

Article

# Other- versus Self-Referenced Social Impacts of Events: Validating a New Scale

Marijke Taks <sup>1,\*</sup>, Daichi Oshimi <sup>2</sup> and Nola Agha <sup>3</sup>

<sup>1</sup> Faculty of Health Sciences, University of Ottawa, Ottawa, ON K1N 6N5, Canada

<sup>2</sup> Department of Sport & Leisure Management, Tokai University, Hiratsuka, Kanagawa 259-1292, Japan; oshimidaichi@tsc.u-tokai.ac.jp

<sup>3</sup> Sport Management Program, University of San Francisco, San Francisco, CA 94117, USA; nagma@usfca.edu

\* Correspondence: marijke.taks@uottawa.ca

Received: 15 November 2020; Accepted: 7 December 2020; Published: 9 December 2020



**Abstract:** Publicly funded sport events are partially justified based on positive social impacts. Past research generally measured social impact for a generic and global “other” with claims such as “Events create new friendships in the community”. These other-referenced (OR) social impacts are generally higher pre-event than post-event and are inflated for both methodological and theoretical reasons. In the pre-event period of the Tokyo 2020 Olympic and Paralympic Games, we empirically tested OR items compared to self-referenced (SR) items, such as “Because of the event, I create new friends in the community” and allowed projection bias to vary between scales. Results of the experiment between an OR-Social Impact Scale (OR-SIS) and a similar SR-SIS confirmed OR-measures to be significantly higher than SR-measures. While artificially inflated OR scores may be useful for event organizers and politicians to gain support for hosting, estimates based on circumscribed self (SR) are a methodologically appropriate measurement of social impact.

**Keywords:** media framing; olympic and paralympic games; projection bias; public discourse; recency bias; Tokyo 2020

## 1. Introduction

Publicly funded sport events require the demonstration of substantial and sustainable positive outcomes for host communities. From a social welfare perspective [1], an efficient allocation of resources suggests that tax-payers, as major co-payers (e.g., 80% of the Olympic Games; [2]), should receive beneficial outcomes. Nevertheless, residents’ quality of life may be directly impacted by the event, be it positive or negative, before, during, and/or after the event. Research has demonstrated that hosting one-off major sporting events generally results in negative ecological impacts [3], as well as a lack of substantial economic impact [4–7] and sustainable tourism impact [8]. This has shifted the justification for hosting mega-events towards trying to demonstrate positive benefits from other features of events such as urban planning [9], or less tangible features such as social impacts [10] and sport participation impacts [11] affecting local communities and residents after the event has taken place. Kellett, Hede, and Chalip emphasized “the need for greater attention to the social value of events and the relationship between events and their host community” [12] (p. 117). Thus, measuring social impacts from events for residents is important to potentially identify if and how event-related outcomes occur and are sustained in host communities. It is equally important to accurately measure those social impacts, but there is no consensus on how to best measure social impact.

Most research to date has asked residents about their perceptions of social impact (e.g., [13–15]). As West and Kenny stated, “if researchers are interested in accuracy of perception, they need to operationally define and measure it” [16] (p. 359). Previous research has defined perception as attitudes

or opinions regarding the social impact of events for a generic and global “other” with questions such as “Events create new friendships in the community”. We refer to these as other-referenced (OR) to distinguish them from questions operationalized through a self-referenced (SR) lived experience, for example, “Because of the event, I created new friends in the community”.

There are two concerns with OR perceptions. First, in relation to time, most research to date has found social impacts measured in terms of OR items to be higher pre-event compared to post-event [13–15,17–19]. These greater predicted pre-event OR values can be criticized as over-estimates because the post-event values are more precise measurements since the actual items being asked about have already occurred. While these pre-event perceptions of social impacts may well serve event organizers and/or local governments to gain event support (e.g., [20,21]), improving ways to measure pre-event social impact more precisely is warranted.

A second area of concern is the wording or point of reference. Questions about a global other (OR) and circumscribed self (SR) lead to different answers [16]. For instance, Fredline, Deery, and Jago [22] found consistently higher OR scores than SR scores for social impacts of tourism. Because of social projection bias [23], a respondent will report that others are more likely to hold their views, resulting in higher scores than if asked about themselves. Therefore, we suggest that the point of reference taken (i.e., circumscribed self versus a global other) can calibrate and provide a more accurate measurement of social impacts pre-event.

To account for these concerns with OR social impact scales, the purpose of this study is to re-word existing questions from OR to SR and to test each simultaneously. Whereas others have looked at the same type of questions (OR) in two different time periods pre- and post-event [13,15,17], our research design looks at different types of questions (OR vs. SR) in the same time period (pre-event), keeping time constant. By holding the time constant to the higher pre-event period, other influencing factors, such as media framing, level of excitement, etc., are also kept constant, allowing us to isolate the effect of words in obtaining a more precise estimate of social impact. Any difference in scores will be because we re-orient respondent perceptions of social impact from others to self.

In what follows, we first elaborate on the social impact of events and the theories underpinning their measurement. We provide an overview of how past research has measured the social impact of events. We discuss the difference between OR and SR outcomes in terms of two distinct points of reference, namely: (1) words (OR vs. SR items), and (2) time (pre- vs. post-event). Next, we conduct an experiment comparing two similar social impact scales for which only the wording has been changed; one scale reflects a Self-Referenced Social Impact Scale (SR-SIS), while the other reflects an Other-Referenced Social Impact Scale (OR-SIS). From data collected in the pre-event time of the Tokyo 2020 Olympic and Paralympic Games (OPG), we validate the scales and present practical insights on the anticipated social impact of this mega event. Finally, the results allow us to advocate for future use of the SR-SIS as a measurement of social impact.

## 2. Social Impact of Sport Events

Mathieson and Wall’s definition of social and cultural impacts of tourism can be easily applied to sport as the ways in which sport events contribute to “changes in value systems, individual behaviour, family relationships, collective lifestyles, safety levels, moral conduct, creative expressions, traditional ceremonies and community organisations” [24] (p. 133). Ritchie [25] described perception of social impact of sports events as enhanced local pride, a sense of community, and enthusiasm for the community among residents of a host community. Holmes and colleagues [26] distinguished between positive and negative social impacts as well as between short-term (during the event) and long-term social impacts (after the event). Examples include prestige for the host community (positive, short-term), increased traffic (negative, short-term), building community pride and community cohesion (positive, long-term), and community alienation (negative, long-term).

### 2.1. Theories Underpinning Scales for Social Impact of Events

Different theories have informed the development of various social impact scales (see Table 1, column 2). Some scales developed based on these theories are predominantly OR-based while others are more SR-based. Social Exchange Theory (SET) is perhaps most commonly used, assessing that residents are willing to become involved in a social exchange if the perceived benefits outweigh the costs of involvement (e.g., [27]). In other words, if residents believe that there will be an overall positive impact from hosting an event locally, they will be in favour of hosting it [28]. The Theory of Reasoned Action (TRA) analyzes the interrelationships between beliefs, attitudes, intention, and behavior with the intent to predict and understand individuals' behavior [29,30]. According to Social Representation Theory (SRT), perceptions of impacts are formed through "direct experiences, social interaction and other sources of information, such as the media" [31] (p. 147). Thus, this requires some form of interaction between residents (i.e., an experience) and information sources to shape perceptions [32]. Overall, researchers utilizing SET, TRA, and SRT predominantly use individuals' perceptions of others to measure social impacts such as overall community values, expectations, and beliefs.

Community Attachment Theory (CAT) posits that OR perceptions of community residents toward hosting a sport event are largely impacted by the extent to which individuals feel connected to and involved in the community at large. Trust and reciprocity, concepts which relate to social capital, are considered to be important factors (e.g., [33,34]). Social Identity Theory (SIT) refers to values and emotional attachment associated with memberships in a particular group (i.e., connection and involvement; [35,36]). Social Anchor Theory (SAT) is used to explain psychological benefits that are sustained among group members after experiencing an event, and consists of social capital and social identity [37]. While different authors define social capital in different ways (e.g., [38]), in the context of events, social capital reflects how the event affects the residents and their relationship to the community [33]. The focus of CAT, SIT, and SAT on feelings requires the application of SR items. In constructing scales of social impact, some researchers have used one theoretical framework [39], while other researchers have integrated several [40].

### 2.2. Measuring Social Impact of Events

Regardless of the theories used to create previous scales, social impact is intangible in nature and therefore challenging to capture. Column 3 in Table 1 provides an overview of measurements of social impacts in the various studies, and illustrates the use of OR and/or SR items and constructs in each of these studies. Most of the studies presented in Table 1 are standalone projects and isolated cases. From the table it is evident that: (1) there is no unified, accepted social impact scale for sport events; and (2) many social impacts rely heavily on OR-based measurements of residents, or occasionally employ a mix of both OR and SR items (see Table 1).

