JOURNAL OF NEW APPROACHES IN EDUCATIONAL RESEARCH
Vol. 7. No. 1. January 2018. pp. 61–69 ISSN: 2254-7339 DOI: 10.7821/naer.2018.1.260

ORIGINAL

Serious play in education for social justice - An exploratory study

Thomas Wartenweiler
Department of Educational Research, Lancaster University, United Kingdom {thomaswartenweiler@gmail.com}

Received on 05 September 2017; revised on 02 November 2017; accepted on 19 November 2017; published on 15 January 2018

DOI: 10.7821/naer.2018.1.260

ABSTRACT
Fair Battles is a 12-week Swiss education for social justice program with the goal of sensitizing high school students about the impact of their consumer habits on society. The pedagogical concept of Fair Battles is to employ the tool of serious play to enhance students' social empathy, which then leads to service learning projects. This exploratory mixed-methods study examined the impact of the program by using pre- and post-program student surveys (n=16) and post program semi-structured qualitative student interviews (n=10). The survey data were analyzed using SPSS and the interview data were analyzed using template analysis. The results were organized according to Kirkpatrick's four-level evaluation model. The quantitative results show a statistically significant increase in the post-program survey scores for the learning and behavior level. The qualitative results suggest that the program had a positive impact on students on all four of Kirkpatrick's levels. The conclusions are that the Fair Battels program is impactful, that social justice education needs to be holistic and that the combination of serious play and service learning elements seems to be effective for social justice education. Further research in the area of social justice education and serious play is recommended.

KEYWORDS: SOCIAL JUSTICE EDUCATION, IMPACT EVALUATION, SERIOUS PLAY, SOCIAL EMPATHY, SERVICE LEARNING

1 INTRODUCTION

This article describes the exploratory evaluation of a new Swiss education for social justice (ESJ) program called "Fair Battles". The goal of Fair Battles is to sensitize students about the impact of their consumer habits on other people, specifically on those producing consumer goods under poor working conditions. The program's intended outcome is greater student awareness about what it means to consume fairly. The pedagogical concept of Fair Battles is to employ the tool of serious play to enhance students' social empathy which then leads to service learning projects. This is supposed to both enhance the students' knowledge about global inequalities and their sense of agency in being able to make a difference. Zembylas' (2013) following two claims are therefore combined in the Fair Battles program:

> Critical compassion is… cultivated, if students begin to understand the conditions (structural inequalities, poverty, globalization etc.) that give rise to suffering and acknowledge some sort of human connection between themselves and others (p. 516).
>
> It is important for students to engage in action, what is meant is that they need to work actively to address structural inequality —which is the foundation of much suffering. (p. 517).

The main aim of the study was to explore the program's overall emotional, cognitive and behavioral impact on the students (n=16) of a small Swiss secondary school class and the general impact on their school as a whole. In addition, the study examined whether the pedagogical elements of serious play and service learning were seen as important or meaningful by key stakeholders. Despite its limitations in scope and generalizability, this explorative evaluation was significant because it was the first external evaluation of the program and because there exists little research on the impact of Swiss ESJ programs as well as on ESJ programs that incorporate serious play elements.

The next section of this article describes the local Swiss context in which the program is implemented. Then, the Fair Battles program and the core pedagogical elements of serious play, social empathy and service learning are introduced. The following section describes the methodology and evaluative approach taken in this study. A further section presents results in a descriptive format and a discussion of these results in light of the research literature. A final section offers reflections and recommendations for the Fair Battles program and for ESJ programs in general.

2 BACKGROUND

2.1 Local Swiss Context and Program Description

Through the introduction of the new national curriculum *Lehrplan 21* in Switzerland in 2015, social justice education has come more into the focus of mandatory schooling. The added curricular area, called "education for sustainable development" (*Bildung für Nachhaltige Entwicklung, BNE)*, explicitly introduces various aspects of education for social justice. The new BNE curriculum requirements have led to a considerable increase of social justice education programs. However, as Helbling and Schwarz (2015) point out, there is little control, coordination or evaluation of these new programs.

Fair Battles is a school program developed by the private Swiss association *Fairp(l)ay.* Its goal is to sensitize students about global economic inequalities and about the impact of their personal consumer choices. The word "battle" symbolizes competition or effort on various levels: in the global economy, between school

*To whom correspondence should be addressed:

Lancaster University, Bailrigg, Lancaster, LA1 4YW, United Kingdom

classes at the foosball table, and in regards to societal change that one needs to battle for (Verein Fairplay, 2017).

The program is divided into two interrelated phases. The first phase uses serious play elements to approach the issues in a light and playful, yet also reflective manner. Students research the background story of one chosen character. They then apply what they have learned unto a foosball table figure. If the chosen character is generally disadvantaged, this is reflected in the designing process of the figure, for example by narrowing its foot. If the character is generally advantaged, its foot will be designed more advantageous, for example by adding width to it. In this way, the entire foosball table is transformed into an "unfair" table with a disadvantaged and an advantaged team, reflecting global economic realities. To avoid stereotyping, this process is continuously reflected with the students during this phase. During an initial "unfair" tournament, the students experience global inequalities through the miniature world of the unfair table. The research of the background stories and the experience of an unfair game ideally lead to the development of social empathy amongst the students. This social empathy is translated into action in the second phase of the program through small-scale service learning projects, i.e. by volunteering in a fair-trade cooperation. The program's theory of change is that serious play leads to social empathy and social empathy in turn leads to active engagement through service learning (Figure 1):

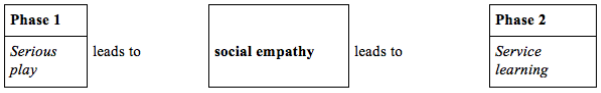

Figure 1. Fair Battles Theory of Change

The program website can be found at www.fairbattles.ch. It has also been previously featured in two positive newspaper articles in local media (Furrer, 2014; Tschopp, 2014).

