[a] Faculty of Arts, Charles University, Jan Palach Square 1/2, 116 38 Prague, Czech Republic, email: Lukáš Kolek kolek@ksvi.mff.cuni.cz / email: Vít Šisler vit.sisler@ff.cuni.cz

[b] Faculty of Education, Charles University, Magdalény Rettigové 4, 116 39 Prague, Czech Republic, email: martinkova@cs.cas.cz

[c] Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 271/2, 182 07 Prague, Czech Republic

[d] Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Prague, Czech Republic, email: brom@ksvi.mff.cuni.cz

Note

This paper has been **ACCEPTED** for publication in the **Journal of Computer Assisted Learning**
First published: 04 August 2021

https://doi.org/10.1111/jcal.12575

Abstract

This study investigates a video game's effects on implicit and explicit attitudes towards depicted historical events in the short- and long-term on a sample of 148 young adults. We used, as an intervention tool, a serious game *Czechoslovakia 38–89: Borderlands* that deals with the expulsion of the Sudeten Germans from the former Czechoslovakia after the WWII. Results showed more negative pretest-posttest explicit attitude changes towards the expulsion on a general level (d = −0.34) and a specific level (d = −0.53) compared to the control group. Over the long-term, group differences in attitude change remained significant for the specific level (d = −0.44), but not for general one (d = −0.16). Exploratory analysis on the item level indicated that especially attitudes towards the expulsion's (un)fairness were affected by the game. However, no significant changes were found in implicit attitudes in the experimental group. This study is the first of such scale to empirically investigate video games' effects on a society's historical awareness.

Keywords

Video games; media in education; game-based learning; explicit attitudes; implicit attitudes; history representation, differential item functioning

Structured Abstract

Background

- Games focusing on recent and still sensitive historical events are increasingly popular. However, the empirical research about the effects of historical video games or video games in general on attitudes towards the topics depicted in game narratives is limited. For instance, virtually nothing is known about historical games' differential impact on implicit or explicit attitude evaluations towards the topics represented in such games.

Objectives

- As the first study of such a scale, this study investigates a video game's effects on implicit and explicit attitudes towards depicted historical events in the short- and long-term.

Methods

- In a between-subjects design study, 148 young adults played an intervention game or a control game for about 50 min. The intervention game was the serious game *Czechoslovakia 38-89: Borderlands* that deals with the expulsion of the Sudeten Germans from the former Czechoslovakia after the WWII. We measured participants' explicit and implicit evaluations of the event (pretest, posttest, one month delayed posttest).

Results and Conclusions

- Our intervention game changed explicit attitudes in the short- and long-term, but no significant changes were found in implicit attitudes after the intervention.

Major Takeaways

- Our results confirm that video games can affect a society's historical awareness. As video games have become a worldwide phenomenon affecting whole societies regardless of social status or age, this study offers new perspectives for understanding and interpretation of the formation of historical narratives in informal settings.

Practitioner notes

What is currently known about the subject matter

- No studies address video games' long-term effects on both explicit and implicit attitudes.
- The research of video games' effects on short- and long-term attitudes is generally limited.
- Video games can change short-term explicit attitudes under certain circumstances.
- So far, the effect of video games on implicit attitudes has not been confirmed.
- There are no studies dealing with the effect of historical video games on attitudes.

What our paper adds to this

- The study is the first to combine long-term explicit and implicit attitude measurements related to games.
- Our video game intervention affected explicit attitudes over the short term and partly over the long term.
- We have not confirmed the effect of our video game intervention on implicit attitude change.
- Based on our results, historical video games can affect short- and long-term explicit attitudes.

The implications of study findings for practitioners

- Our study confirms that perspective-taking can change players' explicit attitudes.
- We demonstrate that exploratory DIF-C analysis can provide deeper understanding of the explicit attitude changes on the item level.
- Our results confirm that video games can affect a society's historical awareness.
- Our results suggest video games have an important role in our understanding of historical and social realities.

1. Introduction

Video games reach broad segments of the population and depict various areas of human life, including historical topics. Games dealing with, or inspired by, historical events were among the top selling titles in 2019 (e.g., Total War: Three Kingdoms, Assassins Creed: Odyssey, Civilization VI, and Kingdome Come: Deliverance) (Valve Corporation, 2019). Games focusing on recent and still sensitive historical events are increasingly popular, as is evident from the commercial success of This War of Mine or media coverage for the recently released Through the Darkest of Times (de Smale et al., 2019; Paintbucket Games, 2019; Yin-Poole, 2014). Can we consider history-focused video games to be agents that shape our historical awareness and attitudes?

Several scholars suggest the importance of history-focused video games for the formation of historical awareness, especially among the younger generation (Chapman, 2016; Kapell & Elliott, 2013; Pötzsch & Šisler, 2019). Players often spend dozens of hours enjoying these games where they are exposed to real or fictitious historical narratives that invite certain understandings of the past as chosen by game designers (Atkins, 2003).[1] Furthermore, games provide players with agency needed to interact with the represented historical narratives; thus allowing them to challenge those representations and form their own conclusions (Pötzsch & Šisler, 2019).

In this study, we distinguish between two types of attitude evaluations to interpret attitude changes—explicit and implicit attitude evaluations. The differences and relationships between them can be explained by the Associative-Propositional Evaluation model (APE model; Gawronski & Bodenhausen, 2006, 2007, 2014). According to this model, explicit attitudes are based on

[1] Average playtimes for games are not public information, but unofficial estimates from *Steam* data (including only PC and Mac players) are available. For example, according to *AStats.nl*, the average playtime on 9 June 2020 for *Civilization VI* was 43.0 h and the average playtime for *Assassins Creed Origins* was 54.6 h.

propositional reasoning, which reflects evaluative judgements derived from information related to the object in question (Bohner & Dickel, 2011; Gawronski & Bodenhausen, 2014). Implicit attitudes are derived from associative evaluations, which are immediate affective evaluative reactions to an object based on the object's relatability to, or familiarity with, other concepts in our memory (ibid.).

There is extensive empirical research about the effect of narratives on attitudes within various non-interactive formats: such as written text, radio or film (see Green et al., 2019). Less is known about the effects of narrative video games on attitudes towards the topics depicted in them. In particular, empirical research about the effect of historical video games or video games in general on attitudes towards the topics in game narratives has shortcomings and large gaps in knowledge exist. For instance, virtually nothing is known about historical games' differential impact on implicit or explicit attitude evaluations towards the topics represented in such games, and thus about the effect of games on our formation of historical awareness.

The vast majority of studies about video games and attitudes focuses on short-term change of explicit attitudes towards the depicted topics (e.g., Cuhadar & Kampf, 2014, 2015; Hawkins et al., 2019; Peña et al., 2018; Pentz et al., 2019). On the contrary, little is known about long-term changes and about changes in implicit attitudes in relation to video games. We found only one study using pretest-posttest design with a control group examining implicit attitudes (Alblas et al., 2018). Also, as far as we know, there are only two studies of video games that have examined long-term explicit attitude evaluations and which have a control condition (Kampf, 2016b; Ruggiero, 2015). To our knowledge, no studies have examined long-term implicit attitude change, nor both explicit and implicit attitude change with a control condition over the long-term in relation to the

topics represented in video games. There is also no empirical study focusing on the effect of video games on attitudes towards the history represented in a video game.

The present study addresses the research gap by offering missing empirical evidence about historical video games' effects on their players' explicit and implicit attitude evaluations towards the represented topic over the long-term. To do so, this study uses as an intervention tool the historical adventure video game Czechoslovakia 38–89: Borderlands. The game depicts the expulsion of the Sudeten Germans from the Czechoslovak borderlands after WWII through perspective-taking, that is, from the perspective of the event's various actors. The paper proceeds as follows. Section 2 elaborates on the study background, we further describe the approach to explicit and implicit attitudes (Section 2.1), we discuss it in the context of narrative video games (Section 2.2) and of the existing empirical studies focusing on video games and attitudes (Section 2.3). We then describe the study design and hypotheses in Section 3, the methodology including our intervention games in Section 4, and results in Section 5. The discussion of the findings and its implications are offered in Section 6 and conclusion is provided in Section 7.

2. Study Background

2.1. Attitudes and the Associative-Propositional Evaluation model

Possessing an attitude towards an object allows us to avoid “the energy-consuming and sometimes painful process of figuring out de novo how we shall relate ourselves to [the object]” (Smith et al., 1956, p. 41). As Vogel and Wanke (2016, p.3) argued, attitude is “a summary evaluation of an object of thought”. Its objects may be abstract or concrete. Attitudes affect our information processing and behaviour (Brannon et al., 2007; Pratkanis, 1989; van Strien et al., 2016; Vogel & Wanke, 2016).

One group of scholars considers attitudes to be stable over time (e.g., Eagly & Chaiken, 2007; Fazio, 2007; Petty et al., 2007). According to them, attitudes are retrieved from information in our long-term memory whenever we evaluate an object. Another group of scholars perceives attitudes to be a concept created "on the spot." Such attitudes are dependent both on information retrieved from our long-term memory and on context-dependent information people have at hand during the evaluation (e.g., Gawronski & Bodenhausen, 2007; Schwarz, 2007).

For a long time, researchers accessed attitudes mostly through self-report scales. Self-reports examine participants' explicit (deliberate) attitude evaluations towards the measured phenomena through introspection. Recently, researchers have also begun to examine implicit attitudes (Greenwald et al., 1998; Greenwald & Banaji, 1995). Implicit attitudes are most commonly accessed through measurements of response times in categorization tasks (Greenwald et al., 1998; Greenwald & Banaji, 1995). According to Gawronski and Brannon (2019), the measurement of both implicit and explicit attitudes allows for "deeper insights into the processes underlying evaluative judgements, the processes by which attitudes influence behaviour, and the processes underlying attitude formation and change "(pp. 188). Explicit and implicit attitudes are guided by different, and often interplaying, processes as described by the Associative-Propositional Evaluation model (APE model; Gawronski & Bodenhausen, 2006, 2007, 2014).

The dual process APE model assumes that attitudes are derived from associative evaluations and propositional reasoning (see Figure 1). Both processes are qualitatively distinct yet interconnected. Within the model (Gawronski & Bodenhausen, 2006, 2014), associative evaluations can be examined through implicit attitude measurements. Associative evaluations form the basis for implicit attitudes and represent a spontaneous affective response to the object (e.g., positive evaluations recalled from memory when exposed to activities related to a healthy lifestyle). The model posits that associative evaluations are "defined as the activation of mental

associations on the basis of feature similarity and spatio-temporal contiguity." (Gawronski & Brannon, 2019, pp. 167). In other words, implicit attitudes are derived from associative evaluations, which are immediate affective reactions to the object based on the object's relatability or familiarity with other concepts in our memory. The APE model assumes the existence of a structure for these mental associations in the long-term memory (Gawronski & Payne, 2010). This structure may be changed by frequent co-occurrence of two concepts in one's environment. The co-occurrence either strengthens an associative link (if it already exists) or creates a new one. Stronger links are easier to activate (Gawronski & Brannon, 2019; Gawronski & Payne, 2010). Also, associative evaluations are highly contingent on the "on the spot" context-dependent information for a decision on which of these stored associations will be activated; that is, recalled from our long-term memory. For example, exposing participants to pictures of successful African-Americans prior to an implicit attitude test measuring racial attitudes will result in temporary more favourable attitudes towards African-Americans (see, e.g., Banaji & Greenwald, 2013). Associative evaluations function independently in relation to what one consciously considers to be the truth.

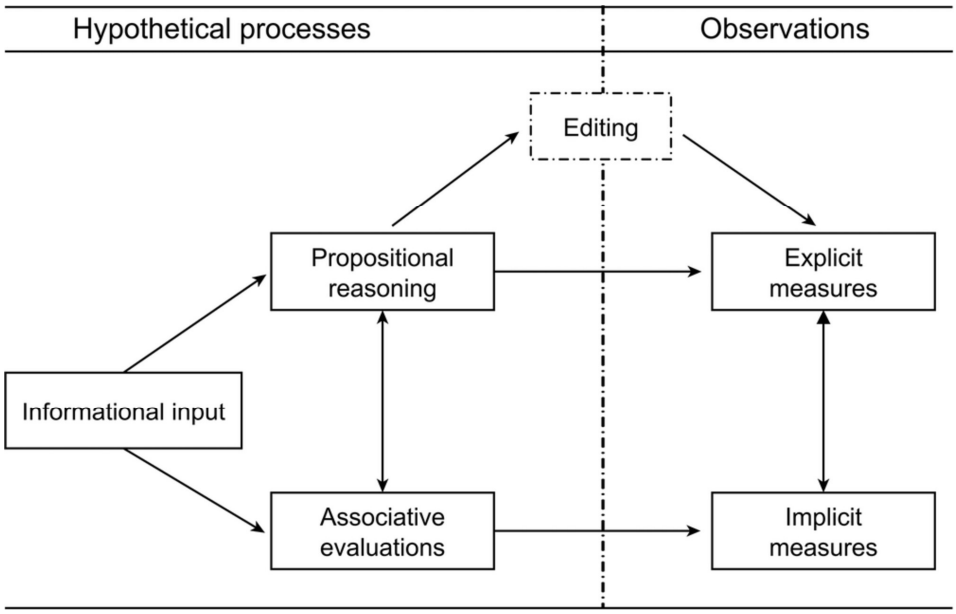

Figure 1. Associative evaluations and propositional reasoning in the APE model (Bohner & Dickel, 2011)

Within the APE model, propositional reasoning is deliberate and forms the basis for explicit attitudes (Bohner & Dickel, 2011). It functions as the validation of information suggested by associative evaluations (Bohner & Dickel, 2011; Gawronski & Bodenhausen, 2014; Gawronski & Brannon, 2019). It is based on logical conclusions derived from information related to the object in question (e.g., exercising is healthy, thus good). The conclusions therefrom are formed based on their consistency with other relevant conclusions related to the object that are stored in our short- and long-term memory. Unlike associative evaluations, propositional reasoning must be in line with what one consciously considers to be truth in relation to knowledge at his/her disposal. One cannot have two contradictory propositional reasonings about the same object over the long-term (Bohner & Dickel, 2011; Gawronski & Bodenhausen, 2014). Therefore, if new propositional reasoning is not consistent with the older one, it will lead to cognitive dissonance and subsequent motivation to restore the consistency. This can happen by rejection of one of the propositions or by acquisition of new information to resolve the created dissonance (Festinger, 1958) resulting in

a belief revision. However, changes in propositional reasoning may not be always reflected in explicit attitude measurements because self-reports (used to assess them) are limited by participants' ability and willingness to share attitudes. This limitation is depicted in Figure 1 by the 'Editing' process.