While there is no globally accepted social impact scale, there are some common, recurring dimensions that capture social impact such as: community spirit, social cohesion, social capital, community involvement, disorder and conflict, and feelings of (un)safety (see Table 1, column 3). Community spirit, the one common dimension across all studies presented in Table 1, refers to feelings of pride and happiness instilled by an event. Some authors call it psychic income (e.g., [33,37]), or a psychological "feel-good-factor" [45]. Social cohesion represents people's perceptions as to how an event affects connectedness between individuals in the community [42,43], while social capital reflects how the event affects the residents and their relationship to the community (e.g., [33,38]). Community involvement indicates to what extent the community is involved with hosting an event and whether their input is solicited and appreciated [14,19]. Disorder and conflict gauges to what extent an event disrupts residents' daily lives [13,22]. Feelings of (un)safety enquire about peoples' feelings of (un)safeness [14]. These six commonly used constructs serve as a basis for this study. Sport participation can be added to this list, given its recent prominence as an acclaimed social outcome of events [11,46].

**Table 1.** Overview of Studies, Theories, and Measurements of Social Impact \* and the Use of Other-Referenced (OR) and Self-Referenced (SR) Items.

Study	Theory(ies)	Examples of Factors and Related Items	Results of Social Impact
<b>Exclusively perception-based</b>			
Balduck et al. (2011) [13]	Social Leveraging Framework; 33 items <sup>(1)</sup> ; 7 Factors	Cultural interest and consolidation: [Event] Reinforced community spirit (OR) External image enhancement: [Event] Enhanced recognition of Ghent internationally (OR) Disorder and conflicts: [Event] Brought conflicts and antagonism between visitors and residents (OR)	Five of the seven factors score above the indifference point of 4 in the pre-event survey. External image enhancement: $M = 5.31$ ( $SD = 1.11$ ) Disorder and conflict: $M = 4.90$ ( $SD = 0.97$ )
Liu et al. (2014) [39]	SET; 35 items <sup>(2)</sup> 7 factors:	Networks and co-operation: OG promoted international cooperation (OR) Environment: OG improved awareness of environment protection (OR) Sport and health: OG promoted sport participation (OR)	All factors score above the indifference point of 3; e.g., Psychic Income and Social Capital: $M = 3.89$ ( $SD = 0.65$ ) Sport and Health: $M = 3.55$ ( $SD = 0.78$ )
Kim et al. (2015) [14]	SET; 23 items <sup>(1)</sup> ; 6 Factors	Economic benefits: [Event] Increased trade for local business (OR) Community development: [Event] Increased interest in internat. sport events (OR) Security risks: [Event] Increased risk of terrorism (OR)	Overall score of positive social impacts dimension $M = 4.33$ ; overall score of negative social impacts dimension $M = 4.10$ ; thus, all above the indifference point of 4.
Mao and Huang (2006) [41]	SET; 27 items <sup>(2)</sup> ; 6 Factors	Community social development: [Event] Reinforced cohesion of community (OR) Personal leisure opportunities and new experiences: [Event] Brought emotional experience to one's life (OR) Community social pressure: [Event] Declined the moral standard (OR)	All items of positive impact scored above the indifference point of 3 (range between $M = 3.45$ and $M = 4.17$ ); Items of negative impact: 6 items scored under the indifference point of 3; 7 items scored slightly over the indifference point of 3
<b>Mainly perception-based</b>			
Ma et al. (2013) [19]	SET, SRT; 23 items <sup>(2)</sup> (21 OR and 2 SR); 4 Factors	General benefits: Because of the World Games I will have more recreational opportunities (SR) Community involvement: I support the World Games because of its vital role in our community (SR) Negative impacts: The World Games will result in traffic congestion (OR)	Scores vary according to segmentation group (neutral, moderately, and positive)

Table 1. Cont.

Study	Theory(ies)	Examples of Factors and Related Items	Results of Social Impact
<b>Mainly perception-based</b>			
Huang et al. (2016) [30]	SET, TRA; 35 items <sup>(2)</sup> (29 OR and 7 SR), 8 factors	Social benefits: <i>Community cohesion</i> (OR) Social costs: <i>Deterioration of social order</i> (OR) Intention to support major events: <i>Support the F1 event as a resident</i> (SR)	Social benefit factor $M = 3.68$ ( $SD = 0.61$ ); above indifference point of 3 Social cost factor $M = 2.93$ ( $SD = 0.61$ ); below indifference point of 3
Kim and Walker (2012) [42]	SET, psychic income; 22 items <sup>(1)</sup> (12 OR and 10 SR); 5 Factors	Community pride as a result of enhanced image: [ <i>The City</i> ] <i>Gained positive recognition</i> (OR) Event excitement: <i>Enjoyed interacting with visitors</i> (SR) Pride in efforts to improve community infrastructure: <i>Improved the quality of police and fire services</i> (OR)	Of the 49 items in the pilot study, 44 items scored higher than the indifference point of 4; 25 scored 5 or higher.
<b>Balanced approach OR and SR</b>			
		<ul style="list-style-type: none"> <li>➤ How has this affected <b>the community as a whole?</b> (14 items, all OR)</li> <li>➤ How has this affected <b>your personal quality of life?</b> (14 items, all SR)</li> </ul>	
Fredline et al. (2006) [22]	Sustainable Tourism; 14 items (all asked OR and SR form)	Disruption: <i>Tourism disrupts the lives of local residents and creates inconvenience.</i> Delinquent Behaviour: <i>Tourism is associated with some people behaving inappropriately, perhaps in a rowdy and delinquent way, or engaging in excessive drinking or drug use or other criminal behaviour.</i> Pride: <i>Tourism makes local residents feel more proud of their town and makes them feel good about themselves and their community.</i>	Effect on personal quality of life was limited and scores on “personal impact” scored consistently lower than “scores on community impact”
<b>Mainly experience-based</b>			
Heere et al. (2013) [35]	SIT; 15 items (mainly SR); 6 Constructs CAT -Social Capital; 21 items (4 OR, and 17 SR); 5 Constructs	Public evaluation: <i>Overall, people hold a favorable opinion of my [nation].</i> (OR) Interconnection to self: <i>When someone criticizes my [nation], it feels like a personal insult.</i> (SR) Behavioral involvement: <i>Changes impacting my [nation] will change my life.</i> (SR) Collective action: <i>Do you regularly attend local community events?</i> (SR) Trust and safety: <i>Does your local area have a reputation for being a safe place?</i> (OR) Value of life: <i>Do you feel valued by society?</i> (SR)	Constructs score well above the indifference point of 4; three constructs (social connections, diversity tolerance and value of life) score above the indifference point of 3; two constructs (trust and safety and collective action) score around or below the indifference point of 3

Table 1. Cont.

Study	Theory(ies)	Examples of Factors and Related Items	Results of Social Impact
<b>Mainly experience- based</b>			
Gibson et al. (2014) [33]	Psychic Income; <b>4 items, all OR; 1 construct</b> CAT-Social Capital <b>21 items (4 OR, and 17 SR); 5 Constructs</b>	Psychic income: <i>The 2010 World Cup will increase community spirit and pride (OR)</i> Psychic income: <i>The 2010 World Cup will increase feelings of national pride and patriotism (OR)</i> (similar to Heere et al., 2013)	Psychic income, all OR, all scores above the indifference point of 3 Social Capital, Collective action, all items SR, all score below the indifference point of 3; all other constructs (mainly OR) score above the indifference point of 3
Inoue and Havard (2014) [43]	SET & Balance Theory; <b>21 items (8 OR and 13 SR); 8 constructs</b> related to social impact	Event Excitement: <i>The [Event] increased my interest in golf (SR)</i> Community Pride: <i>Memphis gained positive recognition by hosting the [Event] (OR)</i> Social Camaraderie: <i>Attending the [Event] allowed me to develop warm relationships with others (SR)</i>	All items score above the indifference point of 4, but all OR items scored above 6.0, and the SR items scored below 6.0 (except for 2 items: “I really enjoy following golf” and “supporting this cause is important to me”)
Schlegel et al. (2017) [44]	Subjective Well-being, and atmosphere; <b>12 items (3 OR and 9 SR); 2 Factors</b>	Subjective well-being: <i>I have felt calm and relaxed (SR)</i> Perceived atmosphere: <i>There are amazing vibes (OR)</i>	Scores of the three OR items are all above 4.0; all other items (SR) score below 4.0 (during the event)
Oja et al. (2018) [37]	SAT; <b>4 Factors (*) (mainly SR)</b>	Social capital: <i>The people I interact with here in Kansas City would share their last dollar with me (SR)</i> Community and neighborhood identification: <i>I am very interested in what others think about Kansas City (SR)</i> Psychic income (see Kim and Walker, 2012 above)— <b>mixed OR and SR</b>	Items means not comparable; participants’ excitement and neighbourhood identification decreased after the event; social capital and team identification increased after the event.

Note. \* All nine social impact of events articles published in *JSM*, *SMR* and *ESMQ* since 2013 are included in Table 1. Two key articles were added from *Tourism Management*, as well as Fredline et al.’s foundational report [22] and Balduck et al.’s article [13] which was the first social impact paper to appear in *ESMQ*. CAT = Community Attachment Theory (Social Capital); SAT = Social Anchor Theory; SET = Social Exchange Theory; SIT = Social Identity Theory; SRT = Social Representation Theory; TRA = Theory of Reasoned Action; OG = Olympic Games; SR = Self-Referenced; OR = Other-Referenced; <sup>(1)</sup> = number of items not provided. Background colour clearly delineate the four distinct approaches to measuring social impact; Italic words represent “Items”.