2.2 Serious Play

The program's approach is to use elements of serious play to introduce the complex topic of global economic connections in a playful way that meets the students on an emotional level. Play has a longstanding tradition as a learning strategy in various forms. "Serious Play", which can be seen as a form of „Guided Play" (Weisberg, Hirsh-Pasek, Golinkoff, Kittredge, & Klahr, 2016), is therefore not a new concept, but it has received renewed attention lately in various areas of education (i.e. Schrage, 2013; Young & Murray, 2017). Serious play distinguishes itself from regular play by following purposefully designed activities to "create opportunities to exchange knowledge that can be instrumental in developing awareness of complex themes, issues, or perspectives." (Hinthorne & Schneider, 2012, p. 2808) Serious play activities always include a critical reflection element and are thus connected to real life events and circumstances while allowing at the same time a light approach to complex and at times heavy topics. "Through making things, and sharing them with others, we feel a greater connection with the world and more engaged with being more active in the environment rather than sitting back and watching" (Gauntlett, 2010, p. 73). In a literature review on play in education, Hinthorne and Schneider conclude that "serious play is particularly good at enhancing critical reflection, encouraging knowledge exchange, and promoting innovative problem solving" (2012, p. 2806). However, not many studies exist about serious play in education, so a solid empirical basis for the validity of serious play in secondary school (Young & Murray, 2017) or in ESJ is still missing.

2.3 Social Empathy

Next to the elements of serious play, the first phase of the Fair Battles program contains thematic inputs and individual research. Here, students encounter real life stories of people's hardship in life. Thus all three elements – serious play, individual research and thematic inputs – are intended to create an emotional connection between the students and the topic at hand. In a parallel learning process, the goal is to achieve greater cognitive understanding of the complexity of the global economy. This combination of empathic feelings with cognitive understanding of the context has been defined by Segal (2011) as *social empathy*: "Social empathy is the ability to understand people by perceiving or experiencing their life situations and as a result gain insight into structural inequalities and disparities." (p. 266) It combines the three elements of empathy, contextual understanding and social responsibility (Segal, 2007). Segal states that "the assumption upon which the concept of social empathy is built is that with socially empathic feelings and knowledge, people are more inclined to work to promote social and economic justice and social well-being." (Segal, Wagaman, & Gerdes, 2012, p. 544) However, social empathy – with Segal as its main proponent - is only one of many concepts of empathy. Critical voices (i.e. Eriksson & Englander, 2017) state that social empathy is insufficiently and too narrowly conceptualized, relying too heavily on simulation theory. Further research will have to show if its place in ESJ is warranted.

2.4 Service Learning

In ESJ, Zembylas (2013) advocates for "reclaim[ing] altruism by inspiring small-scale actions of solidarity that constitute students as active participants of community life" (p. 515). Such service learning projects can thus become an emotional valve and a catalyst for civic learning for students. The Fair Battles program utilizes service learning to motivate the participants to contribute positively to their community through so-called "Fair Actions", small-scale projects that are chosen, planned and implemented by the students themselves. According to Mitchell (2008), service learning programs "encourage students to see themselves as agents of social change, and use the experience of service to address and respond to injustice in communities" (p. 51). Meyers (2009) sees service learning as a pedagogical tool "to promote social justice, and [to] encourage both students' personal development and social engagement.…Service-learning can help students become more self-aware, appreciative of diversity, and agents of social change" (p. 373). While service learning has many proponents, critical voices such as Butin's (2003) find that there is "limited community impact of service learning…limited empirical evidence for defining and articulating best practices that lead to meaningful and sustained student outcomes, and…difficulty of rigorous and authentic assessment of service-learning outcomes." (p. 1674). In this study however, the focus of evaluation was not on the impact of service learning projects on the community, but its benefits for fostering the students sense of agency.

3 METHODS

3.1 Evaluative framework

The exploratory evaluation was planned by using the RUFDATA framework (Saunders, 2000). According to Rogers' (2008)

criteria, the Fair Battles program can be characterized as both a complicated (many parts) and complex (emergent) program. Focusing on a single case seemed appropriate to evaluate such a multi-faceted program. For an "intensive detailed description and analysis of a single project, program, or instructional material in the context of its environment" (Frechtling & Sharp, 1997) can give voice to key stakeholders (students, teacher, principal, program coach) instead of just the most influential ones (Saunders, 2006). As an evaluative framework, Kirkpatrick's (1996) four-level model was chosen to organize the data. Kirkpatrick's levels are reaction criteria, learning criteria, behavior criteria, and results criteria. Praslova (2010) shows how the Kirkpatrick model can be applied to educational research. Despite some critique of the model (i.e. Holton, 1996; Bates, 2004), it seemed especially apt to evaluate an educational program holistically on the emotional, cognitive and behavioral level as well as the impact it has on its wider context.

The study is based on an ontologically relativist and epistemologically interpretivist framework (Grix, 2002; Landridge, 2007), implying that there is not one objective and true view to be found and that the researcher co-creates meaning together with the key stakeholders when interpreting the data.