The main focus of this study is on players' changing evaluations towards the depicted concepts in games due to their play. In this context, attitudes play a significant role when someone processes complex information (Sanbonmatsu & Fazio, 1990); they (attitudes) influence our information selection and the way we interpret obtained information (Case & Given, 2016; Pratkanis, 1989; Vogel & Wanke, 2016); they have also an effect on our evaluation of the credibility of our information sources (van Strien et al., 2016) and on the time we devote to a particular source (Brannon et al., 2007; van Strien et al., 2016). At the same time, information processing forms the core mechanism of attitude formation and attitude change (Pratkanis, 1989; Vogel & Wanke, 2016). As such, attitudes represent an essential concept for understanding mechanisms involved in informal learning and the respective roles of video games in it.

2.2. Narrative Video Games and Attitude Change

Our empirical study focuses on narrative-driven video games. Game narratives are carriers of meaning having the potential to shape players' attitudes towards the story learned or experienced (Green & Jenkins, 2014; Maio & Haddock, 2010; Steinemann et al., 2017).

Similar to some other pop culture media, narrative video games are interactive in the sense that players can shape game narratives and game worlds, perceive the outcomes of this interaction, and also react to these outcomes (Smethurst & Craps, 2014). Thus, the relationship between the players and the game narrative is dynamic: It provides them with agency and an active role in influencing the game environment. Narratives based on real experiences or narratives that are

credible enough to actually happen can effectively influence attitudes (Green & Brock, 2002). Empirical evidence suggests that the more a person has invested in the story, the more difficult it becomes to consider counterarguments to those presented in the narrative (Dal Cin et al., 2004; Green & Brock, 2000, 2002). Narrative games could thus affect explicit and implicit attitude evaluations. Seeing and influencing the consequences of their own actions may create associative links between the narrative and the outcome for players. At the same time, historical games provide players with representations of certain historical events. Players may then challenge this representation in the game in a meaningful way and thus challenge their propositional reasoning related to the event. Narrative games have unconvincing results in relation to cognitive learning outcomes (e.g., Adams et al., 2012; Sýkora et al., 2021), but they may have an effect on attitudes.

2.3. Studies about Video Games and Attitude Change

Despite growing empirical research on attitudes, the role of video games in attitudinal change has not been explored much (see Soekarjo & Oostendorp, 2015). Most of the empirical studies that deal with the subject (e.g., Cuhadar & Kampf, 2014, 2015; Hawkins et al., 2019; Jacobs, 2018; Knol & de Vries, 2011; Tragazikis & Meimaris, 2009) measured attitudes primarily through self-report scales revealing short-term changes in participants' explicit attitudes. We have identified only a few studies which were simultaneously (1) focused on attitude change towards the depicted narrative content and (2) not possessing any external element to the game affecting participants within the research design of the study, for example, driving a car in a video game while making a phone call in a real life and (3) having a control condition. From the perspective of attitude change research, the absence of a control group is problematic as we cannot evaluate the effect of measurement on results; especially as attitudes may be, to various extents, dependent on the information accessible at the particular moment (see Section 2.1).

Concerning short-term explicit attitude change, we have found only five studies with a control condition presenting mixed findings. On the one hand, Kampf's (2015, Exp. 1, Exp. 2; 2016a) three studies confirmed the ability of two different games to affect players' attitudes towards the Israeli-Palestinian conflict. However, all three studies focused explicitly on students of political science and communications. Also, despite having control conditions, two of these studies (Kampf, 2015, Exp. 1, Exp 2.) did not share their results beyond stating that attitudes did not change in the control conditions. Furthermore, a study by Price et al. (2015) confirmed small between group differences concerning the game's ability to change attitudes using two Likert scales, however one of those scales confirmed the effect only due to the shift in the control group. On the other hand, the study by Peña et al. (2018) did not find any significant between-group difference in the experimental and the control group.

Concerning long-term explicit attitude change, only a few studies examined it using self-report scales (e.g., Duncan et al., 2018; Smith et al., 2017). To our knowledge, only two of those with long-term measures included a control condition (Kampf, 2016b; Ruggiero, 2015). Both studies confirmed long-term explicit attitude change: the former towards the Israeli-Palestinian conflict, the latter towards homelessness. However, the latter confirmed the game's effect on attitude change only due to a significant shift in the control condition.

Since the introduction of implicit attitude measurements, studies measuring both implicit and explicit attitude change have started to appear; but they are still rare. The lack of studies using implicit measurements may be due to the limited availability of free software for measuring implicit attitudes and the relatively demanding requirements for the analysis of obtained data (Gerling et al., 2014). Existing studies generally lack data from a control condition (e.g., Alhabash & Wise, 2015) or they focus on college students or psychology students in a posttest-only design

(e.g., Saleem & Anderson, 2013; Yang et al., 2014). The only study with a pretest-posttest design with a control group (Alblas et al., 2018) dealing with implicit attitudes (not explicit attitudes) found significant group difference, which was caused, however, by a significant change in the control group, not in the experimental group. To our knowledge, there are no studies using implicit attitude measurements over the long term, and therefore no long-term studies combining explicit and implicit attitude measurements. At the same time, no studies focus on attitude change towards depicted historical events in games.

3. This Study

3.1. Study Design

This study investigates a historical video game's impact on attitude change towards the historical events depicted in the game using explicit and implicit attitude measurements. It has pretest-posttest design with a delayed posttest 1 month after the intervention; the study has one control and one experimental group. The outcomes of our research are applicable to narratives and narrative game research and their effects on players' attitudes, but the results also contribute to research on attitude change in general.

In the experimental group, we use the modified historical game *Czechoslovakia 38–89: Borderlands* depicting the expulsion of the Sudeten Germans. In the control group, participants play two games from the *Trader of Stories* series; these are narrative, point-and-click adventures from a fantasy world. *The Trader of Stories* series has game mechanics similar to Czechoslovakia 38–89: Borderlands but contains neither sensitive material nor historical references.

As outcome variables, we measure explicit and implicit attitudes towards the expulsion of the Sudeten Germans, both immediately after the intervention and 1 month later. As control

variables, we measure explicit and implicit attitudes before the intervention; together with participants' positive and negative affect and their characteristics such as gender, age and education.

3.2. Rationale for the Choice of the Game

Concerning the depicted historical topic in the modified version of *Czechoslovakia 38–89: Borderlands*, we have intentionally focused on the expulsion as it represents a sensitive topic challenged by the Czech historical narrative. As such, it allows us to predict attitude change.

After WWII, up to 3,000,000 Germans were forced to leave the country; they represented over a quarter of then Czechoslovakia's entire population (Abrams, 1995; Czech Statistical Office, 2018; Glassheim, 2000; Staněk, 2005). It is also estimated that 15,000–40,000 Germans died during the process (Abrams, 1995; Glassheim, 2000; Staněk, 2005). The act of the expulsion remained mostly unchallenged in prevailing historical narratives on the Czech lands during the last century.

Since the creation of Czechoslovakia in 1918, Czech Germans were considered outsiders or immigrants (outgroups) by many Czechs in the new republic (Glassheim, 2000; Rothschild, 1974). Before the Great Depression, 75% of the Czech German population voted for parties that were not organized around nationalist issues (Zahra, 2010). Their predominant allegiance to the nationalist Sudeten German party began after the Great Depression (ibid.). Until the fall of the communist regime in 1989, official Czech historiography argued mainly in moral or legal favour of the expulsion: considering it inevitable due to the Czech Sudeten German party's collaboration with the Nazi occupants of Czechoslovakia during the WWII (Abrams, 1995; Glassheim, 2000). After the Velvet Revolution in 1989, attempts to oppose this historical narrative and condemn the principle of collective guilt remained the exception in contemporary Czech political representation

and provoked many protests (Houžvička, 2005; Kunštát, 1998). The expulsion is still considered justified in general historical discourse. Likewise, the topic arouses passions to this day and plays a role in political struggles in the Czech Republic (CVVM, 2019; Kenety & Janzer, 2018; Kořan et al., 2010).

Therefore, we consider the expulsion to be a historically sensitive topic in the Czech socio-cultural environment: one that affects Czech intergroup attitudes towards Sudeten Germans. We expect our participants to have a rather positive attitude towards the expulsion due to the prevailing historical narrative.

Czechoslovakia 38–89: Borderlands depicts the expulsion from various perspectives; thus challenging the popular historical narrative about the justified expulsion. The game is based on real testimonies and historical research, but the game's characters and story are fictional.

3.3. Hypotheses

We formed the following four hypotheses and one exploratory goal:

H1: The experimental intervention will cause negative short-term explicit attitude change towards the expulsion.

Immediately after exposure to the game, we expect to find significant negative explicit attitude change (i.e., negative pre-post difference) towards the expulsion of the Sudeten Germans in the experimental group. Furthermore, we expect to find significantly more negative pre-post change in the experimental group than in the control group.

Rationale for H1: The highly plausible narrative in *Czechoslovakia 38-89: Borderlands* depicts the expulsion from multiple perspectives. Due to the perspective-taking, the game narrative's high plausibility and revelation of new, topic-related information in the game (which is

new in relation to the prevailing Czech historical narrative), we expect that the intervention will affect propositional reasoning leading to more negative explicit attitudes towards the expulsion.

H2: The explicit attitude change towards the expulsion in the experimental group will be long-term.

In the experimental group, we expect to find significant negative explicit attitude change towards the expulsion of the Sudeten Germans from the time of the pretest to the one-month delayed posttest (i.e., negative pre-del difference). Furthermore, we expect to find significantly more negative pre-del difference in the experimental group than in the control group.

Rationale for H2: Explicit attitudes are based on propositional reasoning about the particular object. This proposition must be in line with other propositions in one's memory and in line with what one considers to be the truth (Gawronski & Badenhausen, 2014). The explicit attitude changes we anticipate will mean a revision of one's beliefs toward the topic (Gawronski & Brannon, 2019), which is a process with a long-term effect. Therefore, expected immediate change in explicit attitudes towards the expulsion will last, i.e. become long-term (see Figure 2).

H3: The experimental intervention will cause a negative short-term implicit attitude change towards the expulsion.

Immediately after exposure to the game, we expect to find significant negative implicit attitude change (i.e., negative pre-post difference) towards the expulsion of the Sudeten Germans in the experimental group. Furthermore, we expect to find significantly more negative pre-post change in the experimental group than in the control group.

Rationale for H3: The intervention contains historical perspectives on the expulsion that are not present in the prevailing historical narrative in the Czech Republic. Because of that, we

assume that the intervention will temporary activate previously existing, but rarely used, associative links between the event of expulsion and other, general concepts stored in participants' memories which they consider to be negative (e.g., collective guilt or attitudes towards property confiscation). These new associative links will influence participants' associative evaluation of the event as a whole. As a consequence, we expect a negative short-term change in implicit evaluations towards the expulsion.

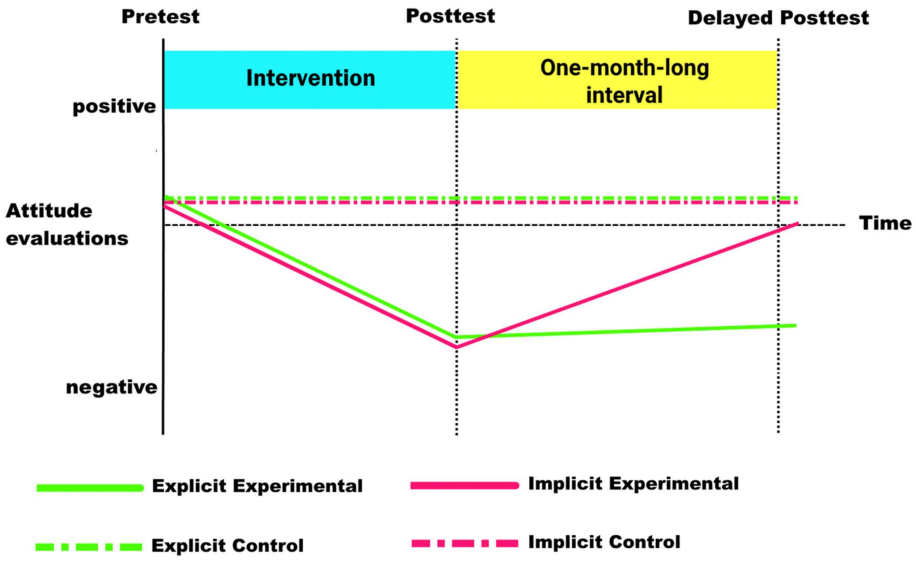

Figure 2. Expected change in implicit and explicit attitude evaluations over time

H4: Implicit attitudes towards the expulsion will not change over the long term.

Rationale for H4: In order to observe long-term negative implicit attitude change, associative evaluations of the expulsion should be affected. This can happen due to the frequent co-occurrence of the expulsion with other concepts participants consider to be negative (e.g. confiscation of property, collective guilt) activating the associative links between them. We do not expect this to happen because the length and character of our intervention will activate the associative links between the expulsion and the other concepts participants consider to be negative

only a few times. As a result, we do not expect our intervention to affect implicit attitudes towards the expulsion over the long term; only in the short term.