### 2.3. Point of Reference in Measuring Social Impact

The fact that social impact studies, including the ones in Table 1, use non-unified social impact scales and pertain to unique events, hinders comparing actual results in numerical terms. However, from column 4 in Table 1 it is apparent that studies utilizing OR measures generally demonstrate average scores above the indifference point for positive social impact perceptions. Results for OR negative social impacts are more variable with some scoring above the indifference point, indicating negative social impact perceptions [13,14]. Other studies show negative social impact with scores both above and below the indifference point (e.g., [41]), or solely below the indifference point [30]. Low scores of negative impact indicate a positive perception. Nevertheless, studies relying on OR perceptions overly demonstrate positive social impact outcomes [13–15,17–19].

In contrast, studies that embedded some SR items demonstrate a range of scores above and below the indifference points (e.g., [33,35]). In cases where SR items score above the indifference point, the measures are consistently lower than OR items (e.g., [22,43,44]). Other studies show SR items scoring below the indifference point and OR measures above the indifference point (e.g., [33]). In both cases, measures with an OR orientation consistently surpass those relying on an SR orientation. Moreover, studies which mainly use SR measures report many more social impact measures below the indifference point.

To try and explain why OR estimates are consistently greater than SR estimates of social impact, we begin by noting the common phenomenon that public perception of positive event impacts remains strong despite research demonstrating that claims of positive economic impact and/or enhanced city image are not supported by the evidence [47]. This disconnect can be explained by noting that events are frequently hosted based on political grounds [48]. Politicians, senior sport administrators, and corporate leaders involved in bids and/or host committees use political discourse to emphasize positive impacts and outcomes of hosting the events, including economic, tourism, city image, social and/or sport participation outcomes [49]. Well-planned public campaigns often highlight the multiple benefits from hosting events, while neglecting the negative outcomes with the intent to gain the support of residents [50].

Similarly, Sant and Mason [49] demonstrated how public opinion is influenced through media framing, which highlights some aspects of an event over others. “Framing refers to the process by which people develop a particular conceptualization of an issue or reorient their thinking about an issue” [51] (p. 104). For instance, framing may reorient the audience’s beliefs and attitude towards an event [49], thereby instilling certain expectations. For example, media emphasizing the millions of dollars an event will generate without reporting about the costs may engender a biased public opinion of positive economic impact from the event. The tendency to overstate the potential economic and social benefits of event hosting has been well documented in the literature [52] and results in public perceptions that are not always in alignment with actual facts. Media framing may influence both OR and SR expectations when estimating social impacts, especially pre-event, when the event experiences have not yet occurred.

Whereas media framing and public discourse are factors external to the individual, projection bias is internal to the survey respondent. Humans have a tendency to believe others hold similar views to themselves, a phenomenon called social projection [23]. When determining a response to an OR global statement such as, “The event increases social interactions in the community” individuals will use their own preferences (which may already have been influenced by external factors) to predict what others are likely to believe. Even though SR responses may be influenced by these external factors, OR responses are even more influenced because, as Van Boven, Judd, and Sherman note, respondents “perceive others as more likely to hold that stance” [53] (pp. 84–85, emphasis added). Thus, there is a theoretical explanation why OR outcomes are higher than SR outcomes.

#### 2.4. Point of Reference and Temporal Effects

##### 2.4.1. Same OR Question, Different Time

When using OR questions in different time periods, there is a strong tendency for OR social impacts to be higher pre-event compared to post-event [13–15,17–19]. In these studies, the post-event measures are presumed to be a better reflection of reality because the event has actually occurred and people are not guessing about the future. No studies argue that the lower post-event numbers are underestimates of the pre-event reality. In other words, time helps us better understand which estimates are closer to reality.

##### 2.4.2. Same SR Question, Different Time

When using SR questions in different time periods, there is more variation in pre-and post-event results [33,35,37,45]. For example, Kavetsos and Szymanski [45] studied the feel-good effect one year before, during, and three months after the 2006 FIFA World Cup, using a four-scale SR question: “On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?” They found no effect one year prior, the highest levels during, and a decrease three months after the event. In their study on the 2010 FIFA World Cup, Heere and colleagues [35] found higher SR scores post- compared to pre-event for private and public evaluation of the event. Gibson and colleagues [33], also studying the 2010 FIFA World Cup, found lower SR scores post-event for three social capital dimensions (i.e., collective action, social connectedness, tolerance of diversity). However, practical significance of this difference was regarded as moderate. The authors found the same SR score for two other social capital dimensions (i.e., trust and safety, and value of life). In their study on the 2012 Major League Baseball All-Star Game, held in Kansas City (MO), Oja and colleagues [37] found some SR increased post-event (e.g., social capital: “Interacting with Kansas Citians makes me want to try new things”), others decreased (e.g., Neighborhood identification: “I am very interested in what others think about Kansas City”). Given this wide variation in higher, lower, or the same scores for SR items, there is no ground to posit that SR items underestimate social impact.

We established in previous sections that OR measures are empirically and theoretically higher than SR measures and we can now establish that this is true whether the data are collected before, during, or after an event. For example, when data are collected in advance of an event, predicted perceptions of community impact on generic others (which have not occurred) will still be higher than predictions of self-experiences (even though they also have not yet occurred). While both OR and SR questions are subject to media framing, SR questions involve no projection bias.

##### 2.4.3. Different Questions, Same Time

Whereas others have looked at the same type of questions in two different time periods, the only study that used both OR and SR questions in the same time period is the study by Fredline and colleagues [22] in the context of tourism. The authors found OR outcomes to be consistently greater than SR outcomes. We developed an experimental research design to further compare SR and OR questions in the same time period. The following section explains the need for this investigation.

#### 2.5. SR Scales as an Improved Measure of Social Impact

According to Fredline and colleagues, social impacts “often have a differential effect on different members of the community” [54] (p. 23). The only way to capture these differential effects is to ask community members about their own experiences. Asking any human to evaluate a situation involves a level of subjectivity [22], but in determining levels of truth within human judgement, West and Kenny [16] make clear that circumscribed responses are more pragmatic and accurate than global ones.

Methodologically, consider that political polls obtain a representative sample of the population and then ask, “Will you vote for candidate x?” The foundation of inferential statistics is that the collective opinions of this sample can be extrapolated to the population. A poll will not ask, “Will others

vote for candidate x?" because asking someone to predict the opinions of others allows more error into the measurement.

Similarly, methodologically, a proper sample of the population includes all residents (not just event attendees or volunteers). When referring to social impact experiences, we want to know how a sport event actually affects a resident's life in general, regardless of whether they associate with the event (e.g., as volunteer or attendee) or not. The internal and subjective response residents have through direct or indirect contact with the event captures the spillover or externality effect the event has on the hosting community as a whole [55]. It is the collective response of all residents responding to questions about themselves that will provide the most accurate measure of social impact.

Moving forward at a practical level, we conduct an experiment comparing two similar social impact scales for which only the wording is different; one scale uses SR items and the other uses OR items. Depending on the timing of the data collection, these items can be reworded as future (expected experiences before the event), present (actual experiences during the event), or past tense (reflecting on past experiences after the event), as is consistent with past social impact scales that use multiple time frames [13,14,32]. In the context of this study, we test social impacts pre-event. Thus, items were worded in the future tense (expected social impacts). SR items are worded in the first person using "me", "my", and "I" to rate how respondents expect that the event will affect their lives personally. For example, when measuring social cohesion using SR, residents are asked to indicate their level of agreement with "the event will strengthen my relationships in the community" measuring expected self-experience. OR items are worded in the third person, reflecting a "generic other"; in this case, the social cohesion item reads, "the event will strengthen relationships in the community", measuring predicted perception of others.

In what follows, we first investigate how the social impact of events differs when measured based on two comparable scales, the OR-SIS (OR items only) and the SR-SIS (SR items only) in a pre-event period that holds political discourse and media framing constant but allows projection bias to vary between scales. This allows us to replicate the tourism findings of Fredline et al. [22] to investigate if OR-worded items result in higher outcomes than SR-worded items in the context of events.

### 3. Materials and Methods

#### 3.1. Context

The event selected for this study had to be significant enough to create some type of shock in the host community to attract attention from residents [56]. The selected event was the Tokyo 2020 OPG in Japan. The candidature file for this event proposed various programs to increase the social impact, including \$39 million USD for cultural development, \$12 million USD for Olympic education, and \$60 million USD for community development. The surveys were collected two years prior to the when the event was supposed to happen. At the time of data collection, there was no mention yet of the 2020 Tokyo Olympic and Paralympic Games being postponed (official announcement made by the International Olympic Committee on 24 March 2020, <https://tokyo2020.org/en/>). With the OPG being the largest international sporting event, attracting attention from residents in the host region in the years leading up to this mega-event, the event offered an appropriate context for the study. Japan was awarded the OPG in 2011, and the country has spent years preparing.