3.2 Design

A mixed methods design was chosen to evaluate the program. Mixed methods can be defined as "the class of research where the researcher mixes or combines quantitative and qualitative research techniques, methods, approaches, concepts or language into a single study" (Johnson & Onwuegbuzie, 2004, p. 17). Thus the quantitative survey results could be corroborated and complemented by key stakeholder voices (students, teacher, principal, program coach). This allowed for evaluating the impact of the program on both the participating students as well as on the school as a whole.

3.3 Research questions and instruments

The main research question of the study was: *What is the impact of the Fair Battles program on the individual participant on a cognitive, behavioral and emotional level?* Secondary questions were: a) *Is there a positive impact of the program on the class or school as a whole?* b) *Were the pedagogical elements of serious play and service learning seen as important or meaningful by key stakeholders?*

The pre- and post-program student survey was developed based on concepts that stood out in the literature review, in collaboration with a secondary school teacher, an education for social justice expert within the Swiss BNE network, and an education for social justice project manager. The goal was to create a questionnaire that would assess whether the program achieved its goals or not. All questions were based on the collaborators' professional experience, on their experience with the program as well as on similar existing surveys (Icon Kids, 2017; Tully, Krug, & Wienenfoet, 2011). The items in the questionnaire seemed to either have face validity or were based on previously validated surveys. Due to the small sample size, the survey was not validated in its current form. The survey items were rated on a 4-point Likert-type scale from zero ("strongly disagree") to three ("strongly agree") and consisted of 24 medium length (10 – 20 words) questions, arranged in 8 item clusters covering the following topics: knowledge about people's working conditions in different countries , knowledge about the origin of various articles available in Switzerland, knowledge about possibilities for actively promoting fairer product production, knowledge about Swiss laws around fair production, talking with others about fair consumption, promoting fair consumption, buying fair trade products, and not buying an article if they knew it was produced unfairly. The hypothesis was that students would on average score significantly higher on all items in the post-program assessment compared to the pre-program assessment. For the data analysis, item clusters one to four were grouped into the "learning criteria" variables (see Table 1). Items five to eight were grouped into the "behavior criteria" variables according to Kirkpatrick's (1996) model.

In the qualitative interviews, the participants were asked the following main questions:

(1) What did you like / not like about the program?

(2) What have you learned through the program?

(3) What have you personally taken away from the program?

(4) Has the program somehow changed / affected your personal life?

(5) Did you observe an impact of the program on the school / class?

Interview questions were developed in collaboration with the various program coaches and with the program manager, based on previous experiences with the program. The interview schedule was piloted with three former program participants.

3.4 Recruitment

The entire student cohort that had participated in the program was asked to complete the pre- and post-program survey (n=16). All key program stakeholders were invited to participate in the interviews. Out of the initial student group of 16, ten students (n=10/16; 62.5%) volunteered to participate. Also, the teacher who was mainly involved in the program, the school's principal and the coach who helped implement the program agreed to participate.

3.5 Participants

The students were all 15-16 years old at the time of the post-program survey and interviews. There were 6 girls and 4 boys who participated in the interviews. Though of different ethnic backgrounds, they had all grown up in Switzerland and spoke German as their mother tongue. Both the principal and main teacher were Caucasian Swiss males. The program coach was a Canadian-Swiss.

3.6 Data collection

Pre- and post-program surveys were administered during class to all the students who had participated in the program (n=16). Qualitative interviews also took place during class time. Once ethical approval had been obtained from the University of Lancaster and the participants (and the students' parents) had given written consent, a semi-structured interview of 20 - 30 minutes was conducted with each participant. In the interviews, the students were asked open-ended questions in the following areas: a) positive and negative memories of the program b) if and how it had a lasting impact on them personally c) if and how it had a lasting impact on their environment (class, school etc.). The teacher, principal and program coach were asked about the impact they had observed on the students. The interviews were conducted in German. All interviews were recorded and transcribed verbatim.

3.7 Ethical issues

During the interviewing process, it was important to be aware of possible power dynamics between the researcher and the students. One important aspect was to make the students feel comfortable. The interviews took place in the room adjunct to their main classroom while class was in session. The door of the interview room remained open during the interview. Participation was entirely voluntary, and the students were encouraged to speak their mind and reminded that there were no wrong answers. The students all seemed eager to participate in the interviews, and it felt like they were speaking freely as well about the negative aspects of the program.

3.8 Data analysis

The qualitative data were analyzed by conducting descriptive statistical analyses as well as paired-samples T-tests to compare the means of the 8 item clusters in the pre-program and post-program survey, using version 23 of IBM's SPSS Statistics. Since this was an exploratory pilot study with a very small number of participants, no statistical tests to determine reliability and validity of the questionnaire were conducted.

Qualitative interview transcripts were analyzed using template analysis (King, 1998). "Template Analysis is a form of thematic analysis which emphasizes the use of hierarchical coding but balances a relatively high degree of structure in the process of analyzing textual data with the flexibility to adapt it to the needs of a particular study. (Brooks, McCluskey, Turley, & King, 2015) Template analysis is useful in conducting content analysis when using an existing theoretical framework for data analysis (see Au, 2007; Minnaar, 2013; Ray, 2009; for examples of template analysis in educational research). "Central to Template Analysis is the development of a coding template, which summarizes themes identified by the researcher(s) as important in a data set, and organizes them in a meaningful and useful manner." (Brooks & King, 2014, p. 4) In this study, Kirkpatrick's (1996) four level model was used as an a priori coding template. "Template analysis allows the researcher to define some themes in advance of the analysis process, referred to as 'a priori' themes.… This is usually because a research project has started with the assumption that certain aspects of the phenomena under investigation should be focused on." (Brooks & King, 2012, p. 3). The data were coded according to the four levels of Kirkpatrick's model (affective, cognitive, behavioral, results). Emerging themes were grouped into categories. The results section portrays the different categories, presents the quantitative data and direct quotes from the interview data (translated into English by the author) in order to achieve a "rich description of the data set" (Braun & Clarke, 2006, p. 83).