Exploratory Goal: To explore which areas of the studied construct are most affected by the intervention, we will analyse between-group differences in pre-post and pre-del change in explicit attitude measurements towards the expulsion on the item level.

Rationale for the exploratory goal: As mentioned before, the topic of the expulsion is a complex one. Therefore, we measured the explicit attitudes through two scales focused on different levels of detail regarding the topic: on the most general level (a Macro level composed of general evaluation of the expulsion as a whole) and on the level of specific topic-related statements (a Micro level composed of particular topics such as confiscation of property). The two scales were composed of multiple items, each focusing on a different perspective of evaluating the expulsion on the respective level. We formed two constructs from these items, but it may be the case that some of the items are affected more than the others by the intervention. We expected no between-group differential item functioning (DIF) on the baseline, and our goal was to explore whether item level differences between groups are present in pre-post and pre-del changes. While item level measurements are less reliable than the scale as a whole, we do not expect to confirm between-group differences in change on the item level for all items. However, any confirmed significant differences may point to areas of the studied construct that are most affected by our intervention.

4. Methods

4.1. Participants

Inclusion criteria for participants consisted of being a Czech-speaking person and completion of Czech elementary/basic school. As such, we can assume that they were exposed to

the prevailing narrative about the expulsion and the aftermath of WWII in the Czech context. Therefore, Czech participants have a unique set of knowledge and attitudes towards the topic as it is close to them. This fact is important as the results from several studies (e.g., Kampf, 2016a) suggest that there is a relation between the observed effect on attitudes and participants' cultural background. Furthermore, we required that all our participants have at least basic knowledge of the English language since the game used in the control condition was in English only.

The optimal sample size for the study was estimated using a power analysis to be 64 persons per group so as to detect medium between-group-difference effect size (Cohen's $d = 0.5$) of attitude change using a two-sample t test with a significance level $\alpha = 0.05$ and power $1-\beta = 0.80$. We collected data from 148 participants: of these, 141 were recruited through portals offering short-term jobs in Prague. The advertisement stated that we offer participation in a study dealing with video games and their influence on game players. The participants were rewarded financially for their time (500 CZK: approx. 18 EUR). The remaining seven were college students who were offered extra credit for their coursework.

We further excluded two participants who were not native Czech speakers and one participant who had already played an earlier version of our game. The remaining 145 participants' ages ranged from 15 to 30 years ($M = 20.9$; $SD = 3.9$; Women: 42.1%) and they had various levels of completed education (elementary: 25.5%; high school: 54.5%; university: 20%). We assigned participants randomly to the experimental ($n = 81$) and control ($n = 64$) groups. We have set an equal number of data collection sessions for both conditions, but the imbalance was caused by the lower ratio between the participants who confirmed their attendance and the participants who actually arrived in the control condition, even though that the hiring process was identical.

One hundred twenty-four participants arrived for the second testing session that occurred 1 month later. From those 124, we further excluded five participants due to external influences[2] resulting in 119 participants taking part in the second testing session (experimental group: n = 73, control group: n = 46).We do not know what caused the higher drop-out rate in one of the conditions as the data collection process was identical.

4.2. Interventions – Video Games

4.2.1. Experimental Intervention.

Our intervention tool in the experimental group was a modified version of the game *Czechoslovakia 38-89: Borderlands* (http://cs3889.com). The original game is a single-player narrative video game from the *Czechoslovakia 38-89* series. It provides its players with a story based on historical research and the personal testimonies of eyewitnesses to specific historical events from the years 1945-1948, i.e. the expulsion of the Sudeten Germans from then Czechoslovakia, the resettlement of the Czechoslovak borderlands, and the communists' rise to power. It is a dialogue-based, full motion video adventure. The game has two different time dimensions: the present and the past. In the present, players take on the role of a clerk deciding on the preservation of a school building in a village in the Czech borderlands. The players' goal is to learn as much as possible from the game characters, eyewitnesses to events from the years 1945-1948, so as to decide the school building's fate. As players interact with the eyewitnesses in the present (Figure 3), these eyewitnesses provide players with fragments of the past through video

[2] Those five participants were strongly affected by Angela Merkel's speech on 21 June 2018 delivered on World Refugee Day and dealing with the topic of the expulsion of Sudeten Germans. They were also impacted by subsequent public debate in Czech media during the period between their posttest and the one-month delayed posttest. See Kenety and Janzer (2018).

sequences, comics and thematic mini-games (Figure 4 and Figu re 5). In the past, players cannot change the history depicted in the game, they can only interact with it.

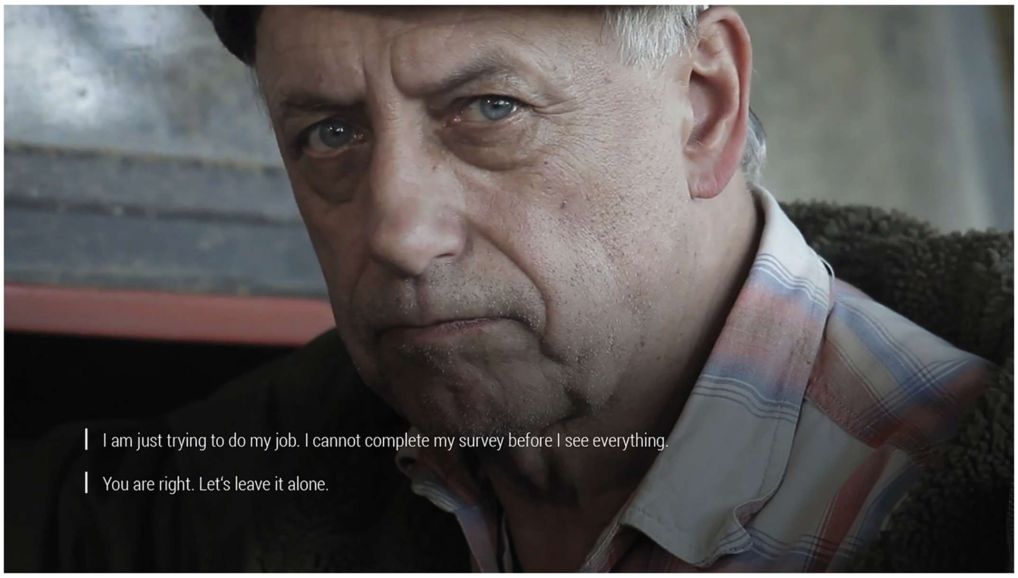

Figure 3. An example of interaction with an eyewitness from the present in *Czechoslovakia 38–89: Borderlands*

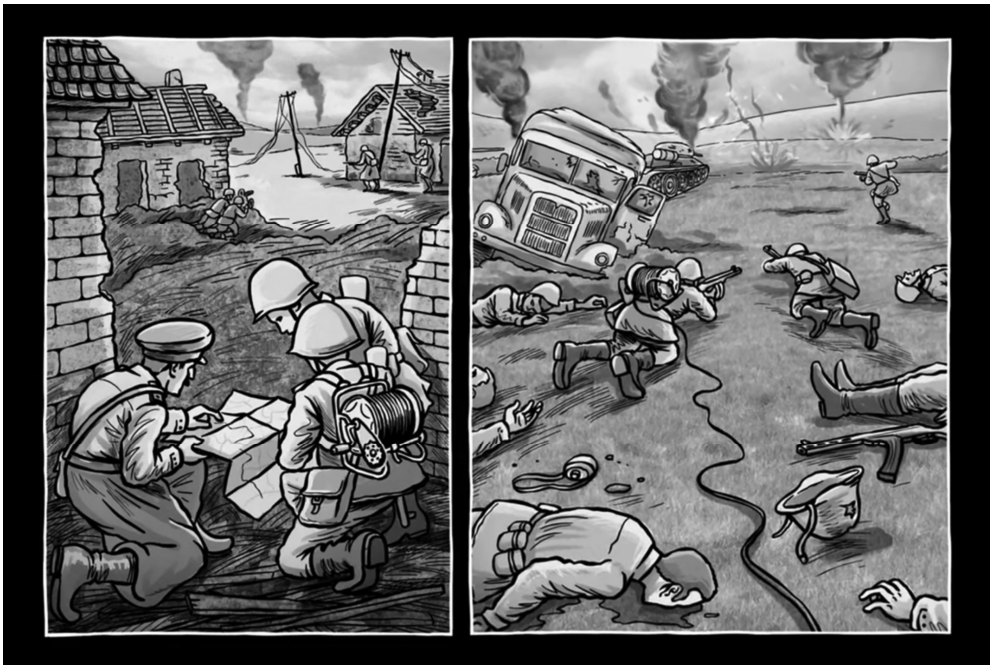

Figure 4. A demonstration of the interactive comics from *Czechoslovakia 38–89: Borderlands*

Figure 5. A demonstration of a point-and-click mini-game focused on exploring attics from *Czechoslovakia 38–89: Borderlands*

The game uses perspective-taking in its game design. Thus, it allows players to approach historical events from various perspectives to uncover different characters' points of view and motivations. In this study's game version, they interact with seven characters: 1) a Volhynian Czech who was allocated a farm in the Czech borderlands after WWII, 2) an entrepreneur who is the son of a former, high-level communist functionary, 3 and 4) siblings of Sudeten German origin who were separated as one was expelled and the other escaped the expulsion, 5) a local amateur historian living in the school building, 6) a former member of the National Security Corps and an anti-Nazi resistance fighter and 7) a Holocaust survivor of Jewish origin.

We developed a modified version of the game that depicts primarily the act of the expulsion of the Sudeten Germans. The game was divided into two parts. The first part was approx. 28 minutes long and contained the introduction to the whole story (provided by game character 1), a point-and-click minigame exploring the school attic with its historical items Figure 5), and stories

about the expulsion of the Sudeten Germans offered by game characters affected mostly negatively by the event (characters 3, 4 and 7).

The second part was approx. 22 minutes long and it focused on the stories of those who took an active part in the expulsion of the Sudeten Germans (game characters 2 and 6) or those who perceived the act as inevitable (game character 5). Game characters also provided players with interactive comics as part of their narration of historical events (see Figure 4).

We modified the original game in order to ensure that participants always encountered all the key information about the act of expulsion with the game still providing them different walkthroughs to finish it. Thus, we created a standardized gaming experience for all players without apparent limitations on players' agency.

4.2.2. Control Intervention.

The intervention tool used in our control group included two games from the *Trader of Stories* series (Trader of Stories, 2017): *A Grain of Truth* and *Trader of Stories Chapter 1*. The series was chosen to provide participants in our control group with relatively the same game experience with respect to the game mechanics of *Czechoslovakia 38-89: Borderlands*, but with a different narrative as these games are unrelated to the expulsion of Sudeten Germans. Both games from the *Trader of Stories* series are strongly narrative, point-and-click adventures. The gameplay in both games consists of collecting items, discovering new locations, people and creatures and collecting their stories through a dialogue system and mini games. The first game used in the control group is about a woman named Myosotis who collects stories from others to regain her memory. The second game released in the series is a prequel to the first game: it focuses on the moment when the main character lost her memory.

Gameplay in the control group was defined by our experimental group's average play times during the pilot studies: 28 minutes for the first part and 22 minutes for the second part. To ensure a standardized course for the experiment, participants played the first game from the series in the first part. They played the second game in the second part.

4.3. Measures

All our empirical data were collected through pen-and-paper-type questionnaires except for the computer-administered Single Category Implicit Association Test (SC-IAT).[3] All the questionnaires in the experiment had a high degree of reliability ($\alpha \geq 0.8$; see Table 1).

Table 1.

Internal Consistency of Measurement Scales

Measurement	Phase	number of items	Cronbach's alpha		
			Alpha	LCI	UCI
Macro explicit	Pretest	7	.88	.84	.90
	Posttest		.88	.84	.91
	Delayed		.84	.80	.88
Micro explicit	Pretest	10	.82	.77	.86
	Posttest		.87	.83	.90
	Delayed		.85	.81	.89
PANASp	Pretest	10	.82	.78	.86
	Posttest		.85	.80	.88
	Delayed		.87	.83	.90
PANASn	Pretest	10	.80	.74	.84
	Posttest		.83	.78	.87
	Delayed		.81	.76	.85

Note. LCL/UCL - lower/upper confidence limit, i.e. lower/upper bound of 95% confidence interval

[3] We verified and adjusted all the measurement tools before the start of the experiment during the two pilot studies ($n = 18$; $n = 25$).

Background demographic questionnaire. The demographic questionnaire yielded data about the participants' gender, age and education.

PANAS. To evaluate participants' positive and negative affect, we used the Positive and Negative Affect Schedule (Watson, Clark & Tellegen, 1988). It contained two, 5-point Likert item mood scales. One scale was for positive affect; the other was for negative affect. Each scale contained 10 adjectives. Each of them coded on a scale of 1 to 5 with smaller values indicating less positive / more negative affect (total score: 10 – 50). Several studies indicated that attitudes and possible changes therein, may be affected by mood (e.g. Schwarz and Clore, 1983). We therefore used this scale as a control variable.

Macro and Micro explicit attitudes measurement. The chosen concept for this study (the expulsion of the Sudeten Germans) represents a complex topic. Similar to the study by Soekarjo and van Oostendorp (2015), we approach it using two explicit attitude constructs, evaluated through two questionnaires. The first one, Macro explicit attitude measurement, pinpoints attitude towards the expulsion as a whole by asking participants about their general opinion about the event. The second one, Micro explicit attitude measurement, focuses on participants' attitudes towards micro-level (specific) topics regarding the expulsion (e.g. confiscation of Sudeten Germans property or the necessity of the expulsion). In the **Macro** questionnaire, participants were asked to evaluate the event of the expulsion of the Sudeten Germans on a macro level using a five-point, semantic differential scale (Osgood, Suci & Tattenbaum, 1957) with seven pairs of bipolar adjectives. Each item consisted of five squares positioned between two bipolar adjectives. Each item was coded as -2 – +2. The pairs of words included the following: unnecessary-necessary, wrong-correct, inadequate-adequate, criminal-righteous, shameful-honest, unfair-fair, unfounded-justified (i.e., total possible score: –14 – +14). In the **Micro** questionnaire, participants assessed

ten evaluative statements about the specific topics related to the expulsion forming the construct. These statements were ranked on a seven-point Likert scale (Likert, 1932) coded 1 to 7; with smaller values meaning less positive / more negative attitude towards expulsion, yielding a total score on a scale from 10 to 70.[4] We specifically asked to what extent participants agreed with the following statements, e.g. "The Sudeten Germans were undeservingly accused of crimes committed by Nazi Germany." (see Table B.1 in the Appendix for a complete list).