#### 3.2. Experimental Design and Survey Instruments

In practice, social impact can be measured before, during, and after events, or in multiple periods; the choice depends on the unique research purpose. The purpose of this paper is to test the wording of survey items, thus, the research design only necessitates we hold the time period constant. We did so by taking advantage of the OPG which were in the pre-event period as we began the study. In terms of experimental design, Fredline and colleagues [22] also held the time period constant and then took an approach in which they first asked respondents whether they agreed that a certain impact occurred.

Subsequently, they asked respondents how they: (1) perceived this impact on the overall community (OR); and (2) experienced the impact on their personal quality of life (SR). Our approach is slightly different in that respondents were randomly classified into two groups: Group A responded to the SR-SIS scale only while Group B responded to the OR-SIS first, followed by the SR-SIS. In both surveys, basic demographic information was collected, as well as some questions related to the respondents' affinity with the Tokyo 2020 OPG (e.g., support and interest). Each of the two surveys was written in English and translated to Japanese; translation validity was checked by two native speakers. Given that the survey took place 28 months before the originally scheduled date of the OPG, the wording in both surveys was written in the future tense to capture anticipated pre-event social impact.

### 3.3. Measurements

Social impact was measured using a scale developed in the context of the 2016 Rio OPG [57]. It contains 23 items, representing seven constructs described in the literature review: social cohesion (SCOH, 4 items), community spirit and feel-good factor (FGF, 3 items), social capital (SC, 4 items), community involvement with regard to the event (CI, 3 items), disorder and conflict (DC, 3 items), feelings of (un)safety (FUS, 3 items), and sport participation and physical activity (SPA, 3 items). In this scale, all items were worded in terms of "I" and "me", representing the SR-SIS. To create the OR-SIS all items were re-worded in terms of global others (see Table 2). Items from both scales were measured on a 7-point Likert scale (1, Strongly disagree to 7, Strongly agree; see also Table 2).

Demographic variables included sex, age, marital status (2 categories), occupation (8 categories), and personal annual income (6 categories). The survey also inquired how long the participants had lived in the city (number of years).

Affinity with the OPG was also measured because previous research has demonstrated that affinity for, or involvement with an event affects social impact [37]. It is included to verify sample bias. Affinity was measured by asking two questions: "I support the Tokyo 2020 Olympic and Paralympic Games as a resident" and "Tokyo should bid for other major sporting events", both measured on a 7-point Likert scale (1 = Strongly disagree to 7 = Strongly agree) [58]. Three additional questions asked about their interest with the OPG, including, "How frequently do you think about Tokyo 2020 OPG (1 = Not at all to 7 = Very frequently)", "How interested are you in Tokyo 2020 OPG" (1 = Not at all to 7 = Very interested), and "How important is it for you to be informed (have knowledge) about Tokyo 2020?" (1 = Not at all to 7 = Very important) [59].

**Table 2.** Descriptive Statistics of Each Scale (Group B).

Constructs and Items	OR-SIS			SR-SIS			AVE	CR
	M	SD	$\beta$	M	SD	$\beta$		
<b>Social Cohesion</b>							0.65–0.72	0.88–0.91
1. The event will strengthen (people's/my) friendships/relationships in the community	3.40	1.47	0.79	3.05	1.55	0.86		
2. The event will create (people's/my) new friendships/relationships in the community	3.52	1.45	0.81	2.96	1.55	0.82		
3. The event will make (people/me) feel strongly connected to one another	3.76	1.49	0.81	3.15	1.53	0.84		
4. The event will strengthen (people's/my) sense of belonging in the community	3.69	1.46	0.79	3.20	1.61	0.86		
<b>Community Spirit</b>							0.68–0.75	0.87–0.90
1. The event will increase (people's/my) feelings of pride because Tokyo is hosting an event	4.17	1.67	0.85	4.05	1.82	0.88		
2. The event will increase (people's/my) feelings of happiness because Tokyo is hosting the event	3.97	1.55	0.85	3.94	1.76	0.88		
3. The event will lift the (my) spirits (of the population)	4.69	1.61	0.78	4.08	1.81	0.84		
<b>Social Capital</b>							0.66–0.70	0.88–0.90
1. The event will inspire people (me) to become more engaged in the community	3.77	1.48	0.81	3.20	1.58	0.86		
2. The event will enhance (people's/my) feelings of trust in the community	3.74	1.45	0.83	3.21	1.52	0.83		
3. The event will inspire people (me) to more regularly attend community events	3.71	1.41	0.80	3.12	1.58	0.84		
4. The event will increase (people's/my) social interactions in the community	3.83	1.48	0.79	3.09	1.54	0.82		
<b>Community Involvement with Regard to the Event</b>							0.47–0.64	0.72–0.84
1. People (I) will be able to express their (my) opinion about the organization of the event	3.13	1.52	0.68	2.76	1.58	0.72		
2. People (I) will discuss the organization of the event with other people in the community	3.34	1.46	0.70	2.69	1.48	0.88		
3. People (I) will have conversations about the organization of the event	3.11	1.56	0.67	2.63	1.51	0.80		
<b>Sport Participation and Physical Activity</b>							0.71–0.74	0.88–0.89
1. The event will inspire people (me) to become more involved in sport and/or physical activity	4.14	1.50	0.78	3.37	1.67	0.83		
2. The event will spark (people's/my) interest in becoming more involved in sport and/or physical activity	4.37	1.57	0.88	3.76	1.77	0.87		
3. The event will increased (people's/my) interest in sport and/or physical activity	4.39	1.54	0.86	3.71	1.74	0.87		
<b>Disorder and Conflict</b>							0.42–0.46	0.68–0.72
1. The event will disturb (people's/my) daily life in terms of peace and tranquility	4.21	1.60	0.71	3.91	1.71	0.76		
2. (I) People will refrain from going to the city because it will be/is too crowded because of the event	4.26	1.51	0.64	4.28	1.74	0.60		
3. (I will experience) The event will cause traffic jams	5.29	1.39	0.60	5.00	1.55	0.66		
<b>Feelings of (Un)Safety</b>							0.60–0.73	0.82–0.89
1. (I) People will feel unsafe because of potential terrorist attacks due to the event	4.87	1.39	0.76	4.68	1.54	0.88		
2. (I) People will feel afraid that the event attracts terrorists	4.79	1.45	0.83	4.71	1.55	0.86		
3. (I am) People will be concerned about the increased levels of security due to the event	4.67	1.44	0.72	4.67	1.55	0.82		

Note.  $\chi^2/df = 2.42$  (898),  $p < 0.001$ , CFI = 0.933, TLI = 0.923, RMSEA = 0.053, SRMR = 0.046.

### 3.4. Study Participants and Data Collection

Data were collected through an Internet-based survey conducted by a Japanese research company in February 2018. The commercial company was paid to collect the data and operated under their own ethical code (<https://monitor.macromill.com/policy/privacy.html>). This study was exempt from ethics approval because it was considered use of secondary data by the University of Ottawa's Office of Research Ethics and Integrity. Data were transmitted to the researchers without identifiers from the respondents. Stratified sampling based on demographic variables (gender and age groups) from the Population Census of Tokyo was performed to establish a representative view of the 1030 participants (successful response rate: 98.7%). Of the respondents, 49.5% were female; the average age was 42.58 years ( $SD = 14.42$ ); 53.7% were employed, 63.9% earned more than 2,000,000 yen; and the average length of time living in Tokyo was 16.64 years ( $SD = 15.46$ ). Among respondents, 38.2% expressed an intention to volunteer for the event. Their support for the Tokyo 2020 OPG was mediocre ( $M = 3.87$ ;  $SD = 1.53$ , on a 7-point Likert), although their interest in the Tokyo 2020 OPG scored slightly higher ( $M = 4.35$ ;  $SD = 1.81$ ). As for the Tokyo census, 50.7% of the population was female, and the average age was 44.76 years (Tokyo Census: <https://www.toukei.metro.tokyo.lg.jp/jsuikai/js-index2.htm>). Although small differences were found in gender (the number of male respondents was slightly higher) and average age (about 2 years lower) comparing the census data with the case description, no significant differences were found between local residents and this sample (see Table 3). These statistically insignificant marginal differences are consistent with previous stratified samples in social impact research [33].

**Table 3.** Descriptive Statistics of the Total Sample and the Participant Groups A and B.

	Total (N = 1030)	Group A (n = 515)	Group B (n = 515)	<i>p</i>
<b>Demographic Characteristics</b>				
Sex (% women)	49.5	49.9	49.1	ns
Age ( <i>M</i> and <i>SD</i> )	42.58 (14.42)	42.77 (14.44)	42.40 (14.41)	ns
Occupation (% employed full time)	53.7	54.2	56.4	ns
Marital status (% married)	51.7	51.5	52.0	ns
Income (% earning > 2,000,000 yen/year)	63.9	63.8	68.4	ns
Residence ( <i>M</i> number of years, and <i>SD</i> )	16.64 (15.46)	16.36 (15.20)	16.93 (15.74)	ns
<b>Affinity with the OG</b>				
Intention to Volunteer (% yes)	38.2	35.5	40.8	ns
Support for the Games <sup>(1)</sup>	3.87 (1.53)	3.88 (1.53)	3.86 (1.54)	ns
<i>I support the 2020 Olympic and Paralympic Games as a resident</i>	3.89 (1.76)	3.87 (1.75)	3.92 (1.76)	ns
<i>Tokyo should bid for other major sporting events</i>	3.85 (1.62)	3.89 (1.61)	3.81 (1.62)	ns
Interest in Tokyo 2020 <sup>(1)</sup>	3.85 (1.56)	3.85 (1.55)	3.84 (1.56)	ns
<i>How frequently you think about Tokyo 2020</i>	3.53 (1.65)	3.53 (1.62)	3.52 (1.69)	ns
<i>How interested you are in Tokyo 2020</i>	4.35 (1.81)	4.33 (1.85)	4.37 (1.77)	ns
<i>How important knowledge of Tokyo 2020 is in your life</i>	3.66 (1.67)	3.68 (1.64)	3.65 (1.71)	ns

Notes: ns = difference between Groups A and B are not significant; <sup>(1)</sup> *M* and *SD* measured on a 7-point Likert scale. Italic words represent "Items".