4 RESULTS

The results section is presented according to Kirkpatrick's (1996) four levels of the evaluative framework. The quantitative data only pertain to the learning criteria and the behavior criteria and are presented in those two sections. Results are first presented in a descriptive form and then further interpreted in the discussion section.

4.1 Reaction criteria

According to Kirkpatrick's (1996) model, the reaction criteria describe someone's affective reaction to a program (how they feel about it). To inquire about the reaction criteria in this study, the students were asked what positive and negative elements they would spontaneously remember about the program. The theme that was identified in the data was that the students mainly remembered the combination between fun and serious content. The words used many times to describe elements of the program were "awesome" (4 times; the foosball table, the fair actions, the fashion show), "cool" (6 times) and "fun" (5 times), but also "learned a lot" (3 times) and "very interesting" (4 times).

The following student's comment reflects this general feeling of hard work and learning that is at the same time enjoyable:

> I remember that we worked hard for the program. And it was definitely great fun.

More specifically, students mentioned the creative aspects that were part of the serious play element as a positive part. The quote below refers to the creation of the foosball table figure.

> We did a lot of crafts. I liked that a lot. But we also spoke about more serious topics. I found it a cool mix between something more serious, but also something that was fun.

The next statement reflects this combination in a slightly different light. The student found the contents at the same time "important" and "fun".

> It was a very cool program where we learned about sides that we didn't know about before, about production and sale. I find that important and it was also great fun.

These quote show that in the students' eyes, fun elements can be combined with hard work and with serious and important content. The affective impact of the program was also observed by the program coach, as shown in the results criteria section below. It becomes clear that Kirkpatrick's (1996) reaction criteria cannot be completely separated from the learning criteria. In general, the students' positive spontaneous recollections of what had stood out to them show that the program had a positive affective impact on the participants.

On the other hand, the negative elements that were remembered about the program did not add up to a consistent theme. Seven students mentioned something they did not like: "long days, too much independent research, program too long, had to invest too much free time, seeing really difficult things, talking to strangers during the fair action" and "getting in trouble for being late". One student reported that she was shocked about some content of the program. However, the analysis of her transcript made it clear that this ended up being a key experience for her in the sense of a "healthy shock":

> I was very much shocked when I read about the conditions in India. I'm from India myself, but from a very different place. And because I've grown up in Switzerland, I have not heard much about what's happening in some places in India.

Overall, the data analysis within the reaction criteria demonstrates that the play element in the program had the desired effect on students of maintaining a fun atmosphere while not avoiding the seriousness of the topic at hand. The service learning elements were also mentioned in a positive light, but less frequently.

4.2 Learning criteria

The learning criteria in Kirkpatrick's (1996) model describe the cognitive learning that took place during a program. For the learning criteria in this study, quantitative data analysis from the

student survey was mixed with qualitative interview data analysis. The results from the quantitative survey data show that after the program, the students knew significantly more about working conditions in different places in the world, about the origin of various articles they would buy in Switzerland, about the possibilities for taking action to promote social justice and about Swiss laws on fair production. Table 1 shows that all compared means have a significant difference statistically ($p = 0.02$ or smaller). All effect sizes (Pearson's r) account for large effects according to Cohen (1992).

The analysis of the interview data pertaining to the learning criteria revealed that the purpose of the program could be articulated very well and precisely by all students. Yet beyond learning new facts, five of the participants revealed a deeper understanding of how some people are disadvantaged and especially of their own advantaged lives because of the program, as shown in the student quotes below:

> I learned how good our lives are here and how bad other people's lives are. And that we should be content with the way we live here.

The following student's statement also reveals that the learning that occurred was not exclusively about facts, but also led to self-reflection:

> I still find the topic interesting. I mean [fair consumption] is not just about other people, it's also about myself in the end.

Many students (n=7) also reported on having learned genuinely new things that they deemed important or relevant to their own lives:

> I have seen things that I never thought were possible….It was good for us to see what is happening in other places in the world… I had no idea."

The serious play elements helped students understand global connections within the overall topic of fair consumption. This was mentioned by 3 students.

> I believe that because of the foosball table, we have seen better that there are people who are disadvantaged.

Finally, half the students reported that the learning impact of the program was sustained over some time after the program:

> I've been mulling over the topic [of social justice] ever since the program….I'm much more aware now.

The interview data confirm the findings from the student surveys that the program has had a significant learning effect.

4.3 Behavior criteria

The behavioral criteria of Kirkpatrick's (1996) model describe the impact of a program on the participants' behaviors (how their behavior changes or how they apply the new knowledge). For the behavioral criteria in this study, survey and interview methods were again mixed. The results from the quantitative survey data show that post program, the students would talk more with their friends about fair production, they would make more of an effort to promote fair production, they would buy more fair trade products, and they would refrain more often from buying a product when they knew that it was produced under unfair circumstances. Table 2 shows that all compared means have a highly significant difference statistically ($p = 0.001$). All effect sizes (Pearson's r) accounted for large effects according to Cohen (1992).