Single Category Implicit Association Test (SC-IAT). To measure implicit attitudes towards one historical concept (Expulsion of the Sudeten Germans), we used a modification of the Implicit Association Test (Greenwald & Banaji, 1995; Greenwald, McGhee & Schwartz, 1998) from Karpinski and Steinman (2006) known as Single Category Implicit Association Test measuring participant reaction times. We used a computer-based version of the test. Of all the options (Wigboldus, Holland & van Knippenberg, 2005), SC-IAT appears to be the most reliable method for implicit attitude measurement towards a single concept (Karpinski & Steinman, 2006).

The SC-IAT reveals participants' positive or negative associative evaluations towards the measured phenomenon. The test measures participants' response times when they are categorizing words into three categories: negative evaluations, positive evaluations and the category containing concepts related to the measured phenomenon; in our case the Expulsion of Sudeten Germans category.

Apart from the practice section, there are always two categories on one side and one category on the other side of the screen. Which category is on which side depends on the test phase as explained below. We used eight words for each of the three categories (Positive adjectives,

[4] Negatively formulated items were reverse-coded for the purpose of the analysis.

Negative adjectives and Expulsion of Sudeten Germans; see Table A.1 in the Appendix) and respected the methodology proposed by Karpinski and Steinman (2006). The routine consisted of instructions (see Figure 6) followed by three activity blocks, in which participants categorized words into two categories (see Figure 7 and Table 2). The three main activity blocks were separated by instruction screens summarizing the instructions for the subsequent block.

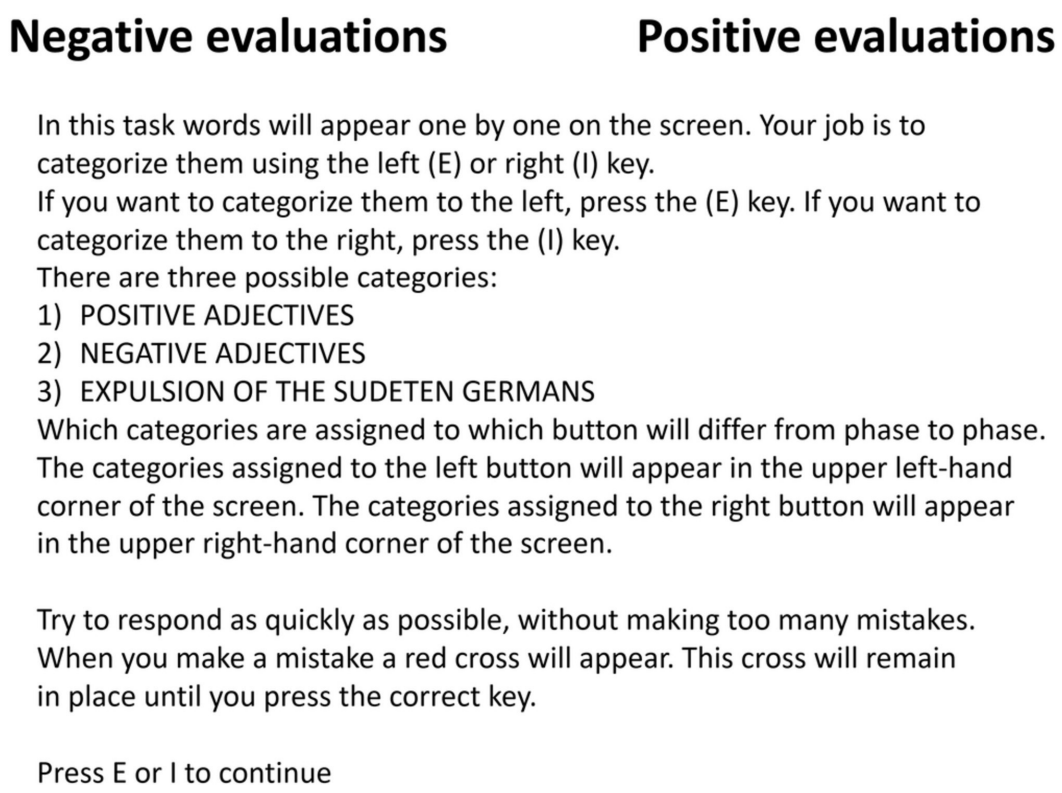

Figure 6. SC-IAT initial instructions

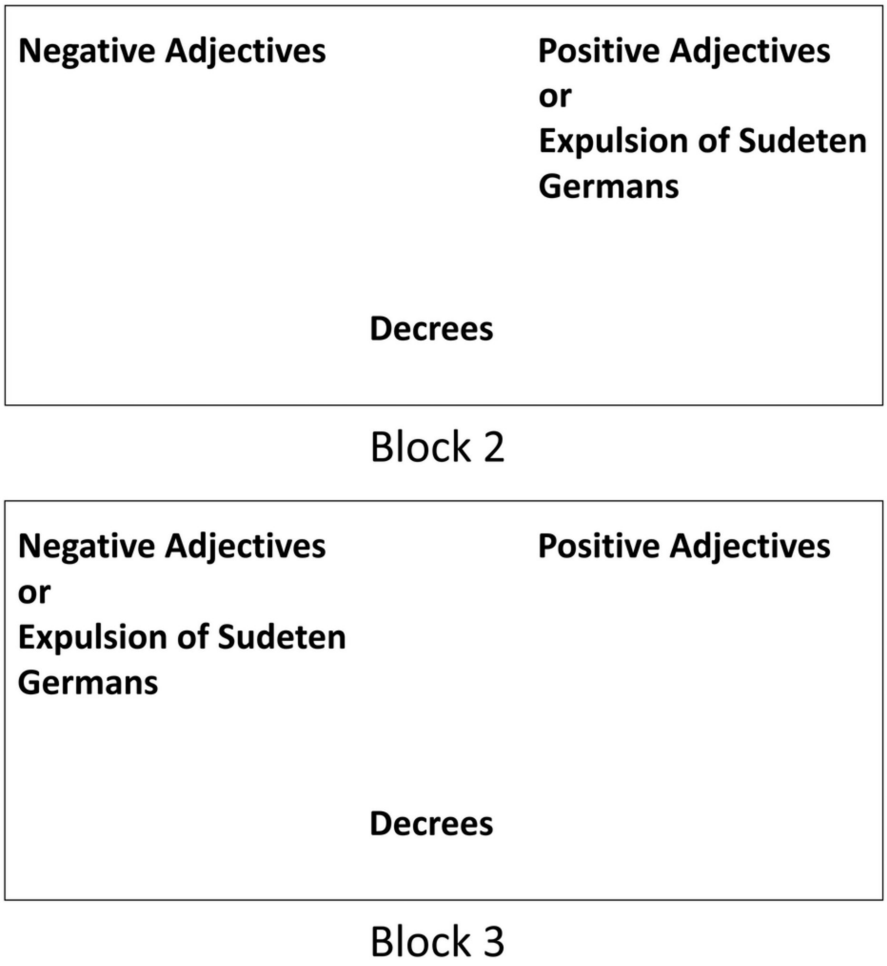

Figure 7. Demonstration of the categorization of the word decrees in the block 2 and the block 3 of the SC-IAT

In the first practice block, participants had to categorize eight words from the Negative adjectives category and eight words from the Positive adjectives category twice (2x2x8 = 32 words in total). This was the only block not used in the evaluation. In the second block, participants categorized eight words from each of the three categories four times (3x4x8 = 96 words in total). The third block also consisted of eight words from each of the three categories, which also appeared four times (3x4x8 = 96 words in total). The first 24 words in the second and the third blocks were considered practice – they were included in the calculation of the final score, but they were

calculated separately from the remaining 72 words in each block. The test was evaluated using the improved scoring algorithm for D-score as proposed by Greenwald, Nosek & Banaji (2003).

Table 2.

Blocks in the Single Category Implicit Association Test

Block	Words	Function	Left-key response categories	Right-key response categories
1	32	Practice	Negative adjectives	Positive adjectives
2a	24	Practice	Negative adjectives	Positive adj. + Expulsion
2b	72	Test	Negative adjectives	Positive adj. + Expulsion
3a	24	Practice	Negative adj. + Expulsion	Positive adjectives
3b	72	Test	Negative adj. + Expulsion	Positive adjectives

The test's assumption was that relative response times would reveal participants' positive or negative associative evaluations of the expulsion of the Sudeten Germans. Participants' objective is to categorize words appearing in the middle of the screen into the correct categories on the sides of the screen as fast as possible (see Figure 7 and Table 2). Participants implicit attitudes are revealed by the comparison of response times in categorization of the words into the respective categories between the Blocks 2 and 3. For instance, as illustrated in Figure 7, assume that participants are significantly faster in categorizing the word *Decrees* to the category Expulsion of Sudeten Germans on the right side of the screen when paired with the category *positive adjectives* than to the left side of the screen when paired with *negative adjectives*. It would reveal that for participants it is easier to categorize the word when associated with positive adjectives than when associated with negative adjectives, i.e., in terms of the APE model, participants link the concept of the expulsion with positive evaluations and the link between the positive evaluations is

stronger with the expulsion than the link with the negative evaluations, which is revealing positive implicit attitudes towards the concept.

We made two modifications to the original methodology proposed by Karpinski and Steinman (2006). All the odd-numbered participants in our study followed the routine described in the previous paragraph. With an aim of minimizing the effect of task order, the order of our blocks two and three was exchanged for each even participant. Another modification involved our first block where participants categorized only positive and negative adjectives (as a means of familiarizing themselves with the mechanism). Our computer-based measurement tool was created using Psychopy2 Experiment Builder (Peirce, 2007; 2009) and by modifying Hussey's (2016) existing script.

4.4. Procedure

The study included two testing sessions; always conducted between 9:00 AM and 12:15 PM (see Figure 8). Participants were tested in groups of three to 12 persons, each sitting at a separate computer in a lab. We began sessions with standardized instructions about the course of the experiment and by signing informed consent forms about participation in the study.

The pretest phase of the first testing session consisted of the following steps (see Figure 8). Firstly, we collected basic demographic data and data on participants' initial positive and negative affect. Next, we evaluated participants' initial implicit attitudes towards the expulsion of the Sudeten Germans using SC-IAT. Next, we collected data on Macro and Micro explicit attitudes.

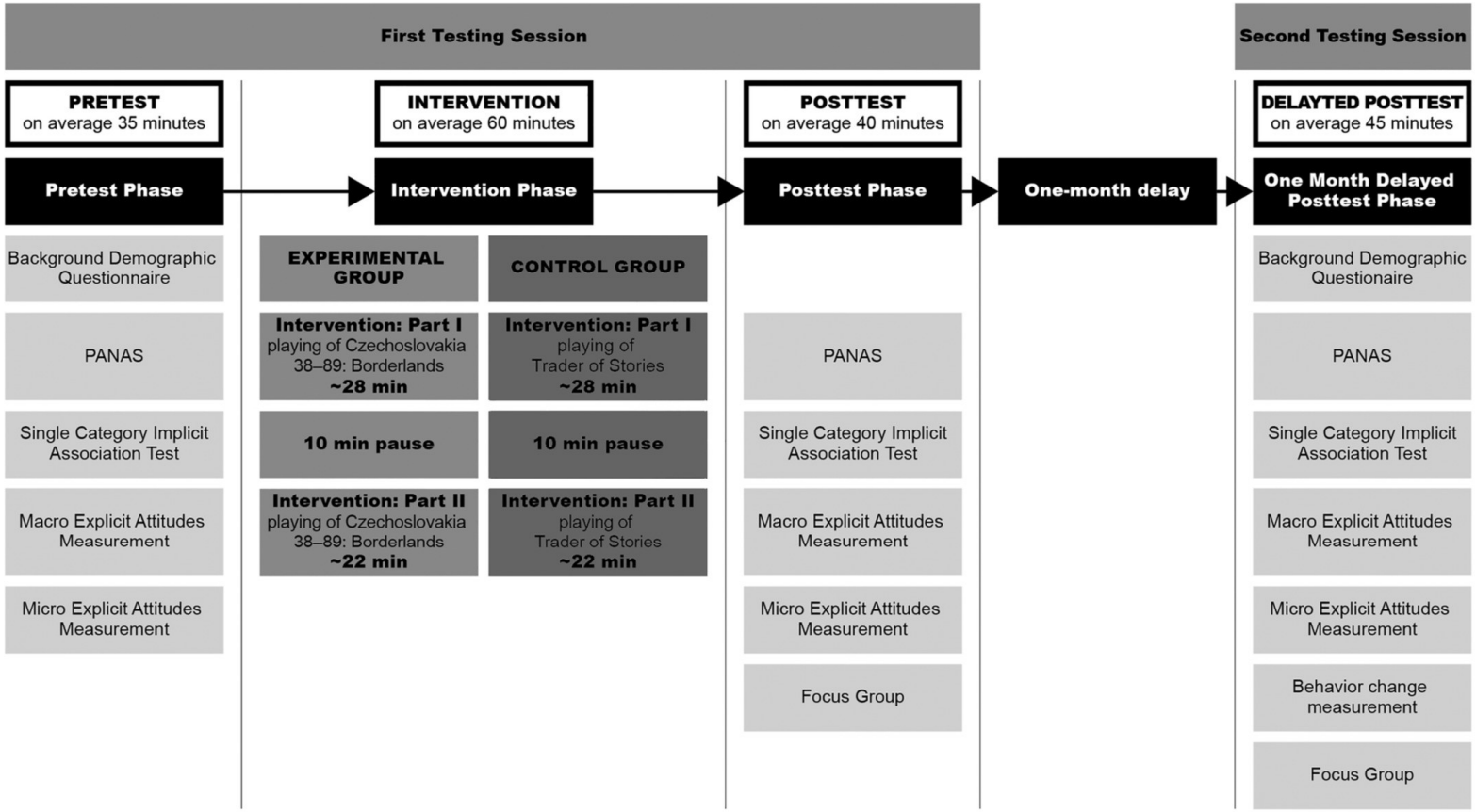

Figure 8. Experiment Procedure

After the pretest phase of the data collection, participants were exposed to the approx. 28-minute-long intervention of playing *Czechoslovakia 38-89: Borderlands* (experimental group) or the game *A Grain of Truth* from the *Trader of Stories* series (control group). After a ten-minute break, they played the second part of the *Czechoslovakia 38-89: Borderlands* game or *Trader of Stories Chapter 1* for approx. 22 minutes. Faster participants were instructed to wait until the last participant finished the game.