### 3.5. Study Design and Data Analysis

The first randomly classified group, Group A ( $n = 515$ ), responded to SR-SIS only; Group B ( $n = 515$ ) responded to the OR-SIS first, followed by the SR-SIS. There was no significant difference between Groups A and B in demographic and event affinity variables, indicating no selection bias (see Table 3). Confirmatory factor analysis was performed for both the SR-SIS and the OR-SIS. Composite Reliability (CR) and Average Variance Extracted (AVE) values were computed for each construct to tests convergent and discriminant analysis. Comparative fit index ( $CFI \geq 0.90$ ), Tucker–Lewis index ( $TLI \geq 0.90$ ), root mean square error of approximation ( $RMSEA \leq 0.08$ ), and standardized root mean square residual ( $SRMR \leq 0.08$ ) were utilized to confirm the goodness-of-fit index criteria [60]. Several comparative analyses were performed using chi-squares and paired independent *t*-tests.

To begin, the samples from Group A and Group B were compared with the Tokyo city census to test for sample bias. Then, the samples from Group A and Group B were compared to eliminate any sample bias between the two sets of respondents (see above). Next, the differences between OR and SR social impact were tested in two ways: first by comparing the SR-SIS from Group A (SR-SIS-A) with the OR-SIS from Group B (OR-SIS-B) using an independent *t*-test between surveys; second, the Group B SR scale (SR-SIS-B) was compared with Group B OR scale (OR-SIS-B) through a paired sample *t*-test. Finally, the SR-SIS from Group A (SR-SIS-A) was compared with the SR-SIS from Group B (SR-SIS-B) to test for response bias and to report social impacts of the Tokyo 2020 OPG.

## 4. Results

### 4.1. Results of the CFA

The results of the global fit indexes, which assessed the proposed model's fit with the data ( $\chi^2/df = 2.42$  (898),  $p < 0.001$ , CFI = 0.933, TLI = 0.923, RMSEA = 0.053, SRMR = 0.046) showed that the measurement models fit the data (see Table 2). Moreover, the computed CR and AVE values for the 14 constructs (seven OR and seven SR factors) ranged from 0.68 to 0.91 for CR and from 0.42 to 0.75 for AVE, indicating reliability and convergent validity with some limitations (i.e., the recommended level of 0.70 for CR and 0.50 for AVE; [61]). Regarding discriminant validity, though most of the squared correlations did not exceed the AVE values, indicating discriminant validity in every construct, five pairs of factors (between OR-SIS/SCOH—OR-SIS/SC; OR-SIS/FGF—SR-SIS/FGF; OR-SIS/DC—SR-SIS/DC; OR-SIS/FUS—SR-SIS/FUS; and SR-SIS/SCOH—SR-SIS/SC) exceeded the AVE values (see Table 4). Thus, we compared the chi-square value of a measurement model with the correlation constrained to equal one to a baseline model without this constraint [62]. We performed a chi-square difference test for those five pairs of factors (a total of five tests in all), and every case resulted in a significant difference, suggesting that all the measures of constructs in the measurement model achieve discriminant validity. Thus, the validity and reliability of this scale were acceptable.

### 4.2. SR versus OR Social Impact

Independent *t*-tests to examine significant differences between the SR items of Group A (SR-SIS-A) and the OR items of Group B (OR-SIS-B; Table 5) showed that all scores on the OR factors are significantly higher than the scores for the SR factors, except for feelings of (un)safety. The same is true for differences in OR and SR for the same participants in Group B, as demonstrated by the paired *t*-test results in Table 5.

In terms of specific results, the ranking of factors according to OR and SR scores for the participants in Group B were similar. In the OR-SIS, four of the seven factors scored above the 4-point indifference level, compared to three factors in the SR-SIS. In both the OR-SIS and SR-SIS, the concepts pertaining to negative social impact factors were over the indifference threshold, namely feelings of (un)safety, and disorder and conflict. The two other factors in the OR-SIS which scored higher than 4 were sport participation and physical activity ( $M = 4.30$ ;  $SD = 1.38$ ) and community spirit ( $M = 4.28$ ;  $SD = 1.43$ ) which indicate some level of positive OR social impact perceptions. In the SR-SIS, only one positive social impact, community spirit, reached a 4-point score ( $M = 4.00$ ;  $SD = 1.66$ ).

**Table 4.** Discriminant Validity (Average Variance Extracted Value for Each Construct with the Squared Correlations between the Respective Constructs).

	OR-SIS/ SCOH	OR-SIS/ FGF	OR-SIS/ SC	OR-SIS/ CI	OR-SIS/ SPA	OR-SIS/ DC	OR-SIS/ FUS	SR-SIS/ SCOH	SR-SIS/ FGF	SR-SIS/ SC	SR-SIS/ CI	SR-SIS/ SPA	SR-SIS/ DC	SR-SIS/ FUS
OR-SIS/ SCOH	0.65													
OR-SIS/ FGF	0.57	0.68												
OR-SIS/ SC	0.74	0.61	0.66											
OR-SIS/ SPA	0.45	0.29	0.41	0.47										
OR-SIS/ CI	0.50	0.59	0.56	0.27	0.71									
OR-SIS/ DC	0.07	0.10	0.06	0.05	0.05	0.42								
OR-SIS/ FUS	0.01	0.00	0.01	0.01	0.00	0.39	0.60							
SR-SIS/ SCOH	0.61	0.37	0.55	0.39	0.31	0.05	0.01	0.72						
SR-SIS/ FGF	0.55	0.72	0.57	0.29	0.52	0.12	0.01	0.48	0.75					
SR-SIS/ SC	0.54	0.36	0.55	0.35	0.31	0.06	0.02	0.81	0.48	0.70				
SR-SIS/ CI	0.32	0.16	0.27	0.46	0.14	0.03	0.01	0.51	0.23	0.49	0.64			
SR-SIS/ SPA	0.49	0.40	0.50	0.26	0.50	0.08	0.01	0.55	0.52	0.56	0.32	0.74		
SR-SIS/ DC	0.11	0.16	0.10	0.08	0.07	0.54	0.28	0.05	0.13	0.06	0.02	0.06	0.46	
SR-SIS/ FUS	0.01	0.01	0.01	0.01	0.00	0.32	0.63	0.00	0.01	0.01	0.00	0.00	0.43	0.73

Note: OR-SIS = other-referenced social impact scale; SR-SIS = self-referenced social impact scale; SCOH = social cohesion; FGF = community spirit; SC = social capital; CI = community involvement; SPA = sport participation and physical activity; DC = disorder and conflict; FUS = feelings of (un)safety.

**Table 5.** Comparison between the Self-Referenced Social Impact Scale (SR-SIS) and the Other-Referenced Social Impact Scale (OR-SIS), and Response Bias.

Factors	Self-Referenced vs. Other-Referenced									Response Bias		
	SR-SIS-A vs. OR-SIS-B			OR-SIS-B vs. SR-SIS-B						SR-SIS-A vs. SR-SIS-B		
				Independent <i>t</i> -Tests			Paired <i>t</i> -Tests			Independent <i>t</i> -Tests		
	SR-SIS-A	OR-SIS-B	SR-SIS-B	<i>t</i> -Value	<i>r</i>	<i>p</i>	<i>t</i> -Value	<i>r</i>	<i>p</i>	<i>t</i> -Value	<i>r</i>	<i>p</i>
<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )	<i>M</i> ( <i>SD</i> )										
Social cohesion	3.06 (1.34)	3.59 (1.26)	3.11 (1.40)	6.56	0.2	0.000 ***	12.15	0.47	0.000 ***	0.61	0	0.543
Community spirit	4.08 (1.65)	4.28 (1.43)	4.00 (1.66)	2.1	0.1	0.036 *	7.94	0.33	0.000 ***	1.03	0	0.306
Social capital	3.19 (1.39)	3.76 (1.25)	3.12 (1.36)	6.93	0.2	0.000 ***	15.58	0.57	0.000 ***	0.87	0	0.386
Community involvement with regard to the event	2.65 (1.22)	3.20 (1.22)	2.74 (1.36)	7.2	0.2	0.000 ***	9.93	0.4	0.000 ***	1.09	0	0.274
Sport participation	3.59 (1.56)	4.30 (1.38)	3.63 (1.58)	7.69	0.2	0.000 ***	13.22	0.5	0.000 ***	0.38	0	0.701
Disorder and conflict	4.29 (1.28)	4.59 (1.18)	4.51 (1.33)	3.85	0.1	0.000 ***	1.98	0.09	0.048 *	2.63	0	0.009 **
Feelings of (un)safety	4.63 (1.37)	4.78 (1.22)	4.74 (1.36)	1.86	0.1	0.063	0.91	0.04	0.361	1.33	0	0.183

Note. \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ .

The four remaining positive social impact factors scored below the 4-point indifference level. For instance, in the SR-SIS, participants disagreed that the event will positively impact their personal sport participation or physical activity levels ( $M_{GroupA} = 3.59$ ;  $SD = 1.56$ ;  $M_{GroupB} = 3.63$ ;  $SD = 1.58$ ). Social capital, social cohesion, and community involvement scored low in both the SR-SIS and OR-SIS, but significantly lower in the SR-SIS. Thus, participants disagree that the event will positively impact social cohesion, social capital, and/or community involvement with regard to the event if this is measured through OR measures, but disagree even more when measured through SR measures. Overall, the results confirm that in the period roughly two years before the OPG, OR measures of perceived future social impact are consistently higher than SR measures of perceived future social impact, except for feelings of (un)safety.

#### 4.3. Social Impact of Tokyo 2020 OPG Based on SR Measures

The SR-SIS measured the pre-event social impact of the Tokyo 2020 OPG in two different sample groups. No significant differences were found in SR items between Group A (SR-SIS-A) and Group B (SR-SIS-B), except for disorder and conflict (see Table 5). However, with an effect size of 0.8, this difference is unimportant. Thus, overall, the response bias between Group A and Group B is negligible.

In terms of specific results, amongst the conceptual categories, feelings of (un)safety scored highest, followed by disorder and conflict. Both reflect negative social impact factors. Community spirit is ranked third and was the only other factor that scored around the 4-point indifference threshold. All other social impact factors scored below the indifferent threshold, from a 3.06 ( $SD = 1.34$ ) for social cohesion to 3.63 ( $SD = 1.58$ ) for sport participation and physical activity, showing little evidence for positive social impact of the Tokyo 2020 OPG based on projected SR social impact.

## 5. Discussion

### 5.1. The SR-SIS versus OR-SIS

Scale testing confirmed that the 23 items represented seven predetermined social impact factors in both the SR-SIS and the OR-SIS, as was the case for the original social impact scale [57]. The comparison between the SR-SIS and OR-SIS allowed us to replicate past work on social impact from tourism [22] in the context of sport events. Our results confirm the theoretical prediction that OR items are higher, both in positive and negative ways. Specifically, except for feelings of (un)safety, all other OR measures of collective social impacts are significantly higher than SR measures, whether this was measured between groups (A and B), or within the same group of respondents (Group B). The order in which the scales were provided for group B (i.e., OR first, SR second) cannot explain the differences, as the SR social impacts reported by group A and B were not significantly different. Thus, changing the point of reference, in terms of wording, matters when measuring social impacts of sport events.

The higher OR scores indicate that respondents consistently believed that others were more likely to experience certain potential social impacts than if asked about themselves, confirming an internal bias of social projection [23,53]. We can find no evidence or theory to explain why the SR measures would be underestimates of reality and therefore conclude that OR measures are overestimates because projection bias leads to inaccuracy in measurement. We therefore posit that the point of reference as the circumscribed self (SR) calibrates and provides a more accurate measurement of social impacts than providing a perspective of a global other (OR). As a result, based on the experimental design, the SR-SIS is an appropriate scale to use in the pre-event period and can be generalized to other one-off, mega-events similar to the OPG.

SR scores are also more accurate because they involve a more appropriate scientific survey methodology (asking each person in the sample and generalizing to the population). Each person matters. People who see zero social benefits are given a voice. They can respond zero and be averaged in with those who see medium and high social impacts. On the other hand, in OR questions, these people

who do not see a personal benefit, are asked if others see a benefit. Of course, if they are paying attention to media, they will answer yes, that there are others who see a benefit. Indeed, external factors such as media framing and public discourse can be powerful tools to shape perceptions [49,63,64], and possibly affect pre-event expectations of OR social impacts more than SR. While OR scales are upwardly biased, they do grasp public opinion in more general terms, reflecting a general mood that lingers among a population which may influence event organizers and policy makers. For instance, high OR social impact scores may be useful for sport event organizers and policy makers to gain support for hosting, but the reverse can also happen. Specifically, for the Tokyo OPG, residents have been strongly influenced by the media regarding COVID-19, negatively impacting their attitudes towards the Games. A poll indicated that 51.7% of Tokyo residents no longer wanted to see Tokyo host the OPG in 2021 [65]. While these attitudes and public opinions may matter to gain support for hosting or not, they remain generic and affected by projection bias.

### 5.2. Practical Results for the Tokyo 2020 OPG

In terms of absolute scale values we found little evidence for perceived potential positive social impact whether measured as SR or OR items, and negative social impacts prevailed. Thus, positive social impacts measured two years prior to the originally scheduled Tokyo OPG are negligible, whether measured based on OR or SR. The only positive social impact factor which hovered just above (in the OR-SIS) or around (in the SR-SIS) the 4-point indifference threshold is community spirit. This refers to the short-lived feel-good factor which residents and the community at large experience in the context of events, confirming previous studies [37,44,45]. Except for sport participation in the OR-SIS, the remaining four positive social impact factors score below the 4-point indifference level, and significantly lower in the SR-SIS compared to the OR-SIS. In summary, two years prior to the planned occurrence of the Tokyo 2020 OPG, residents did not anticipate that the event would contribute positively to social impact, whether measured through perceptions of others, and even less so when measured through perceptions of self. This may be due to the absence of recency [66]; data were collected two years prior to the event. In previous studies where social impacts of events were prevalent, results showed the highest levels during the event [37,44,45].

The two negative social impact factors, disorder and conflict and feelings of (un)safety, showed the highest scores pre-event, regardless of how they were measured. This result shows that negative social impact perceptions and experiences prevail over positive social impacts (see also [13,14]), at least when measured two years prior to the event. Residents have no control over the occurrence of the negative social impact factors. However, people can react to aspects of conflict and disorder, and decide for themselves if they want to circumvent the inconveniences created by the event, such as deciding not to go to the city because of traffic jams. Thus, it is up to the individual how they deal with or act upon the situation. Overall, there may be an expectation that people will try to avoid the potential inconveniences leading to higher scores for both SR and OR measures of disorder and conflict. In contrast, there is little residents can do against feelings of unsafety. For example, if terrorist attacks occur, anybody can be affected, and there is very little individuals can do to avoid this threat. This may explain why we do not find a significant difference between pre-event OR and SR measures for feelings of (un)safety; if it occurs everybody will experience it, as is the case with COVID-19. The data for this study were collected pre-COVID-19 and we recognize that if repeated during COVID, the results could be very different.

### 5.3. Limitations and Future Research

The literature has demonstrated that pre-event OR scores are generally higher than post-event scores [13–15,17–19], indicating a potential problem with using OR questions before an event. However, in keeping time constant in the research design, both OR and SR data were collected pre-event when social impacts were not yet experienced and respondents could only anticipate how they believed what they would experience the events and how they would react. It can be argued that asking residents to

evaluate the impacts of an event based on their projected self-experience is self-reflective and subjective in nature [22]. Nevertheless, estimates of pre-event social impact are a regular occurrence in both research and practice. Whether or not we agree with the practice of predicting a possible future social impact, we know it occurs, and this research offers a solution to make the data collection as precise as possible.

Clearly, more research is warranted to understand the longitudinal and sustainable nature of experienced social impacts. Assuming manifestations of social impact experiences do occur during the event, future research will have to determine whether the social impact gap between OR and SR items remains, increases, or diminishes when data are collected during the event, and even post-event. Thus, to provide deeper insight and to further test our argument, we recommend repeating the same study design using both the SR-SIS and OR-SIS in future research during and post-event. This will allow to test the actual accuracy of SR scores and find support if the SR-SIS is indeed a valid scale in other event time periods. This will also provide further insights in the wide variation of higher, lower, or the same scores for SR items, found in the literature pre- and post-event [33,35,37,45]. Thus, in order to test the robustness of the scales and the findings, the data collection should be repeated during and after the Tokyo 2020 OPG to test the recency effect and take into account that media attention [49,64] and political discourse [63], increasing as the event gets closer. We also recognize that COVID-19 has drastically changed the context for the Tokyo OPG [65], which may hinder a longitudinal comparison pre-, during, and post-event.

Although Robbins and Krueger [23] reported no differences in the ordering of scales, as in SR-SIS versus OR-SIS in Group B, ideally, to provide a clearer distinction between OR and SR outcomes, data would have been collected from four groups: a group responding to the OR-SIS only; a group responding to the SR-SIS only; a group responding to the OR-SIS first, followed by the SR-SIS; and finally, a group responding to the SR-SIS followed by the OR-SIS. This expanded experimental design could have confirmed the findings of Robbins and Krueger [23] and would have controlled for any potential bias of first responding to OR items followed by SR items and vice versa. The fact that no significant differences appeared between the SR-SISs of Group A and Group B, except for disorder and conflict, alleviated some of these concerns.