The two main themes that could be identified in the interview data pertaining to the behavioral criteria were: a) fair consumption becoming a topic at home, and b) impact on personal consumer habits. As for sharing about the program at home and how it chan-

Table 1. Survey results learning criteria items

	M pre-program	SD	M post-program	SD	*t*	*p*	*r*
Working conditions	1,13	1,025	2,13	0,619	-6,325	0,001	0,85
Origin of articles	1,19	1,109	2,38	0,5	-6,333	0,001	0,85
Possibilities for action	1,31	1,138	2,38	0,5	-5,506	0,001	0,82
Swiss laws	0,19	0,403	0,5	0,516	-2,611	0,02	0,56

Table 2. Survey results behavioral criteria items

	M pre-program	SD	M post-program	SD	*t*	*p*	*r*
Talking about fair production	1,06	1,063	1,88	0,619	-4,961	0,001	0,79
Promoting fair production	0,44	0,892	1,63	1,025	-6,333	0,001	0,85
Buying fair trade	1	1,265	2,38	0,719	-6,214	0,001	0,85
Not buying unfairly produced articles	0,94	0,998	2,13	0,619	-7,251	0,001	0,88

ged some of the family consumer habits, the students shared that they talked about it quite intensively at home:

> Clothes…and food. I'm more aware…I've also shared about it much at home.

At the same time, the following student comment shows that sharing her new insights at home was not always a straight-forward endeavor. Again, we see that the behavioral criteria cannot be neatly separated from the results criteria. As the students' behavior changed, this would in turn affect the people around them.

> I've been driving my Mom crazy because I'm saying 'you know, this is not fair' and she says 'I know, but it's also cheaper'.

There was a clear impact on personal shopping habits among about half the students. They reported buying more fair trade products now: "I'm looking carefully when shopping…I'm more careful now when I know, it's child labor, it's not been paid fairly." Also, they reported realizing that there is a connection between price and production: "Yes, when I'm at a store, I think twice about what I buy…not just the cheap stuff, but I look at what's inside." Further, they also report not buying items that are clearly produced unfairly: "I'm more careful now when buying clothes. There is a store that I avoid, where I don't want to go anymore." These statements corroborate the survey results presented above.

Seeing the behavioral impact of the program is especially important because the main program goal was to sensitize students about fair consumption and their own consumer behavior with the hope that this would lead to lasting changes in their consumer habits. This was also observed by the program coach:

> I guess that half the students would say it had an effect on their lives. There is something that they think about when they're going to buy things. And some of the students will say 'I've decided not to buy one of these shirts because it is obvious to me from what we have researched that this shirt is not from a source that I can trust.'

4.4 Results criteria

The results criteria in Kirkpatrick's (1996) model describe the outcome of the program. According to Praslova's (2010) application of the model for educational settings, this means that there is a general benefit of the program not only for the student, but for society in general. The data of this study suggest that the program had an impact on the individual students, on the class, on the school and on the students' families. The program coach described the impact on class ambience:

> [One student] presented what she had found, and the class was a fairly rambunctious class, you know, it was not always easy to keep them focused, but the room was completely still and she told from her perspective what she discovered and it landed. And that's not always the case that you'll get this sort of respect for the truth…. Even for the mini-gangsters in the classroom that would actually prefer to say "nothing matters to me. I'm too tough to let this feeling get to me" you can still engage them enough, and then at a later point, playing with the playing field once it's finished, they come to a certain realization that you couldn't have required them to do by telling them or scolding them for not taking it seriously earlier (program coach).

The students' main teacher commented on how the serious play element initiated a positive classroom dynamic.

> The positive dynamic was incredible, the students' enthusiasm and passion about being in a competition, to make points and at the same time to serve other human beings.

Even the school's principal who was only marginally involved in the project reported how the program had a positive impact on the entire class:

> A great atmosphere; many children that played foosball with enthusiasm. It really bonded them together as a class. I can recommend it very much. I can congratulate every school that decides to do this program.

The program coach noted that there was a positive impact beyond the core class. The entire school seemed to have profited from the program:

> In the framework of this program, there was a new tone set that it actually is important to feel for somebody else, to show empathy….This even affected students that were not directly connected to the program. There was a certain vibe, you know, 'there are unjust situations here'. And that was something that other teachers started to notice in certain cases.

Two-thirds of the students reported that they brought the topic of fair consumption home and that it had an impact on their families:

> I've talked about it with my parents. They are now more attentive, especially when shopping for groceries.

Both qualitative and quantitative data show that the Fair Battles program had a definite impact on the participating students, on their class as a whole, and even to a certain degree on their school and their families. The following section discusses and interprets these findings in light of existing research literature.

5 DISCUSSION

The quantitative data show a significant increase of survey mean scores from program start to program end. They not only highlight an increased knowledge about fair consumption and social justice but also a student cohort that sees itself as increasingly active in applying this knowledge in practical ways. This indicates that students' cognitive knowledge and attitudes as well as their behavior can be positively impacted by education for social justice programs. Various authors such as Brown (2006) or Cross-Denny and Heyman (2011) come to a similar conclusion when analyzing social justice programs in the US. However, there is a dearth of research in this area in Switzerland.