The posttest phase of the first testing session followed the same pattern as the pretest phase; apart from the absence of the questionnaire collecting demographic data. Afterwards, we organized focus groups with participants (irrelevant for present purposes).

We organized our second testing session three to five weeks after the intervention. It followed the same pattern as the pretest and afterwards included one extra questionnaire irrelevant for the present purpose (this was a behaviour change measurement with eight questions). Then, we collected qualitative data through focus groups: this time we focused on participants' activities related to the topic of the expulsion (this was done in order to consider whether or not to exclude participants from the delayed testing session). Apart from the possible effect of Ms. Merkel's speech about the topic on five participants during the monitored period, this data did not reveal other important factors and it was not analysed further. In the end, we debriefed our participants.

4.5. Data Analysis

The reliability of questionnaires used was assessed in terms of internal consistency by means of Cronbach's alpha. We used an X^2 test for binary variables (gender) and two-sample t tests otherwise to compare the experimental and control groups on the baseline.

To assess changes in total scores with respect to the group, we used t tests as an initial analysis, and more complex regression models to provide more precise description including the

characteristics of respondent and time as covariates. We used paired *t* tests to assess changes in explicit and implicit attitudes towards expulsion between the pretest and posttest, as well as changes between the pretest and delayed posttest. We used two-sample *t* tests to compare the experimental and control groups in their changes. Further, to account for respondent characteristics and initial PANAS (which may have influenced changes in respondents' implicit and explicit attitudes), we fit linear mixed-effect regression models clustered by ID in which post-intervention explicit/implicit attitude score (Micro, Macro, SC-IAT) was modelled by the following: its pre-intervention counterpart, time (posttest vs. delayed posttest), PANAS (positive and negative), education, age, gender and group membership (experimental/control), their interactions with pretest score, and interaction between group and time. We have then excluded insignificant effects one by one, until all remaining effects or their interactions were significant. The assumptions of normality, linearity, homoscedasticity, and independence of residuals were evaluated for all models. To support our hypotheses, we checked for significance of group effect and its interactions in the final models.

Finally, to provide more detailed information on between-group differences on the item level, we checked for differential item functioning (DIF, Holland & Wainer, 1993, Martinková et al., 2017) and we analysed changes in individual items in the macro and micro explicit attitudes using analysis of differential item functioning in change (DIF-C, Martinková et al., 2020) based on a cumulative logit ordinal regression model (see Appendix C for a technical description of the DIF-C analysis). Significant DIF-C signalized that two respondents with the same explicit attitude pretest score but from different groups (experimental vs. control) had different probabilities for assigning certain ratings to a given item in the posttest or delayed posttest measurement.

In all statistical tests and models, the assumptions were checked, and we used the Benjamini-Hochberg correction to account for multiple comparisons. Analyses were performed

using R version 4.1.0 (R core team, 2021) and its expansion packages including ShinyItemAnalysis 1.3.7 (Martinková & Drabinová, 2018) and difNLR version 1.3.7 (Hladká & Martinková, 2020).

5. Results

There were no significant between-group differences in any of the respondent characteristics and the initial PANAS, even though there was a trend towards difference in age (Table 3). In pretests, all attitude measurements were close to the midpoints in both groups (i.e.: Macro explicit: 0; Micro explicit: 40; SC-IAT: 0). For Pearson correlations between explicit and implicit attitude measurements, see Table E.1 in the Appendix.

5.1. Hypothesis 1: Short-term explicit attitude change

As expected, we found significant between-group difference in explicit attitude pre-post change towards the expulsion of the Sudeten Germans: both in the macro ($d = -0.34$, $p = .022$) and micro measurement ($d = -0.53$, $p = .001$), thus supporting Hypothesis 1. In both groups, the posttest macro attitude measurements showed significant negative shift in attitudes towards the expulsion compared to the pretest, while the shift in micro attitude measurements was significant only in the experimental group (see Table 4). Significant post-intervention group differences were supported by regression models accounting for both respondent characteristics and pretest scores (see Tables 5 and 6). In micro measurement, the post-intervention score was on average two points lower for respondent from the experimental group than for respondent with the same pretest score and characteristics but from the control group. In macro measurement, the difference depended significantly on pretest score; for example, for a male respondent from experimental group with pretest score 1SD above average, the post-intervention score was 1.62 lower than for their male peers from control group with the same pretest score (see Figure 9).

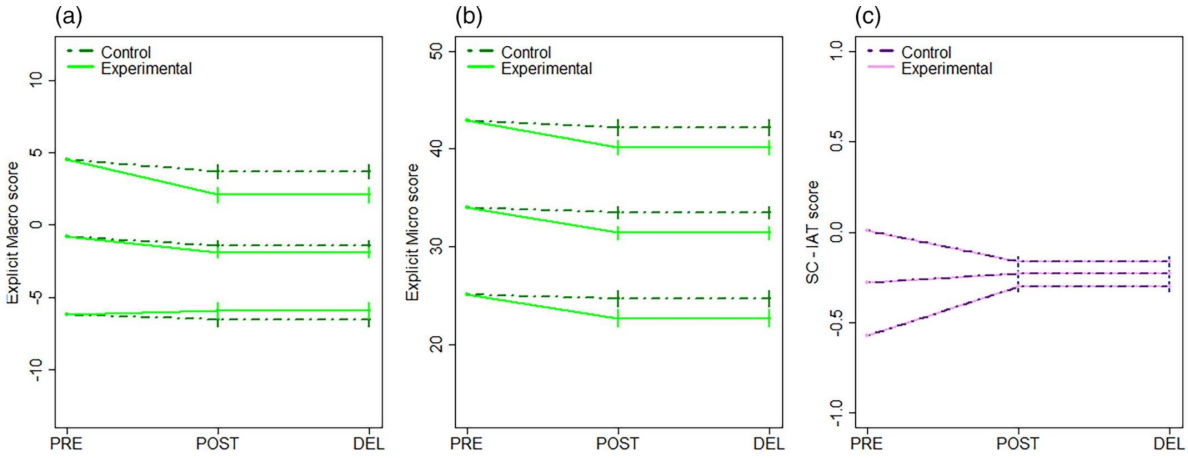

Figure 9. Predicted values and confidence intervals for final mixed effect models for (a) macro score, (b) micro score, and (c) SC-IAT score. Displayed for male with pretest values equal to mean value, 1SD above, and 1SD below mean

5.2. Hypothesis 2: Long-term explicit attitude change

Between-group differences remained significant in the micro explicit attitude measurements when considering the shift from the pretest to the 1 month delayed posttest (d = −0.44, p = .014; see Table 7), however, these between-group differences were not significant for macro explicit attitude measurements (d = −0.16; p = .266; see Tables 7 and 8). Therefore, we can partially confirm Hypothesis 2: with respect to micro rather than macro attitudes. The effect of time (i.e, the difference between posttest and delayed posttest) was found negligible by the regression models both for macro and micro measurements, see Figure 9.

5.3. Hypothesis 3 Short-term Implicit attitude change

Despite the fact that implicit evaluations differed significantly in pre-post comparisons between the experimental and the control groups (d = −0.32, p = .029), there was no significant pre-post change in the experimental group itself (d = 0.08, p = .764) and the between-group difference was rather due to changes in the control group (see Table 4). In a regression model for

posttest/delayed-posttest implicit attitudes, no significant effect was found other than the pretest scores (see Table 9 and Figure 9). Thus, we cannot conclude that Hypothesis 3 was supported.

5.4. Hypothesis 4: Long-term Implicit attitude change

There was neither a significant pre-post, nor pre-del change in implicit attitude measurements in the experimental group. We also did not observe a significant pre-del change in the control group (Table 7). These results are consistent with Hypothesis 4.

5.5. Exploratory goal: Item level differences

Looking in more detail at the individual items in the macro and micro explicit attitude measures, we found, as expected, no significant between-group differences in item functioning before the intervention, that is, no significant DIF in the pretest.

After the intervention, we observed significant between-group differences in item functioning in change, that is, significant DIF-C, for some items. In the micro explicit attitude measure, we found significant DIF-C only in Item 10, that is, “The borderland eviction was carried out fairly.” Respondents from the control group were more likely to give higher scores than their peers from the experimental group who had the same score on the micro pretest. For example, a respondent from the experimental group with 35 points on the pretest would have a 43% probability of assigning three or more points (measured on a scale of 1–7) to this item on a posttest. Meanwhile, a respondent from the control group with the same pretest score would have a 66% probability of assigning three or more points (see dot-dashed lines in Figure 10a).

Table 3.

Descriptive Statistics for the experimental and the control group in pretest

	Experimental group			Control group			Experimental vs. Control group				
	n	Male	Female	*n*	Male	Female	w	Chi2 stat			*p*
Gender (female ratio)	81	41	40	64	43	21	0.15	3.38			.204
	n	*M*	*SD*	*n*	*M*	*SD*	M_{diff}	*SD*	*d*	*t* stat	*p*
Age	81	20.17	3.26	64	21.77	3.66	-1.59	3.52	-0.45	-2.73	.058
Education (1 ES, 2 HS, 3 U)	81	1.86	0.68	64	2.05	0.65	-0.18	0.67	-0.27	-1.64	.204
PANAS+	80	29.54	5.76	63	27.96	6.36	1.58	6.06	0.26	1.53	.204
PANAS-	80	15.41	5.16	63	14.70	4.07	0.71	4.71	0.15	0.92	.405
Macro explicit	81	-0.42	5.02	64	-1.32	5.75	0.90	5.36	0.17	0.98	.405
Micro explicit	81	35.15	8.28	64	32.61	9.56	2.54	8.93	0.28	1.68	.204
SC-IAT	81	-0.26	0.26	63	-0.30	0.32	0.04	0.29	0.14	0.83	.405

Table 4.

Differences between the experimental and the control group in posttest-pretest

	Experimental group									Control group									Experimental vs. Control group				
		PRE		POST		PRE-POST difference					PRE		POST		PRE-POST difference								
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*p*	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*p*	M_{diff}	*SD*	*t* stat	*d*	*p*
Macro	81	-0.42	5.02	-2.10	4.77	-1.67	3.62	-0.46	<.001	64	-1.32	5.75	-1.95	5.91	-0.63	2.00	-0.32	.021	-1.04	3.05	-2.20	-0.34	.022
Micro	81	35.15	8.28	31.48	9.02	-3.67	5.98	-0.61	<.001	64	32.61	9.56	31.78	10.05	-0.83	4.04	-0.21	.105	-2.84	5.38	-3.40	-0.53	.001
SC-IAT	81	-0.26	0.26	-0.24	0.28	0.03	0.33	0.08	.764	63	-0.30	0.32	-0.17	0.27	0.13	0.34	0.40	.008	-0.11	0.33	-1.92	-0.32	.029

Table 5.

Parameter estimates in final linear mixed effects model for posttest and delayed posttest macro explicit measurement

	Estimate	*SE*	*t*	*p*
Baseline (Group C, Gender M, Pretest Macro 0)	-0.65	0.36	-1.82	.072
Pretest Macro explicit	0.96	0.06	15.16	<.001
Group Experimental	-0.67	0.44	-1.50	.136
Gender Female	-1.25	0.45	-2.81	.006
Pretest Macro : Group Experimental	-0.21	0.08	-2.56	.011
Pretest Macro : Gender Female	-0.23	0.08	-2.78	.006
sd(ID)	1.94			

Note: Final model: ScoreMacro ~ PreMacro + Group + Gender + PreMacro:Group + PreMacro:Gender, random=~1|ID.

Table 6.

Parameter estimates in final linear mixed effects model for posttest and delayed posttest micro explicit measurement

	Estimate	*SE*	*t*	*p*
Baseline (Group C, Gender M, Pretest Micro 0, Posttest)	0.21	2.08	0.10	.918
Pretest Micro explicit	0.98	0.06	16.99	<.001
Group Experimental	-2.04	0.78	-2.61	.010
Gender Female	8.57	3.03	2.83	.005
Pretest Micro : Gender Female	-0.30	0.09	-3.49	.001
sd(ID)	3.82			

Note: Final model: ScoreMicro ~ PreMicro + Group + Gender + PreMicro:Gender, random=~1|ID

Table 7.

Differences between the experimental and the control group in delayed posttest-pretest

	Experimental group									Control group									Experimental vs. Control group				
		PRE		DEL		PRE-DEL difference					PRE		DEL		PRE-DEL difference								
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*p*	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*p*	M_{diff}	*SD*	*d*	*t* stat	*p*
Macro	73	-0.20	5.09	-2.07	4.30	-1.87	4.00	-0.47	<.001	46	-0.84	5.69	-2.09	5.50	-1.25	3.50	-0.36	.059	-0.62	3.82	-0.16	-0.89	.266
Micro	73	35.56	8.15	32.15	8.38	-3.41	6.42	-0.53	<.001	46	33.40	9.29	32.51	9.05	-0.89	3.98	-0.22	.207	-2.52	5.72	-0.44	-2.65	.014
SC-IAT	72	-0.27	0.25	-0.27	0.30	-0.01	0.35	-0.02	.881	43	-0.31	0.26	-0.25	0.32	0.07	0.39	0.19	.213	-0.08	0.37	-0.22	-1.12	.266

Table 8.