## 6. Conclusions

Past research on social impact of events relies heavily on asking respondents about a generic and global “other” which we referred to as other-referenced (OR) measures of event impact on the community at large (e.g., “The event strengthens friendships/relationships in the community”; [13,14]). To obtain more accurate measures of social impact, the solution is to use word scale items in terms of self-referenced (SR) social impacts, using the first person using “me”, “I”, or “my” (e.g., “The event strengthens my friendships/relationships in the community”). Two years prior to the planned Tokyo 2020 OPG, we found that anticipated social impact measures of others (OR) were generally significantly higher than projected social impact measures of self (SR). Projection bias, and to some extent media framing and socio-political discourse, explain why OR measures are skewed upwards, and therefore overestimate social impacts. Moreover, methodologically, by asking each person in the sample about their expected personal experiences, generalizations to the population are a more truthful reflection of reality. We therefore posit that the point of reference as the circumscribed self (SR) calibrates and provides a more accurate measurement of social impacts when used to measure mega-events in the pre-event period. Using the SR-SIS moving forward and testing it in other time periods will generate deeper insights and strengthen the generalizability of the instrument.

Two years prior to the anticipated hosting of the Tokyo 2020 OPG, there is little evidence for expected positive social impacts, except for a short-lived anticipated enhanced community spirit (i.e., feel-good factor) and perceived negative social impacts prevail. The negligible positive social impacts can be partially explained by recency bias [66], as stronger effects may become apparent closer to the event.

The present study asserts that measuring pre-event social impact through expected individual experiences using a self-referenced social impact scale provides a more accurate assessment of possible social impact than using an other-referenced social impact scale. A self-referenced social impact scale helps proponents of events inform host communities more realistically about how events contribute socially (or not). Overestimating social impact claims through measures based on perceptions of others may raise residents' expectations; if not delivered, this could negatively impact perceived benefits. According to social exchange theory, residents form negative attitudes toward the object (i.e., the event) when the perceived benefits are lower than the perceived costs (e.g., [13]). Thus, accurate assessment is practically important for event organizers to host sustainable sport events.

**Author Contributions:** Conceptualization, M.T., D.O. and N.A.; Methodology, M.T. and D.O.; Software, D.O.; Validation, D.O. and M.T.; Formal Analysis, D.O. and M.T.; Data collection, D.O.; Resources, M.T., D.O. and N.A.; Data Curation, D.O.; Writing—Original Draft Preparation, M.T. and D.O.; Writing—Review & Editing, N.A., M.T. and D.O.; Visualization, M.T.; Project Administration, D.O. and M.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Sasakawa Sports Research Grant from Sasakawa Sports Foundation [grant numbers 170A1-023, 2017-04-01-2018-03-31].

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Arrow, K.J.; Debreu, G. Existence of an equilibrium for a competitive economy. *Econometrica* **1954**, *22*, 265–290. [[CrossRef](#)]
2. Preuss, H.; Andreff, W.; Weitzmann, M. *Cost and Revenue Overruns of the Olympic Games 2000–2018*; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2019; ISBN 978-3-658-24995-3.
3. Lee, J.W. A thin line between a sport mega-event and a mega-construction project: The 2018 Winter Olympic Games in PyeongChang and its event-led development. *Manag. Sport Leis.* **2020**, 1–18. [[CrossRef](#)]
4. Baade, R.A.; Matheson, V.A. Going for the gold: The economics of the Olympics. *J. Econ. Perspect.* **2016**, *30*, 201–218. [[CrossRef](#)]
5. Coates, D.; Humphreys, B.R. Do economists reach a conclusion on subsidies for sports franchises, stadiums, and mega-events? *Econ J. Watch* **2008**, *5*, 294–315.
6. Humphreys, B.R.; Prokopowicz, S. Assessing the impact of sports mega-events in transition economies: EURO 2012 in Poland and Ukraine. *Int. J. Sport Manag. Mark.* **2007**, *2*, 496. [[CrossRef](#)]
7. Kesenne, S. The economic impact, costs and benefits of the FIFA World Cup and the Olympic Games: Who wins, who loses? In *International Handbook on the Economics of Mega Sporting Events*; Edward Elgar Publishing: Cheltenham, UK, 2012; pp. 270–278.
8. Heere, B.; Wear, H.; Jones, A.; Breitbarth, T.; Xing, X.; Paramio Salcines, J.L.; Yoshida, M.; Derom, I. Inducing destination images among international audiences: The differing effects of promoting sport events on the destination image of a city around the world. *J. Sport Manag.* **2019**, *33*, 506–517. [[CrossRef](#)]
9. Lauermaun, J. Visualising sustainability at the Olympics. *Urban Stud.* **2020**, *57*, 2339–2356. [[CrossRef](#)]
10. Chalip, L. Towards social leverage of sport events. *J. Sport Tour.* **2006**, *11*, 109–127. [[CrossRef](#)]
11. Weed, M.; Coren, E.; Fiore, J.; Wellard, I.; Chatziefstathiou, D.; Mansfield, L.; Dowse, S. The Olympic Games and raising sport participation: A systematic review of evidence and an interrogation of policy for a demonstration effect. *Eur. Sport Manag. Q.* **2015**, *15*, 195–226. [[CrossRef](#)]
12. Kellett, P.; Hede, A.-M.; Chalip, L. Social policy for sport events: Leveraging (relationships with) teams from other nations for community benefit. *Eur. Sport Manag. Q.* **2008**, *8*, 101–121. [[CrossRef](#)]
13. Baldock, A.-L.; Maes, M.; Buelens, M. The social impact of the Tour de France: Comparisons of residents' pre- and post-event perceptions. *Eur. Sport Manag. Q.* **2011**, *11*, 91–113. [[CrossRef](#)]
14. Kim, W.; Jun, H.M.; Walker, M.; Drane, D. Evaluating the perceived social impacts of hosting large-scale sport tourism events: Scale development and validation. *Tour. Manag.* **2015**, *48*, 21–32. [[CrossRef](#)]
15. Oshimi, D.; Harada, M.; Fukuhara, T. Residents' perceptions on the social impacts of an international sport event: Applying panel data design and a moderating variable. *J. Conv. Event Tour.* **2016**, *17*, 294–317. [[CrossRef](#)]