The qualitative results demonstrate that Fair Battles is a holistic program that affects students emotionally and cognitively and impacts their behavior and their classroom culture. While the quantitative data allow for the general conclusion that the Fair Battles program is impactful, the qualitative data pinpoint toward some elements that seem especially relevant for the key stakeholders. The interviews' data pertaining to the reaction criteria indicate that in the students' eyes the serious play elements were particularly important (the "fun" elements). The ethical question

of whether it is okay to have fun while learning about very serious topics can be answered with "yes", for it was exactly the fun elements that met the students on an emotional level. While the strength of including such emotional elements in ESJ is promoted by other authors (i.e. Cohen, 2006), only few evaluations of ESJ programs that include serious play elements exist (i.e. Adelman, Rosenberg, & Hobart, 2016). This study's findings show that serious play has a positive impact on student learning in ESJ.

Most students (n=7) in this study also mentioned the service learning projects in a positive light. This leads to the assumption that a combination of serious play and service learning is especially impactful. However, the data only directly show that each element by itself was seen as important by the key stakeholders. What can be taken from this evaluation in this regard is that both the play and service elements seem to have an impact precisely because they engage the students in a holistic way. The program successfully engages emotions as "critical and transformational forces to produce better learning opportunities" (Zembylas & Chubbuck, 2009, p. 343). Thus i.e. both "becoming the Bangladeshi garment factory worker for the game" and "creating our own fair fashion show" create meaningful and contextualized memories that can be more easily recalled than i.e. mere facts about global injustice. This is precisely where social empathy comes into play. Many students reported this combination of emotional connection with cognitive understanding. This solidifies the assumption that serious play may enhance social empathy. The students cognitively understood the unfairness of global economic realities when looking at the changed foosball table while experiencing the unfairness at the same time during the game. The literature shows that ESJ is not effective when certain topics are only talked about in the classroom. This is especially true for highly emotional messages about human suffering. Zembylas (2013) warns against the *bystander effect,* which desensitizes students and can lead to indifference when they are overwhelmed by a topic. Tester (2001) uses the term *compassion fatigue* as a consequence of compassion overload in the classroom. Moyer-Gusé, Jain and Chung (2012) speak of *reactance* when ESJ programs have a counterproductive impact. Studer (2015) shows some impressive examples of the *boomerang effect*, when ESJ leads to the opposite of the intended consequences. Both serious play and service learning can help here by offering a valve for student emotions.

Another interesting discovery in the interview data is how much the topic of fairness impacted the students not only on an individual, but also on a communal level. That this impact was recognized by the principal who had not been involved in the program implementation is noteworthy. This learning transfer shows that the program impact went not only beyond cognitive learning, but also beyond individualized learning. The program became for this school what Hackman (2005) calls a "tool for action and social change" (p. 106). This aspect of the program could be strengthened further by more intentionally involving other members of the students' families or school community. Chupp and Joseph (2010) state that "maximizing the impact on the local community requires engagement of community members, not merely as recipients …, but as partners in the design, implementation, and assessment of the activity." (p. 209) Thus far, the impact of the program beyond the core class has been seen as a positive side-effect rather than a goal that would be actively pursued.

6 LIMITATIONS AND FURTHER RESEARCH

The present study is an exploratory pilot study. A new survey instrument was developed and then administered to a small number of program participants (n=16). Such small numbers generally do not allow for testing the validity and reliability of the survey items (Lance, Butts, & Michels, 2006). This clearly prevents generalizability of this study's results. As the Fair Battles program is usually only implemented with small groups of students, a next step would be to re-use the survey instrument with at least four more program cohorts in order to collect sufficient data. In a subsequent step, the survey instrument could be tested for reliability and validity, conducting e.g. factor analysis, Cronbach's alpha test for reliability, Kolmogorov-Smirnov test for normality. Similarly, the qualitative data analysis would profit from a larger number of interviews. While template analysis seems to be an effective way of analyzing the data and Kirkpatrick's model seems to offer an appropriate framework, a larger data set would allow for a more nuanced development of themes and for greater depth in analysis. A further recommendation for evaluating serious play in social justice education would be to develop a survey instrument with items that more deliberately cover the serious play elements. In the present study, the quantitative data only confirms that the program had a positive learning effect. The serious play elements that contributed to this learning effect were only analyzed in the qualitative interview data. The development of a valid and reliable quantitative instrument for serious play in social justice education would allow for greater generalizability of results and would help hopefully help in promoting serious play as a valid educational strategy in social justice education.

It is also important to note other limitations of the study. What does not come to light as much in the findings are problematic aspects of the program. It is important to reflect on some of the possible reasons. One of them could be that participants may have emphasized the facts in favor of the program because that's what they may have believed the interviewer wanted to hear. Another factor could be that looking back to the program, a memory bias occurred where the participants recalled the positive aspects more than the negative ones. Repeating this study with a different school class might shed more light on elements of the program that do not work well. This would be important for improving the program.

A further limiting aspect is that the participating student cohort was academically quite strong and may have therefore been particularly well suited for the program. Due to its complexity, the Fair Battles program tends to work better in such academically advanced classes and is at times less well received in academically weaker classes. The results therefore only reflect what happened in this one case. Also, while the study shows that all three pedagogical core elements (serious play, social empathy, service learning) were impactful, the collected data are not comprehensive enough to allow for the conclusion that the theory of change (serious play and learning about less fortunate people's stories leads to social empathy, which in turn leads to action) is correct.