Differences between the experimental and the control group in delayed posttest-posttest

	Experimental group									Control group									Experimental vs. Control group				
		POST		DEL		POST-DEL difference					POST		DEL		POST-DEL difference								
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*p*	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*d*	*p*	M_{diff}	*SD*	*d*	*t* stat	*p*
Macro	73	-1.92	4.75	-2.07	4.30	-0.15	2.77	-0.05	.877	46	-1.61	5.93	-2.09	5.50	-0.48	3.60	-0.13	.388	0.33	3.11	0.11	0.53	.900
Micro	73	31.68	9.15	32.15	8.38	0.47	4.56	0.10	.877	46	31.96	9.43	32.51	9.05	0.55	4.31	0.13	.388	-0.08	4.45	-0.02	-0.10	.919
SC-IAT	72	-0.23	0.27	-0.27	0.30	-0.04	0.29	-0.14	.877	43	-0.16	0.27	-0.25	0.32	-0.08	0.40	-0.20	.388	0.04	0.34	0.12	0.58	.900

Table 9.

Parameter estimates in final linear mixed effects model for posttest and delayed posttest implicit measurement (SC-IAT)

	Estimate	*SE*	*t*	*p*
Baseline (PreIATeffect 0, Posttest)	-0.16	0.03	-5.87	<.001
PreIATeffect	0.24	0.07	3.49	.001
sd(ID)	0.15			

Note: Final model: IATeffectScore ~ PreIATeffect , random=~1|ID

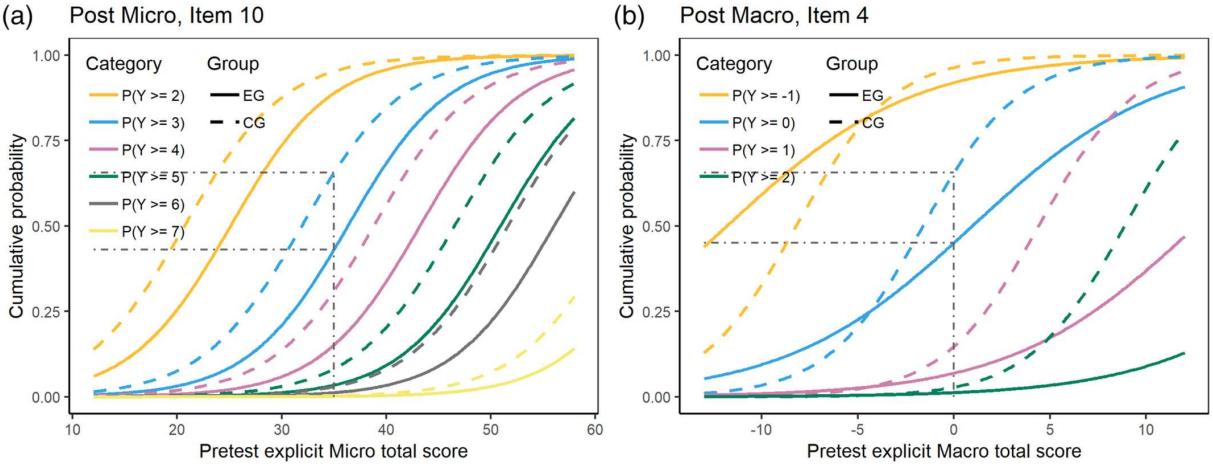

Figure 10. (a) DIF-C for pre-post change in micro attitudes—Item 10: Cumulative probabilities are higher for control groups uniformly for all levels of pretest scores (b) non-uniform crossing DIF-C for pre-post change in macro attitudes—Item 4: Cumulative probabilities are higher for control group for the upper part of levels (of about −5 and higher) of the pretest micro total score and are lower for the lower pretest micro total score levels

With respect to pre-del change, the between-group differences in item functioning (DIF-C) in Item 10 remain significant. Moreover, we see significant DIF-C in Item 6, that is, "Unjustified violence was committed against the Sudeten Germans during the post-war arrangements" (see Figure 10).

Analogously, in the macro explicit attitude measure in the posttest, we found significant DIF-C only in Item 4, that is, "criminal–righteous". For example, when observing two respondents with a macro pre-score of 0, a respondent from the experimental group had a 45% probability of assigning 0 or more points to this item on a post-test, while a respondent from the control group had a 66% probability of assigning 0 or more points to this item (see Figure 10b).

These results suggest that items related to fairness and justness were affected by the intervention more than the others. While these results are statistically significant even after

correction for multiple comparisons, caution is needed here: this analysis was exploratory without a priori formulated hypotheses and, as such, may need further confirmation by follow-up studies.

6. Discussion

This study investigated a video game's effect on implicit and explicit attitude change of Czech young adults towards a depicted historical event over the short and long term. The video game we used as the intervention was Czechoslovakia 38–89: Borderlands, which deals with the expulsion of the Sudeten Germans from Czechoslovakia after WWII using perspective-taking as its core game mechanic.

6.1. Main Findings

First, the results support H1: they confirm a short-term negative explicit attitude change in the experimental group and a greater attitude change in comparison to the control group: both in macro and micro attitude measurements. In other words, playing a serious historical game presenting the personal accounts of expelled persons, as well as those actively engaged in the expulsion, immediately negatively affects attitudes towards both the expulsion as a whole as well as particular aspects thereof: deportations, property confiscation, and violence. The exploratory DIF-C analysis provided a deeper understanding of explicit attitude changes on the item level, suggesting that the intervention affected especially items related to the (un)fairness of the expulsion. We assume that this particular effect on items related to (un)fairness may be caused by the depiction of the event in our intervention and how it challenges the popular historical narrative about the justified expulsion.

Second, the results partly support H2: they confirm a long-term negative explicit attitude change compared to the control group: but in micro attitude measurements only. In other words, in the long term, being exposed to personal accounts of the expulsion in a video game seems to affect negatively the evaluation of its particular aspects, yet not the expulsion as a whole. Also, exploratory DIF-C analysis confirmed between-group differences in item functioning with respect to pre-post and pre-del change in items of explicit measures related to the fairness and justness of the expulsion. These results may provide a basis for a more nuanced explanation of the intervention effect in future studies.

Third, implicit attitude measurements did not significantly change in the experimental group from the pretest to the posttest (H3) and from the pretest to the delayed posttest (H4). In other words, being exposed to the serious historical video game about the expulsion changed the deliberate logical conclusion about the expulsion (i.e., explicit attitudes), yet not the spontaneous affective response thereto (i.e., implicit attitudes).

6.2. Interpreting the results

Although there is no study of a similar scale, there exist several studies that provide partial data about video games and attitude change: studies to which we can relate our findings.

6.2.1. Short-term explicit attitude change

Most of the empirical studies that deal with the subject (e.g., Cuhadar & Kampf, 2014, 2015; Pentz et al., 2019) measured attitudes primarily through self-report scales and without a control group. They revealed short-term changes in participants' explicit attitudes. However, in studies without a control group, we cannot exclude the effects of external factors; especially in the context of an ongoing debate about the nature of attitudes and their context-dependency. Therefore, we will relate our findings only to studies with control groups (see Table 10).

Table 10.

List of studies with a pretest and a control group to which we relate our findings

Study	Explicit Measurements	Implicit Measurements	Short-Term Data	Long-Term Data	Topic	Persuasive Mechanic
Kampf, Exp 1, (2015)	Yes	No	Yes	No	Israeli-Palestinian conflict	Perspective-taking
Kampf, Exp 2, (2015)	Yes	No	Yes	No	Israeli-Palestinian conflict	Perspective-taking
Kampf, (2016a)	Yes	No	Yes	No	Israeli-Palestinian conflict	Perspective-taking
Kampf, (2016b)	Yes	No	Yes	Yes	Israeli-Palestinian conflict	Perspective-taking
Peña et al., 2018	Yes	No	Yes	No	Helping Immigrants	Perspective-taking
Price et al., 2015	Yes	No	Yes	No	Knowledge of body processes	Various mechanics with meaningful choices
Ruggiero, 2015	Yes	No	Yes	Yes	Homelessness	Pre-scripted economical model
Alblas et al., 2018	No	Yes	Yes	No	Healthy nutrition	Evaluative conditioning

Regarding short-term explicit attitude change, there exist several studies measuring it while using a pretest-posttest design with a control group. No studies focused on attitudes towards historical topics. To our knowledge, five of these studies (Kampf, 2015, Exp. 1, Exp. 2; 2016a, 2016b; Peña et al., 2018) used games utilizing perspective-taking in their game design. On the one hand, all Kampf's studies (using the games *Peacemaker* and *Global Conflicts*) confirm the effect of this game mechanic on attitude change towards the sides in the Israeli-Palestinian conflict in comparison to the control group. On the other hand, the study by Peña et al., (2018) did not find this effect as it did not reveal any significant between-group attitude change. The game they used, *Papers, Please*, provides to some extent perspective-taking in its persuasive content focusing on measuring attitudes towards helping immigrants. In that game, players take on the role of a border guard controlling who can enter the country. Players decide whether immigrants can enter the country or not. This decision possibly results in financial consequences for the family of the player's character. As such, significantly more space is given to the emotional impact of a player's decisions on the family of the player's character than to the emotional dimensions of the random encounters with immigrants towards which are attitudes measured. Thus, the game's perspective-taking towards the measured phenomenon (immigrants) is limited. We assume that this is the main reason why this study did not report any attitude changes towards the measured topic.

Another group of studies did not use perspective-taking in their design (Price et al., 2015; Ruggiero, 2015). The first, one, a study by Price et al. (2015), confirms video games' ability to change attitudes. The study used the educational video game *Code, Fred* incorporating various game mechanics with meaningful choices in relation to the measured topic—the importance of knowledge about what is happening in one's body. The authors measured attitude change using two questions on a five-point Likert scale. Both questions resulted in small between-group

differences. However, the significant difference in the second question was caused only by changes in the control group. Attitudes towards the importance of knowledge about the processes inside one's body (as depicted in the game) were measured. However, these processes are all automatic in the human body in real life, thus creating a dissonance potentially limiting the game's persuasive potential.

The second one, the study by Ruggiero (2015), did not reveal any significant between-group explicit attitude changes. The study used game *SPENT* simulating life in poverty and measured attitudes towards homelessness. In the game, players react to various life events (e.g., buying insurance, food) while living from a limited budget given them. The game is based on a pre-scripted scenario, which is formed on the economic model interpreting the effects of the players' decisions on the outcomes in the game, that is, falling or not falling into poverty. Thus, the persuasive potential is pursued through this pre-scripted economical model. The reasons that players end up leading a life of poverty or becoming homeless in this game are reduced to bad financial decisions with no further context. We assume that is the reason why the study did not reveal any significant short-term between-group attitude changes.

Findings on the ability of video games to change attitudes are possibly diverse because games' persuasive mechanics differ. Previous studies indicate that perspective-taking, which was used in our intervention, has been most explored as a persuasive mechanic. It has also been shown to have significant effects on explicit attitude change in relation to depicted topics when the game provides adequate perspective. There is currently no other game mechanic with such clear results.

6.2.2. Long-term explicit attitude change

Regarding long-term explicit attitude change, there exist, to our knowledge, only two studies that measured it while having a control group (Kampf, 2016b; Ruggiero, 2015). Neither of

them focused on attitudes towards historical topics. As already said, the former examined attitudes towards homelessness using a pre-scripted economical model as a persuasive mechanic; the latter used perspective-taking as a persuasive mechanic concerning attitudes towards Israelis and Palestinians in their ongoing conflict. Similar to our study, Kampf's study confirmed the persistence of explicit attitude change over the long term. However, his study collected data only from a homogeneous sample: students studying political science. We extended these findings to a more general, young adult audience. Ruggiero's study confirmed between-group difference over the long term due to a shift in the control group between the posttest and the delayed posttest. There was no clear explanation for this effect, and more data is needed for its interpretation.

In our study, we observed smaller between-group explicit attitude change in the long term compared to Kampf's study (Kampf, 2016b) which used game design principles similar to those in our intervention. Ruggiero's study also confirmed between-group differences in the long term, but only due to a shift in the control group. At the same time, our game intervention was shorter than the one in Kampf's study[5] and longer than that in Ruggiero (2015). Plus, our study confirms significant differences in comparison to the control groups in micro attitude measurements in pretest-delayed posttests, but not in macro attitude measurements. Considering our results, (Kampf, 2016b), and Ruggiero's unclear results (Ruggiero, 2015), they so far indicate that the length of the intervention could play a key role in long-term explicit attitude change; but this idea needs further examination.

We assume the length of the intervention itself is not the only factor influencing a game's persuasive potential. Attention should also be paid to public discourse about the particular topic

[5] The length of the video game intervention was not specified by Kampf (2016b), but the whole experiment took 3 h: including two short questionnaires, a short demo introduction to the game and the video game intervention. Therefore, we assume that the video game intervention was longer than 50 min.

depicted in the game; especially when dealing with attitudes towards sensitive topics as in our study. Our results revealed long-term between-group attitude change in micro explicit attitudes and not macro explicit attitudes. Our suggested explanation for this is as follows. The historical narrative in the Czech lands during the last century spoke mostly in favour of the expulsion and framed it as inevitable. This affected the socio-cultural perception of the event; mostly on the general level as the topic was not discussed in comprehensive detail by the public. Also, the concept of the expulsion on the general (macro) level is, to some extent, abstract. Agreeing or disagreeing with something on the abstract level is less likely to cause dissonance in one's propositions about the topic. Attitudes on the specific (micro) level focused on concepts which are more precise/detailed and are relatable for one's own evaluations, that is, participants at some point of their lives had already formed attitudes towards those relatable concepts (the confiscation of property or the principle of collective guilt). The game itself used personal, relatable stories about particular issues thus humanizing those relatable concepts through depictions of their impact on real people. For that reason, we assume the effect of our intervention on the macro level was weaker in the short term and not significant over the long term. Meanwhile, it was significant on the micro level. The results of our explorative DIF-C analysis suggest that our game's effect could have possibly been larger if measured as relates to the fairness or the justness of the expulsion rather than towards the expulsion as a whole. Therefore, the context of the depicted topic in the game seems also to be an important factor concerning the game's persuasive potential.