16. West, T.V.; Kenny, D.A. The truth and bias model of judgment. *Psychol. Rev.* **2011**, *118*, 357–378. [[CrossRef](#)]
17. Kim, H.J.; Gursoy, D.; Lee, S.-B. The impact of the 2002 World Cup on South Korea: Comparisons of pre- and post-games. *Tour. Manag.* **2006**, *27*, 86–96. [[CrossRef](#)]
18. Kim, S.S.; Petrick, J.F. Residents' perceptions on impacts of the FIFA 2002 World Cup: The case of Seoul as a host city. *Tour. Manag.* **2005**, *26*, 25–38. [[CrossRef](#)]
19. Ma, S.C.; Ma, S.M.; Wu, J.H.; Rotherham, I.D. Host residents' perception changes on major sport events. *Eur. Sport Manag. Q.* **2013**, *13*, 511–536. [[CrossRef](#)]
20. Gursoy, D.; Kendall, K.W. Hosting mega events. *Ann. Tour. Res.* **2006**, *33*, 603–623. [[CrossRef](#)]
21. Kaplanidou, K.; Karadakis, K.; Gibson, H.; Thapa, B.; Walker, M.; Geldenhuys, S.; Coetzee, W. Quality of life, event impacts, and mega-event support among South African residents before and after the 2010 FIFA World Cup. *J. Travel Res.* **2013**, *52*, 631–645. [[CrossRef](#)]
22. Fredline, L.; Deery, M.; Jago, L. *Development of a Scale to Assess the Social Impacts of Tourism within Communities*; CRC for Sustainable Tourism: Gold Coast MC, Australia, 2006; ISBN 978-1-920704-60-5.
23. Robbins, J.M.; Krueger, J.I. Social projection to ingroups and outgroups: A review and meta-analysis. *Personal. Soc. Psychol. Rev.* **2005**, *9*, 32–47. [[CrossRef](#)]
24. Mathieson, A.; Wall, G. *Tourism: Economic, Physical, and Social Impacts*; Longman: London, UK; New York, NY, USA, 1982; ISBN 978-0-582-30061-3.
25. Ritchie, J.B. Assessing the impact of hallmark events: Conceptual and research issues. *J. Travel Res.* **1984**, *23*, 2–11. [[CrossRef](#)]
26. Holmes, K.; Hughes, M.; Mair, J.; Carlsen, J. Chapter 7: Socio-cultural impacts of events. In *Events and Sustainability*; Routledge: New York, NY, USA, 2015; pp. 98–112.
27. Karadakis, K.; Kaplanidou, K. Legacy perceptions among host and non-host Olympic Games residents: A longitudinal study of the 2010 Vancouver Olympic Games. *Eur. Sport Manag. Q.* **2012**, *12*, 243–264. [[CrossRef](#)]
28. Smith, J.A.; Anderson, D.H.; Davenport, M.A.; Leahy, J.E. Community benefits from managed resource areas: An analysis of construct validity. *J. Leis. Res.* **2013**, *45*, 192–213. [[CrossRef](#)]
29. Ajzen, I.; Fishbein, M. *Understanding Attitude and Predicting Social Behavior*; Prentice Hall: Englewood Cliffs, NJ, USA, 1980.
30. Huang, H.; Min, S.D.; Wang, T.R.; Mao, L.L. Social exchange process in collectivistic countries: An examination of sporting events in China. *Eur. Sport Manag. Q.* **2016**, *16*, 172–189. [[CrossRef](#)]
31. Ritchie, B.W.; Shipway, R.; Cleeve, B. Resident perceptions of mega-sporting events: A non-host city perspective of the 2012 London Olympic Games. *J. Sport Tour.* **2009**, *14*, 143–167. [[CrossRef](#)]
32. Lorde, T.; Greenidge, D.; Devonish, D. Local residents' perceptions of the impacts of the ICC Cricket World Cup 2007 on Barbados: Comparisons of pre- and post-games. *Tour. Manag.* **2011**, *32*, 349–356. [[CrossRef](#)]
33. Gibson, H.J.; Walker, M.; Thapa, B.; Kaplanidou, K.; Geldenhuys, S.; Coetzee, W. Psychic income and social capital among host nation residents: A pre-post analysis of the 2010 FIFA World Cup in South Africa. *Tour. Manag.* **2014**, *44*, 113–122. [[CrossRef](#)]
34. Onyx, J.; Bullen, P. Measuring social capital in five communities. *J. Appl. Behav. Sci.* **2000**, *36*, 23–42. [[CrossRef](#)]
35. Heere, B.; Walker, M.; Gibson, H.; Thapa, B.; Geldenhuys, S.; Coetzee, W. The power of sport to unite a nation: The social value of the 2010 FIFA World Cup in South Africa. *Eur. Sport Manag. Q.* **2013**, *13*, 450–471. [[CrossRef](#)]
36. Tajfel, H. *Human Groups and Social Categories: Studies in Social Psychology*; Cambridge University Press: Cambridge, UK; New York, NY, USA, 1981.
37. Oja, B.D.; Wear, H.T.; Clopton, A.W. Major sport events and psychic income: The social anchor effect. *J. Sport Manag.* **2018**, *32*, 257–271. [[CrossRef](#)]
38. Peachey, J.W.; Bruening, J.; Lyras, A.; Cohen, A.; Cunningham, G.B. Examining social capital development among volunteers of a multinational sport-for-development event. *J. Sport Manag.* **2015**, *29*, 27–41. [[CrossRef](#)]
39. Liu, D.; Broom, D.; Wilson, R. Legacy of the Beijing Olympic Games: A non-host city perspective. *Eur. Sport Manag. Q.* **2014**, *14*, 485–502. [[CrossRef](#)]
40. Vargas-Sanchez, A.; Porrás-Bueno, N.A.; Plaza-Mejía, M. Explaining residents' attitudes to tourism: Is a universal model possible? *Ann. Tour. Res.* **2010**, *38*, 460–480. [[CrossRef](#)]

41. Mao, L.L.; Huang, H. Social impact of Formula One Chinese Grand Prix: A comparison of local residents' perceptions based on the intrinsic dimension. *Sport Manag. Rev.* **2016**, *19*, 306–318. [[CrossRef](#)]
42. Kim, W.; Walker, M. Measuring the social impacts associated with Super Bowl XLIII: Preliminary development of a psychic income scale. *Sport Manag. Rev.* **2012**, *15*, 91–108. [[CrossRef](#)]
43. Inoue, Y.; Havard, C.T. Determinants and consequences of the perceived social impact of a sport event. *J. Sport Manag.* **2014**, *28*, 295–310. [[CrossRef](#)]
44. Schlegel, A.; Pfitzner, R.; Koenigstorfer, J. The impact of atmosphere in the city on subjective well-being of Rio de Janeiro residents during (vs. Before) the 2014 FIFA World Cup. *J. Sport Manag.* **2017**, *31*, 605–619. [[CrossRef](#)]
45. Kavetsos, G.; Szymanski, S. National well-being and international sports events. *J. Econ. Psychol.* **2010**, *31*, 158–171. [[CrossRef](#)]
46. Thomson, A.; Cuskelly, G.; Toohey, K.; Kennelly, M.; Burton, P.; Fredline, L. Sport event legacy: A systematic quantitative review of literature. *Sport Manag. Rev.* **2019**, *22*, 295–321. [[CrossRef](#)]
47. Horne, J. The four 'knowns' of sports mega-events. *Leis. Stud.* **2007**, *26*, 81–96. [[CrossRef](#)]
48. Clark, R.; Misener, L. Understanding urban development through a sport events portfolio: A case study of London, Ontario. *J. Sport Manag.* **2015**, *29*, 11–26. [[CrossRef](#)]
49. Sant, S.-L.; Mason, D.S. Framing event legacy in a prospective host city: Managing Vancouver's Olympic bid. *J. Sport Manag.* **2015**, *29*, 42–56. [[CrossRef](#)]
50. Misener, L. A media frames analysis of the legacy discourse for the 2010 Winter Paralympic Games. *Commun. Sport* **2013**, *1*, 342–364. [[CrossRef](#)]
51. Chong, D.; Druckman, J.N. Framing Theory. *Annu. Rev. Polit. Sci.* **2007**, *10*, 103–126. [[CrossRef](#)]
52. Gratton, C.; Henry, I.P. *Sport in the City: The Role of Sport in Economic and Social Regeneration*; Routledge: New York, NY, USA, 2001.
53. Van Boven, L.; Judd, C.M.; Sherman, D.K. Political polarization projection: Social projection of partisan attitude extremity and attitudinal processes. *J. Pers. Soc. Psychol.* **2012**, *103*, 84–100. [[CrossRef](#)]
54. Fredline, L.; Jago, L.; Deery, M. The development of a generic scale to measure the social impacts of events. *Event Manag.* **2003**, *8*, 23–37. [[CrossRef](#)]
55. Downward, P.; Dawson, A.; Dejonghe, T. *Sports Economics: Theory, Evidence and Policy*; Routledge: New York, NY, USA, 2009.
56. Taks, M.; Chalip, L.; Green, B.C. Impacts and strategic outcomes from non-mega sport events for local communities. *Eur. Sport Manag. Q.* **2015**, *15*, 1–6. [[CrossRef](#)]
57. Taks, M.; Rocha, C. Maakte Rio 2016 de Brazilianen gelukkig? [Did Rio 2016 make the Brazilians happy?]. In *The Story of Rio 2016: De Maatschappelijke Betekenis van de Olympische en Paralympische Spelen 2016 [The story of Rio 2016: The Societal Meaning of the Olympic and Paralympic Games 2016]*; Hoover, P., Breedveld, K., Eds.; Mulier Instituut: Utrecht, The Netherlands, 2017; pp. 119–127.
58. Prayag, G.; Hosany, S.; Nunkoo, R.; Alders, T. London residents' support for the 2012 Olympic Games: The mediating effect of overall attitude. *Tour. Manag.* **2013**, *36*, 629–640. [[CrossRef](#)]
59. Hill, B.; Green, B.C. Repeat attendance as a function of involvement, loyalty, and the sportscape across three football contexts. *Sport Manag. Rev.* **2000**, *3*, 145–162. [[CrossRef](#)]
60. Hu, L.T.; Bentler, P.M. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct. Equ. Model.* **1999**, *6*, 1–55. [[CrossRef](#)]
61. Fornell, C.; Larcker, D.F. Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *J. Mark. Res.* **1981**, *18*, 382. [[CrossRef](#)]
62. Algesheimer, R.; Dholakia, U.M.; Herrmann, A. The social influence of brand community: Evidence from European car clubs. *J. Mark.* **2005**, *69*, 19–34. [[CrossRef](#)]
63. Anderson, J.; Taks, M. Urban governance of non-mega sport events: A socio-political discourse analysis. In *Research Handbook on Sports Governance*; Winand, M., Anagnostopoulos, C., Eds.; Edward Elgar Publishing: Cheltenham, UK, 2018; in press.
64. Preuss, H.; Alfs, C. Signaling through the 2008 Beijing Olympics—Using Mega Sport Events to Change the Perception and Image of the Host. *Eur. Sport Manag. Q.* **2011**, *11*, 55–71. [[CrossRef](#)]
65. Bruton, M. Poll: More than half of Tokyo residents don't want city to host 2021 Olympics. 2020. Available online: <https://www.forbes.com/sites/michellebruton/2020/06/30/poll-more-than-half-of-tokyo-residents-dont-want-city-to-host-2021-olympics/?sh=3775fdcd55ea> (accessed on 8 December 2020).

66. Martin, E. *Survey Questionnaire Construction Research Report Series; Survey Methodology #2006-13; Research Report Series*; U. S. Census Bureau: Washington, DC, USA, 2006.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).