Further research of different classes in different settings might shed a more varied and nuanced light on the impact of the program. In a subsequent evaluation, more nuanced questions could explore in greater depth how and why students' perspectives had been changed by the program. This could help to assess whether the program's theory of change is correct or not. In regards to adolescent development, it might also be interesting to reflect this study's findings in light of Kohlberg's (1984) moral development stages. A further investigation could look into what stage of moral reasoning the adolescents were at during the program and if the program helped them progress toward a higher stage.

7 CONCLUSIONS AND RECOMMENDATIONS

The general conclusion of this study is that the Fair Battles program has a positive emotional, cognitive and behavioral impact on students' and beyond students on the school as a whole. What can be learned in particular from this study is that a "lighter" approach to social justice education is appreciated by students and may have a more lasting impact than an entirely "serious" approach and that fun elements that open the students up to playful engagement with social justice should be included. Furthermore, it seems important to offer the program participants an outlet for becoming active in contributing some way to making the world a better place. Thus, the boomerang effect can be avoided, and students learn about the importance of their own agency. Social justice education needs to be holistic. It is about emotional exposure as much as it is about cognitive understanding. It is about changed behavior as much as it is about learning. Empathy that is not expressed in positive action runs the risk of turning into compassion fatigue.

The first recommendation for teachers and activists in school settings is to not just talk to students about social justice, but to reach them also on an emotional and experiential level. This may require alternative approaches such as simulations, role play etc. The second recommendation is to create space for action and to encourage action. (However, such action should not be forced upon the students. Service learning should be a bottom-up process that is as much student-directed as possible.) A third recommendation would be to include the community (school community and wider local community) in the planning in order to maximize the program's impact beyond the individual student. Finally, it is important to keep in mind that even difficult and heavy topics do not necessarily need to exclude the element of fun, but on the contrary, fun elements may make social justice issues accessible to students that would otherwise not be reached.

REFERENCES

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, *36*(5), 258-267. doi:10.3102/0013189X07306523

Adelman, M., Rosenberg, K. E., & Hobart, M. (2016). Simulations and social empathy: domestic violence education in the new millennium. *Violence Against Women*, *22*(12), 1451-1462. doi:10.1177/1077801215625850

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, *27*(3), 341-347. doi:10.1016/j.evalprogplan.2004.04.011

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. doi:10.1191/1478088706qp063oa

Brown, K. M. (2006). Leadership for social justice and equity: Evaluating a transformative framework and andragogy. *Educational Administration Quarterly*, *42*(5), 700-745. doi:10.1177/0013161X06290650

Brooks, J., & King, N. (2012). Qualitative psychology in the real world: the utility of template analysis. In *2012 British Psychological Society Annual Conference. London, UK.* Retrieved from http://eprints.hud.ac.uk/id/eprint/13656/1/Brooks__King_QMiP_2012_Final.pdf

Brooks, J., & King, N. (2014). Doing Template Analysis: Evaluating an End of Life Care Service. *Sage Research Methods Cases*. doi:10.4135/9781446273050135 12755

Brooks, J., McCluskey, S., Turley, E., & King, N. (2015). The utility of template analysis in qualitative psychology research. *Qualitative Research in Psychology*, *12*(2), 202-222. doi:10.1080/14780887.2014.955224

Butin, D. W. (2003). Of what use is it? Multiple conceptualizations of service learning within education. *Teachers College Record*, *105*(9), 1674-1692. doi:10.1046/j.1467-9620.2003.00305.x

Chupp, M. G., & Joseph, M. L. (2010). Getting the most out of service learning: Maximizing student, university and community impact. *Journal of Community Practice*, *18*(2-3), 190-212. doi:10.1080/10705422.2010.487045

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155. doi:10.1037/0033-2909.112.1.155

Cohen, J. (2006). Social, emotional, ethical, and academic education: Creating a climate for learning, participation in democracy, and well-being. *Harvard Educational Review*, *76*(2), 201-237. doi:10.17763/haer.76.2.j44854x1524644vn

Cross-Denny, B., & Heyman, J.C. (2011). Social justice education: Impacts on social attitudes. *Journal of Aging in Emerging Economies*, *3*(1), 4-16.

Eriksson, K., & Englander, M. (2017). Empathy in social work. *Journal of Social Work Education*, 1-15.

Frechtling, J., & Sharp, L. (Eds.) (1997). *User-Friendly Handbook for Mixed Method Evaluations*. Directorate for Education and Human Resources: Washington, DC.

Furrer, S. (2014, December 11). Modeschau gegen Hungerlöhne. *Oberbaselbieter* Zeitung, 1.

Gauntlett, D. (2010). Creativity, Participation and Connectedness: An Interview with David Gauntlett. In S. Sonvilla-Weiss (Ed.), *Mashup Cultures* (pp. 65–78). Vienna: Springer Vienna. doi:10.1007/978-3-7091-0096-7_5

Grix, J. (2002). Introducing students to the generic terminology of social research. *Politics*, *22*(3), 175–186. doi:10.1111/1467-9256.00173

Hackman, H. W. (2005). Five essential components for social justice education. *Equity & Excellence in Education*, *38*(2), 103-109. doi:10.1080/10665680590935034

Helbling, R., & Schwarz, V. (2015). Bildung für nachhaltige Entwicklung soll curricular werden. Zum Stand der Diskussion in der Schweiz. *Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik*, *28*(3), 27-32.

Hinthorne, L. L., & Schneider, K. (2012). Playing with purpose: using serious play to enhance participatory development communication. *International Journal of Communication*, *6*(24), 2801–2824.