6.2.3. Implicit attitude change

Regarding implicit attitude change, there exists, as far as we know, only one study that measured it while using a pretest-posttest design with a control group (Alblas et al., 2018). The study measured short-term implicit attitudes towards healthy nutrition using evaluative

conditioning as a persuasive mechanic within the intervention. In agreement with our study, it does not show significant implicit attitude change in the experimental group. Surprisingly, we found implicit attitude change (pre-post) in the control group, that is, similar to Alblas et al. (2018). We assume that our game did not affect implicit attitudes as it was only 50 min long. It was longer than Alblas et al.'s intervention which lasted just 10 minutes, but it was not long enough to create new, strong, associative evaluations or to strengthen existing ones. Concerning the unexpected findings in the control group compared to the experimental group, it may be the case that the effect of measurement shifted both groups positively towards the neutral middle point. At the same time, the effect of our intervention in the experimental group caused a more negative shifting of implicit attitudes (similar to explicit attitudes). As a result, implicit attitudes in the experimental group did not change significantly, since the effect of the intervention was cancelled out by the effect of the measurement. Meanwhile, implicit attitudes in the control group changed significantly as they were affected only by the effect of measurement. The implicit attitudes in both groups in the delayed posttest were not significantly different to their pretest values, thus indirectly supporting our assumption about the effect of measurement. Significant difference between the two groups in the pre-post difference could then be considered proof that our video game has the ability to affect implicit attitudes. However, more data is needed to evaluate the effect of measurement on implicit attitudes. So far, the empirical evidence concerning the effects of video games on implicit attitudes is inconclusive.

6.3. Significance of results and their implications

On the empirical level, our study (as far as we know), is the first one to provide evidence that historical video games are capable of changing both short- and long-term explicit attitudes

towards the historical events they depict. It should be noted though that, over the long term, we confirmed this change in the micro attitude measurement only.

Our study also has a methodological implication. Exploratory DIF-C analysis suggests that some aspects of the expulsion mentioned in our explicit attitude change questionnaires were affected more than others. For future studies, we suggest that attitudes should be considered as a multidimensional construct; especially, attitudes towards sensitive topics.

A review by Todd and Galinsky (2014) suggested that perspective-taking results in a more favourable attitude towards the group whose perspective was taken. However, this review did not include any game-based learning study. The empirical implication of our study and studies by Kampf is that the findings of Todd and Galinsky (2014) on the effect of perspective-taking can be extended to the domain of games. For example, studies by Kampf (2015, Exp. 1, Exp. 2) using games depicting a recent conflict from multiple perspectives caused more favourable attitudes among participants towards certain sides in the Israeli-Palestinian conflict. The study by (Kampf, 2016b) also confirmed this effect of perspective-taking given that more favourable attitudes were measured towards the opposite side of the conflict than the side representing participants' origins (e.g., Palestinian towards Israelis). On the level of explicit evaluations, our study supports the effect of perspective-taking on attitude change and its leading to more favourable attitudes towards the perspective depicted in the game.

6.4. Limitations

Several of the study's limitations are worth mentioning. First, our experimental intervention lasted only 50 minutes. However, narrative historical video games are often played for dozens of hours. Second, our number of participants decreased between the first (experimental group n = 81; control group n = 64) and the second data collection (experimental group n = 73; control group n =

46): mostly in our control group. Third, our research sample did not allow for the analysis of other possible socio-cultural moderators/mediators, as our participants were a pre-defined group between 15–29 years and as the whole study was conducted in the country's capital city which might possibly be more liberal compared to the rest of the country. Fourth, we did not measure participants' knowledge of the depicted topic. The reason for this was that the questionnaire measuring knowledge could have created uncertainty about the topic and resulted in changes to participants' information seeking behaviour related to the topic, thus affecting our data about the attitudes (Henefer & Fulton, 2005).

7. Conclusion

Several scholars have suggested on a theoretical level that historical video games have the potential to determine how we think about, understand and negotiate the past (Chapman et al., 2017). Nevertheless, despite the importance of these claims, there has not been, until recently, enough empirical evidence to support them.

Our study offers evidence to support the notion that historical video games can change players' attitudes on the level of explicit attitude measurements towards the depicted topics. However, it did not demonstrate video games' effects on implicit attitudes. Changes in micro explicit attitudes (but not in macro attitudes) compared to the control group were stable and thus confirmed historical video games' potential to affect the formation of society's historical awareness. DIF-C analysis provided a more detailed view; suggesting a shift in items discussing the (un)fairness of the expulsion of the Sudeten Germans.

As video games have become a worldwide phenomenon affecting whole societies regardless of social status or age, this study offers new perspectives for our understanding and interpretation of the formation of historical narratives in informal settings. As far as we know, there

is currently no other similar study of this scale providing empirical data about video games' effects on explicit and implicit attitude change over the long term towards the depicted topics.

Future research should collect more long-term empirical data concerning the effect of games on players' attitudes. We currently know very little about the effects of various video game elements on player's attitudes and information behaviour and more studies should focus on the effects of particular game elements on players. Furthermore, very few studies collect complex data including players' socio-cultural background and playing styles, which would allow a more comprehensive analysis with regard to these characteristics; especially the playing styles may possibly have an effect on players' experience with games. The replication of our study with a different research sample and context could provide valuable additional results in this area.

Acknowledgment

LK, CB, and VŠ were primarily supported by the PRIMUS/HUM/03 project at Charles University. LK was also partly supported by student grant FF/VG/2017/115 at the Charles University's Faculty of Arts. VŠ was further supported by the European Regional Development Fund Project, “Creativity and Adaptability as Conditions for the Success of Europe in an Interrelated World” (No. CZ.02.1.01/0.0/0.0/16_019/0000734) and the Charles University Program Progress Q15. PM was partly supported by the PRIMUS/17/HUM/11 projects at Charles University and by the Czech Science Foundation. Grant Number: 21-03658S. The game Czechoslovakia 38-89: Borderlands was developed at Charles University and the Czech Academy of Sciences with support from the Czech Ministry of Culture. All authors had complete access to data supporting the manuscript. The human data were collected with APA ethical principles in mind. We would like to thank Petra Plintová for her help with organizing and conducting the

experiment, Michaela Vařejková and Adéla Hladká for computational assistance, and Ondřej Paška and David Vávra for their consultations on software development.

Competing interests

Czechoslovakia 38-89: Borderlands was developed by Charles University and the Czech Academy of Sciences. Vít Šisler and Lukáš Kolek were members of the game's development team. Charles University plans on receiving revenues from a significantly enhanced commercial version of the game that is to be released in 2021. None of these factors influenced how the study was conducted at any stage of its implementation. Other authors declare no conflict of interests.

Data availability statement

The human data were collected with APA ethical principles in mind. Participant data was anonymized by assigning numbers to each participant. The differential item functioning analysis presented here is implemented within an open-source package for R ShinyItemAnalysis (Martinková & Drabinová, 2018) with an interactive Shiny web application available from https://shiny.cs.cas.cz/ShinyItemAnalysis/. The data that support the findings of this study beyond those presented and the selected code to replicate the analyses can be obtained from Figshare at https://doi.org/10.6084/m9.figshare.14690934 (Kolek, Šisler, Martinková & Brom, 2021).

References

Abrams, B. F. (1995). Morality, Wisdom and Revision: The Czech Opposition of the 1970s and the Expulsion of the Sudeten Germans. *East European Politics and Societies, 9*(2), 234-255. https://doi.org/10.1177/0888325495009002003

Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology, 104*(1), 235–249. https://doi.org/10.1037/A0025595

Alblas, E., Folkvord, F., Anschütz, D. & 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. doi: 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

Atkins, B. (2003). *More than a game: The computer game as fictional form*. Manchester University Press.

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

Bohner, G., & Dickel, N. (2011). Attitudes and Attitude Change. *Annual Review of Psychology, 62*, 391-417. https://doi.org/10.1146/annurev.psych.121208.131609

Brannon, L. A., Tagler, M. J., & Eagly, A. H. (2007). The moderating role of attitude strength in selective exposure to information. *Journal of Experimental Social Psychology*, *43*(4), 611-617. https://doi.org/10.1016/j.jesp.2006.05.001

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*. London: Routledge.

Chapman, A., Foka A., & Westin J. (2017). Introduction: What is historical game studies?. *Rethinking History, 21*(3): 358–71. https://doi.org/10.1080/13642529.2016.1256638

Cuhadar, C.E., & Kampf, R. (2014). Learning about Conflict and Negotiations through Computer Simulations: The Case of Peacemaker. *International Studies Perspectives 15*(4), 509-524. https://doi.org/10.1111/insp.12076

Cuhadar, E., & Kampf, R. (2015). Does Conflict Content Affect Learning from Simulations? A Cross-National Inquiry into the Israeli-Palestinian and Guatemalan Conflict Scenarios. *Negotiation and Conflict Management Research*, *8*(4), 243-260 doi: https://doi.org/10.1111/ncmr.12062

CVVM SOÚ AV ČR, (2019). *Tisková zpráva: Občané o odsunu a Benešových dekretech – říjen 2019*. Centrum pro výzkum veřejného mínění: Sociologický ústav AV ČR, v. v. i.., edited by Spurný, M.

Dal Cin, S., Zanna, M. P., & Fong, G. T. (2004). Narrative Persuasion and Overcoming Resistance. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion* (pp. 175-191). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers. https://doi.org/10.4324/9781410609816

Duncan, L. R., Hieftje, K. D., Pendergrass, T. M., Sawyer, B. G., & Fiellin, L. E. (2018). Preliminary investigation of a videogame prototype for cigarette and marijuana prevention in adolescents. *Substance Abuse*, *39*(3), 275-279. https://doi.org/10.1080/08897077.2018.1437862

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition*, *25*(5), 582-602. https://doi.org/10.1521/soco.2007.25.5.582

Fazio, R. H. (2007). Attitudes as Object-Evaluation Associations of Varying Strength. *Social cognition*, *25*(5), 603–637. https://doi.org/10.1521/soco.2007.25.5.603

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.

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. http://dx.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. http://dx.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

Gawronski, B., & Brannon, S. M. (2019). Attitudes and the implicit-explicit dualism. In D. Albarracín & B. T. Johnson (Eds.), *The handbook of attitudes. Volume 1: Basic principles* (2nd edition, pp. 158-196). New York, NY: Routledge.

Gawronski, B., & Payne, B. K. (Eds.). (2010). *Handbook of implicit social cognition: Measurement, theory, and applications.* New York, NY: Guilford Press.

Gerling, K.M., Mandryk, R. G., Birk M. V., Miller, M., & Orji, R. (2014). The effects of embodied persuasive games on player attitudes toward people using wheelchairs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, New York, NY, USA, 3413-3422. https://doi.org/10.1145/2556288.2556962

Glassheim, E. (2000). National Mythologies and Ethnic Cleansing: The Expulsion of Czechoslovak Germans in 1945. *Central European History 33*(4), 463-486. https://doi.org/10.1163/156916100746428

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701-721. http://dx.doi.org/10.1037/0022-3514.79.5.701

Green, M. C., & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange, & T. C. Brock (eds.). *Narrative impact: Social and cognitive foundations* (pp. 315-341). Mahwah. NJ: Erlbaum.

Green, M. C., & Jenkins, K. M. (2014), Interactive Narratives: Processes and Outcomes in User-Directed Stories. *Journal of Communication, 64*(3), 479-500. https://doi.org/10.1111/jcom.12093

Green, M., Bilandzic, H., Fitzgerald, K., & Paravati, E. (2019). Narrative Effects. In Oliver, M. B., Raney, A. A., & Bryant, J. (Eds.), *Media effects* (pp. 130-146). New York, NY: Routledge.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27. http://dx.doi.org/10.1037/0033-295X.102.1.4

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464-1480. http://dx.doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Hawkins, R., Mendes Ferreira, G. A. R., & Williams, J. (2019). The development and evaluation of 'Farm Animal Welfare', an educational computer game for children. *Animals*, *9*(3). https://doi.org/10.3390/ani9030091

Henefer, J., Fulton, C. (2005). Krikelas's Model of Information Seeking. In K., E., Fisher, S., Erdelez, & L. E.,.McKechnie, (Eds). *Theories of information behavior* (pp. 225-229). Published for the American Society for Information Science and Technology by Information Today.

Hladká A., & Martinková P. (2020). difNLR: Generalized Logistic Regression Models for DIF and DDF Detection. R Journal, 12(1), 300-323. https://doi.org/10.32614/RJ-2020-014

Holland, P. W., & Wainer, H. (Eds.). (1993). *Differential item functioning.* Lawrence Erlbaum Associates, Inc.

Houžvička, V. (2005). *Návraty sudetské otázky*. Praha: Karolinum.

Hussey, I. (2016). Single Category Implicit Association Test (SC-IAT) programmed in Psychopy. GitHub, accessed 28 April 2018, https://github.com/ianhussey/SingleCategoryImplicitAssociationTest

Jacobs, R. S. (2018). Play to win over: Effects of persuasive games. *Psychology of Popular Media Culture, 7*(3), 231-240. http://dx.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. https://doi.org/10.1080/1369118X.2014.982142

Kampf, R. (2016a). 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. DOI: 10.13189/ujer.2016.040807.

Kampf, R. (2016b). 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

Kapell, M. W., & Elliott A.B.R. (2013). *Playing with the Past: Digital Games and the Simulation of History*. London: Bloomsbury Academic.