Holton, E. F. (1996). The flawed four level evaluation model. *Human Resource Development Quarterly*, *7*(1), 5-21. doi:10.1002/hrdq.3920070103

Icon Kids (2017). *Trend Tracking Kids 2017: Ergebnisse zu High Interest Themen bei 6- bis 19-jährigen Kindern und Jugendlichen in Deutschland.* Munich, Germany: Icon Kids.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, *33*(7), 14-26.

King, N. (1998). Template analysis. In G. Symon & C. Cassell (Eds). *Qualitative methods and analysis in organizational research: A practical guide* (pp. 118-134). Thousand Oaks, CA: Sage Publications.

Kirkpatrick, D. L. (1996). Techniques for evaluating training programs. *Classic Writings on Instructional* Technology, *1*(192), 119-132.

Kohlberg, L. (1984). *The psychology of moral development: The nature and validity of moral stages (essays on moral development, Volume 2)*. Harper & Row.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, *9*(2), 202-220. doi:10.1177/1094428105284919

Landridge, D. (2007) *Phenomenological psychology: theory, research and method.* Essex, England: Pearson.

Meyers, S. A. (2009). Service learning as an opportunity for personal and social transformation. *International Journal of Teaching and Learning in Higher Education*, *21*(3), 373-381.

Minnaar, A. (2013). Challenges for successful planning of open and distance learning (ODL): A template analysis. *The International Review of Research in Open and Distributed Learning*, *14*(3), 81-108. doi:10.19173/irrodl.v14i3.1387

Mitchell, T. D. (2008). Traditional vs. critical service-learning: Engaging the literature to differentiate two models. *Michigan Journal of Community Service Learning*, *14*(2), 50-65.

Moyer-Gusé, E., Jain, P., & Chung, A. H. (2012). Reinforcement or reactance? Examining the effect of an explicit persuasive appeal following an entertainment-education narrative. *Journal of Communication*, *62*(6), 1010-1027. doi:10.1111/j.1460-2466.2012.01680.x

Praslova, L. (2010). Adaptation of Kirkpatrick's four level model of training criteria to assessment of learning outcomes and program evaluation in higher education. *Educational Assessment, Evaluation and Accountability*, *22*(3), 215-225. doi:10.1007/s11092-010-9098-7

Ray, J. M. (2009). A template analysis of teacher agency at an academically successful dual language school. *Journal of Advanced Academics*, *21*(1), 110-141. doi:10.1177/1932202X0902100106

Rogers, P. J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, *14*(1), 29-48. doi:10.1177/1356389007084674

Saunders, M. (2000). Beginning an evaluation with RUFDATA: theorizing a practical approach to evaluation planning. *Evaluation*, *6*(1), 7-21. doi:10.1177/13563890022209082

Saunders, M. (2006). The 'presence' of evaluation theory and practice in educational and social development: Toward an inclusive approach. *London Review of Education*, *4*(2), 197-215. doi:10.1080/14748460600855526

Schrage, M. (2013). *Serious play: How the world's best companies simulate to innovate*. Harvard Business Press.

Segal, E. A. (2007). Social empathy: A new paradigm to address poverty. *Journal of Poverty*, *11*(3), 65-81. doi:10.1300/J134v11n03_06

Segal, E. A. (2011). Social empathy: A model built on empathy, contextual understanding, and social responsibility that promotes social justice. *Journal of Social Service Research*, *37*(3), 266-277.

Segal, E. A., Wagaman, M. A., & Gerdes, K. E. (2012). Developing the social empathy index: An exploratory factor analysis. *Advances in Social Work*, *13*(3), 541-560.

Studer, A. (2015). Das Problem des Moralisierens im Unterricht. In S. Bietenhard, D. Helbling, & K. Schmid (Eds), *Ethik, Religionen, Gemeinschaft: Ein Studienbuch* (pp. 164 – 172). Verlag.

Tester, K. (2001). *Compassion, morality and the media*. Buckingham: Open University Press.

Tschopp, S. (2014, October 31). An diesem „Töggelikasten“ ist Geiz nicht geil. *Basellandschaftliche Zeitung*, 31.

Tully, C., Krug, W., & Wienefoet, V. (2011). Jugendkonsum in globalen Handlungsbezügen. *Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik*, *34*(4), 13-19.

Weisberg, D. S., Hirsh-Pasek, K., Golinkoff, R. M., Kittredge, A. K., & Klahr, D. (2016). Guided Play Principles and Practices. *Current Directions in Psychological Science*, *25*(3), 177-182. doi:10.1177/0963721416645512

Young, F., & Murray, G. (2017). Designing for Serious Play. In *Multidisciplinary Perspectives on Play from Birth and Beyond* (pp. 163-179). Springer Singapore. doi:10.1007/978-981-10-2643-0_10

Zembylas, M. (2013). The “crisis of pity” and the radicalization of solidarity: Toward critical pedagogies of compassion. *Educational Studies*, *49*(6), 504-521. doi:10.1080/00131946.2013.844148

Zembylas, M., & Chubbuck, S. (2009). Emotions and social inequalities: Mobilizing emotions for social justice education. In *Advances in Teacher Emotion Research* (pp. 343-363). Springer US. doi:10.1007/978-1-4419-0564-2_17

How to cite this article: Wartenweiler, T. (2018). Serious play in education for social justice - An exploratory study. *Journal of New Approaches in Educational Research*, *7*(1), 61-69. doi: 10.7821/naer.2018.1.260