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91*(1), 16-32. http://dx.doi.org/10.1037/0022-3514.91.1.16

Knol, E., & de Vries, P. W. (2011). EnerCities, a Serious Game to Stimulate Sustainability and Energy Conservation: Preliminary Results. *E-learning papers*, *25*, 1-10.

Kořan, M. (2010). Česká zahraniční politika v roce 2009: Analýza ÚMV, Praha: Ústav mezinárodních vztahů.

Kunštát, M. (1998). Czech-German Relations after the Fall of the Iron Curtain. *Sociologický časopis, 6*(2), 149-172.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22 140,* 55.

Maio, G. R., & Haddock, G. (2010). *The Psychology of Attitudes and Attitude Change*. SAGE Publications Ltd. http://dx.doi.org/10.4135/9781446214299

Martinková, P., Drabinová, A., Liaw, Y. L., Sanders, E. A., McFarland, J. L., & Price, R. M. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE—Life Sciences Education*, 16(2), rm2.

Martinková, P., & Drabinová, A. (2018). ShinyItemAnalysis for Teaching Psychometrics and to Enforce Routine Analysis of Educational Tests. *The R Journal*, 10(2), 503-515. http://dx.doi.org/10.32614/RJ-2018-074

Martinková, P., Hladká, A., & Potužníková, E. (2020). Is academic tracking related to gains in learning competence? Using propensity score matching and differential item change functioning analysis for better understanding of tracking implications. *Learning and Instruction, 66*(April). https://doi.org/10.1016/j.learninstruc.2019.101286

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Peirce, J. W. (2007). Psychopy-Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8-13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform 2*(10). https://doi.org/10.3389/neuro.11.010.2008

Pentz, M. A., Hieftje, K. D., Pendergrass, T. M., Brito, S. A., Liu, M., Arora, T., et al. (2019). A videogame intervention for tobacco product use prevention in adolescents. *Addictive Behaviors*, *91*, 188-192. https://doi.org/10.1016/j.addbeh.2018.11.016

Peña, 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" [Online]. *Cyberpsychology, Behavior, and Social Networking*, *21*(11), 687-693. https://doi.org/10.1089/cyber.2018.0030

Petty, R. E., Briňol P., & DeMarree, K. G. (2007). The Meta-cognitive model (MCM) of attitudes: implications for attitude measurement, change, and strength. *Social cognition, 25*(5), 657-686. https://doi.org/10.1521/soco.2007.25.5.657

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.

Price, C., Gean, K., Christensen, C., Beheshti, E., Pernot, B., Segovia, G., . . . Ward, P. (2015). Casual Games and Casual Learning About Human Biological Systems. *Journal of Science Education and Technology, 25*(1), 111-126. https://doi.org/10.1007/s10956-015-9580-6

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rothschild, J. (1974). *East Central Europe Between the Two World Wars.* University of Washington Press.

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., & Anderson, C. A. (2013). Arabs as terrorists: Effects of stereotypes within violent contexts on attitudes, perceptions, and affect. *Psychology of Violence, 3*(1), 84-99. http://dx.doi.org/10.1037/a0030038

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.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*(3), 513-523. http://dx.doi.org/10.1037/0022-3514.45.3.513

Schwarz, N. (2007). Attitude construction: Evaluation in context. *Social Cognition, 25*(5), 638-656. http://dx.doi.org/10.1521/soco.2007.25.5.638

de Smale, S., Kors, M. J. L., & Sandovar, A. M. (2019). The Case of This War of Mine: A Production Studies Perspective on Moral Game Design. *Games and Culture*, *14*(4), 387–409. https://doi.org/10.1177/1555412017725996

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

Smith, C., Ryder, P., Blais, M., & Schneck, R. (2017). Evaluation of two different poverty simulations with professional phase pharmacy students. *Currents in Pharmacy Teaching And Learning*, *9*(5), 903-910. https://doi.org/10.1016/j.cptl.2017.05.018

Smith, M. B., Bruner, J. S., & White, R. W. (1956). *Opinions and Personality*. 1st eds., New York: John Wiley.

Soekarjo, M., & 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

Staněk, T. (2005). Poválečné "excesy" v českých zemích v roce 1945 a jejich vyšetřování. Prague: Ústav pro soudobé dějiny (Institute of Contemporary History).

Steinemann, S. T., Iten, G. H., Opwis, K., Forde, S. F., Frasseck, L., & Mekler, E. D. (2017). Interactive narratives affecting social change: A closer look at the relationship between interactivity and prosocial behavior. *Journal of Media Psychology: Theories, Methods, and Applications, 29*(1), 54-66. http://dx.doi.org/10.1027/1864-1105/a000211

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.

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. doi:10.1111/spc3.12116

Tragazikis, P., & Meimaris, M. (2009). Engaging Kids with the Concept of Sustainability Using a Commercial Video Game – A Case Study. In: Pan Z., Cheok A.D., Müller W., & Chang M. (eds) Transactions on Edutainment III. *Lecture Notes in Computer Science, 5940*, Berlin: Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-11245-4_1

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

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

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063-1070. http://dx.doi.org/10.1037/0022-3514.54.6.1063

Wigboldus, D.H.J., Holland, R.W., & Van Knippenberg, A. (2005). Single target implicit associations. Unpublished manuscript.

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

Zahra, T. (2010). Imagined Noncommunities: National Indifference as a Category of Analysis. *Slavic Review, 69*(1), 93-119.

Dataset

Kolek, L., Šisler, V., Martinková, P., & Brom, C. (2021). Supplementary Material for article. "Can Video Games Change Attitudes towards History? Results from a Laboratory Experiment Measuring Short- and Long-term Effects". https://doi.org/10.6084/m9.figshare.14690934

Online sources

Czech Statistical Office (2018, June 5). Obyvatelstvo – roční časové řady. Retrieved on 25 May from https://www.czso.cz/csu/czso/obyvatelstvo_hu

Kenety, B., & Janzer, T. (2018, June 21). Merkel Calls Sudeten German Expulsion "Immoral", Drawing Czech Ire. Retrieved 2019, May 25, from https://www.radio.cz/en/section/curraffrs/merkel-calls-sudeten-german-expulsion-immoral-drawing-czech-ire

Paintbucket Games (2019). Press kit of Through the Darkest of Times. Retrieved 2019, December 4, from https://press.paintbucket.de/

Trader of Stories (2017, January 1). Let the Adventure Begin. Retrieved 2019, May 25, from http://traderofstories.blogspot.com/

Valve Corporation (2019). Best of 2019 Top Sellers: The top 100 games as measured by gross revenue this year. Retrieved 2020, June 6, from https://store.steampowered.com/sale/2019_top_sellers

Yin-Poole, W. (2014, November 27). It took just two days for This War of Mine to make its money back. Retrieved 2019, March 21, from https://www.eurogamer.net/articles/2014-11-27-this-war-of-mine-made-its-money-back-in-just-two-day

Appendices

Appendix A. List of SC-IAT words

Table A.1

List of SC-IAT Words in Three Categories

Positive adjectives	Negative adjectives	Expulsion of Sudeten Germans
Just	Criminal	Germans
Good	Wrong	Borderlands
Right	Bad	Sudetenland
Fair	Evil	Decrees
Well-managed	Humiliating	Deportation
Merciful	Unfair	Displacement
Moral	Disgusting	German
Honest	Shameful	Expulsion

Appendix B. List of phrases used Micro explicit attitude measurement

Table B.1

Phrases Used in Micro Explicit Attitude Measurement

Item	Phrase
1	The expulsion of Sudeten Germans from the borderlands was the right decision.
2	The Czechs behaved harshly towards the Sudeten Germans after the war.
3	The displacement of Sudeten Germans was historically necessary.
4	The deportation of Sudeten Germans was a crime.
5	Sudeten Germans were displaced justifiably.
6	Unjustifiable violence was committed against the Sudeten Germans during the postwar arrangements.
7	The Sudeten Germans were undeservingly accused of crimes committed by Nazi Germany.
8	Czechoslovaks' confiscation of Sudeten German property after the war was fair.
9	Postwar settling of matters in the borderlands resulted in an unnecessary number of casualties of Sudeten Germans.
10	Borderland evictions were carried out fairly.

Appendix C. Technical description of the analysis of differential item functioning in changes in Micro and Macro explicit measurements

First, in order to check for differential item functioning (DIF) on baseline, we fit a cumulative logit ordinal regression model on each pretest item

$$P\left(Y_{pi} \geq k \middle| X_p, G_p\right) = \frac{\exp(\beta_{0ik}+\beta_{1i}+\beta_{0DI} \cdot G_p+\beta_{1DIFi} \cdot X_p \cdot G_p)}{1+\exp(\beta_{0ik}+\beta_{1i}+\beta_{0DIFi} \cdot G_p+\beta_{1DIFi} \cdot X_p \cdot G_p)} \qquad \text{(Eq. C.1)},$$

modelling the probability of person p from group G_p (0 Control, 1 Experimental) to assign k or more points Y_{pi} to item i, given their total score on the pretest was X_p. We used a likelihood-ratio test to compare the model (1) with a sub-model in which parameters β_{0DIFik} and β_{1DIFi} were set to zero. DIF was detected when group effect was significant; this DIF was referred to as non-uniform if the effect of interaction between the group and pretest total scores was also significant.

Subsequently, we tested for differential item functioning in change (DIF-C) using analogous cumulative logit ordinal regression models in which Y_{pi} were item scores assigned by person p on item i from posttest or delayed posttest, given their total score on pretest $X_{p.}$ Again, DIF-C was detected when group effect was significant, this DIF-C was referred to as non-uniform if the effect of interaction between the group and pretest total scores was also significant.

Appendix D. DIF-C Micro attitudes analysis

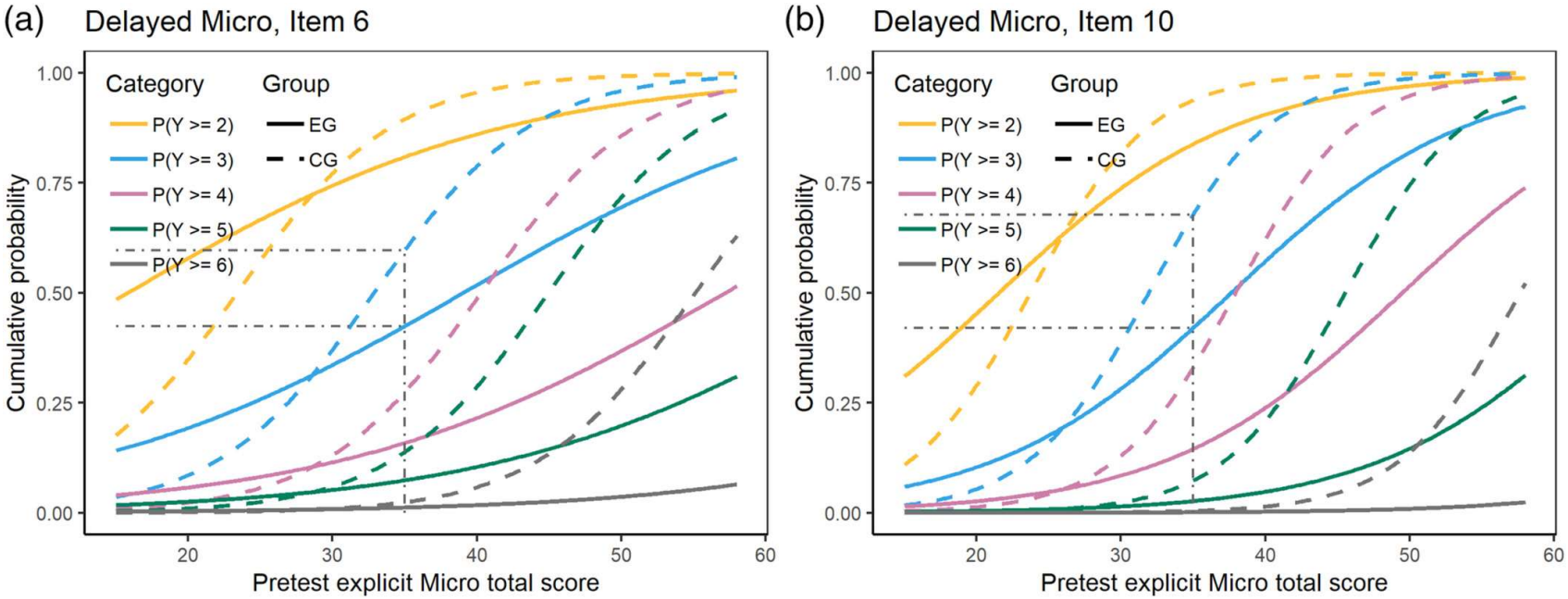

Figure D.1 DIF-C Micro attitudes – item 6 (A) and Micro attitudes – item 10 (B). Dot-dashed black lines depict between-group differences of probability assigning the same score in respondents with the same pretest score.

Appendix E. Pearson Correlations

Table E.1

Pearson Correlations between Explicit and Implicit Attitude Measurements

	Pre Macro	Post Macro	Del Macro	Pre Micro	Post Micro	Del Micro	Pre SC-IAT	Post SC-IAT	Del SC-IAT
Pre Macro	1	.84	.72	.81	.68	.59	.03	-.07	.06
Post Macro	.84	1	.81	.79	.83	.73	.03	0	.02
Del Macro	.72	.81	1	.79	.81	.83	-.09	-.07	-.05
Pre Micro	.81	.79	.79	1	.83	.78	.03	-.03	-.01
Post Micro	.68	.83	.81	.83	1	.88	.08	.05	.03
Del Micro	.59	.73	.83	.78	.88	1	-.02	-.06	-.05
Pre SC-IAT	.03	.03	-.09	.03	.08	-.02	1	.29	.17
Post SC-IAT	-.07	0	-.07	-.03	.05	-.06	.29	1	.32
Del SC-IAT	.06	.02	-.05	-.01	.03	-.05	.17	.32	1