European Sport Management Quarterly

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/resm20

Expectation-achievement gaps and satisfaction in World Cup football supporters: a quasi-experiment

Mark A. Robinson, Matthew C. Davis & Kerrie L. Unsworth

To cite this article: Mark A. Robinson, Matthew C. Davis & Kerrie L. Unsworth (2023) Expectation-achievement gaps and satisfaction in World Cup football supporters: a quasi-experiment, European Sport Management Quarterly, 23:5, 1304-1327, DOI: 10.1080/16184742.2021.2007278

To link to this article: https://doi.org/10.1080/16184742.2021.2007278

EUROPEAN SPORT MANAGEMENT QUARTERLY
2023, VOL. 23, NO. 5, 1304–1327
https://doi.org/10.1080/16184742.2021.2007278

Routledge
Taylor & Francis Group

Expectation-achievement gaps and satisfaction in World Cup football supporters: a quasi-experiment

Mark A. Robinson, Matthew C. Davis and Kerrie L. Unsworth

Leeds University Business School, University of Leeds, Leeds, UK

ABSTRACT

Research question: Football is the world's most popular spectator sport so supporters' satisfaction and happiness (wellbeing) is of considerable interest. We examined how gaps between supporters' expectations and teams' achievements affect supporters' wellbeing, and how social identity moderates this. Addressing previous methodological limitations, we examined such gaps objectively in a meaningful real-world scenario.

Research methods: We conducted a quasi-experiment around the 2018 FIFA World Cup with 278 supporters and 63 control participants, collecting data before, during, and after the championship. We also examined curvilinear effects and the moderating effect of social identity using polynomial regression surface modelling.

Results and findings: Supporters' satisfaction increased when their national team's achievements exceeded expectations, decreased when achievements failed to exceed expectations, and remained stable for a control group. Happiness was unchanged in these conditions, however. Polynomial regression indicated that expectations and achievements explained a substantial 35% of incremental variance in satisfaction, and their three-dimensional curvilinear interaction accounted for a further 6% (a considerable 41% overall). The moderating effect of social identity on this relationship also approached significance.

Implications: Increases in achievements increase satisfaction, and disproportionately so for high achievements. However, satisfaction is also increased by lowering expectations, particularly for supporters identifying strongly with their teams. For supporters identifying less strongly, though, moderate expectations increase satisfaction most.

ARTICLE HISTORY
Received 28 January 2021
Accepted 9 November 2021

KEYWORDS
Expectation-achievement gap; expectations; satisfaction; happiness; prospect theory

Football is the world's most popular spectator sport (Palacios-Huerta, 2004), with a television audience of 4.7 billion (Eurosport, 2011) and 211 national teams affiliated to the global governing body Fédération Internationale de Football Association (FIFA, 2018). It is therefore of interest to sport management professionals, for two reasons. First, financially, as the combined revenue of the ten leading football clubs exceeds $7 billion

CONTACT Mark A. Robinson m.robinson@lubs.leeds.ac.uk Leeds University Business School, University of Leeds, Leeds LS2 9JT, UK

Supplemental material for this article can be accessed at https://doi.org/10.1080/16184742.2021.2007278.

annually (Deloitte, 2020). Second, psychologically, as football teams have a considerable impact on their supporters (Park et al., 2009), which influences both spending (Biscaia et al., 2012; Lee & Kang, 2015) and wellbeing (Mutz, 2019). We therefore examine this, specifically how expectations and achievements of national football teams in the 2018 FIFA World Cup (FWC2018) influence supporters' satisfaction and happiness.

Given football's popularity, it greatly influences many supporters' lives. During host France's championship-winning campaign in FWC1998, suicide rates in France decreased by 10% overall and by 20% on days following their games (Encrenaz et al., 2012). Neurologically, fMRI scans indicate activity in the brain's emotional pleasure centres of supporters watching goals being scored (McLean et al., 2009) and contrasting neural activity when watching their team win or lose (Park et al., 2009). Furthermore, Spanish supporters' increased positive emotions following Spain's victory in FWC2010 endured beyond the championship itself (Jones et al., 2012). Evidently, then, FWC2018 is of major personal significance and represents an ideal context for examining psychological wellbeing.

Regarding expectations and achievements, surveys of Portuguese supporters found joy predicted satisfaction, which then predicted consumer intentions, but expectation fulfilment was only measured within satisfaction (Biscaia et al., 2012). Analysis of Spanish football results suggested gaps between teams' expected and actual performances could measure satisfaction, but expectations and satisfaction were inferred rather than measured (González-Gómez & Picazo-Tadeo, 2010). Thus, there are hints but little definitive evidence about how expectations and achievements affect football supporters' wellbeing.

Expectations, achievements, and psychological wellbeing

Satisfaction is a composite state of attitudinal, cognitive, and affective favourability (Weiss, 2002) towards either: (a) a product or service (Hult et al., 2019), or (b) a person's own life generally, known as 'life satisfaction' (Pavot & Diener, 2008). Happiness also comprises attitudinal, cognitive, and affective components, alongside perceived purpose and agency for one's life (Hills & Argyle, 2002). Both (life) satisfaction and happiness are facets of wellbeing (Pavot & Diener, 2008). Given football's importance in supporters' lives, our research focuses on supporters' satisfaction with their football team's performance and their happiness, which we collectively call wellbeing.

Research has investigated the relationship between achievements and psychological wellbeing (Bücker et al., 2018), with salary often representing achievement. Two alternative findings have emerged, yielding different explanations. Early research supported an *absolute* explanation, with higher salary levels linked to higher psychological wellbeing (Diener et al., 1993; Veenhoven, 1991). More recently, Killingsworth (2021) found life satisfaction and experienced wellbeing both increased linearly from very low ($15,000) to very high ($480,000) annual household incomes.

Other research supports a *relative* explanation, though, finding satisfaction is determined more by salary rank within immediate reference groups rather than by absolute salary (Boyce et al., 2010; D'Ambrosio & Frick, 2007). This approach predominantly uses expectation disconfirmation theory (Lankton et al., 2014; Oliver, 1980), finding that satisfaction increases when achievements exceed expectations and decreases when achievements fail to exceed expectations, across diverse domains (Derksen et al., 2018;

Green et al., 2018; Vannier & O'Sullivan, 2017). Similarly, experiments by Shepperd and McNulty (2002) tested decision affect theory (Mellers et al., 1997) finding participants' wellbeing (happiness and positive emotions) was highest when financial and health outcomes exceeded expectations and lowest when outcomes were lower than expected. This supports both decision affect theory and expectation disconfirmation theory, which both suggest that it is therefore the change from prior expectations to subsequent achievements —the expectation-achievement gap (Li & Fung, 2012)—which influences wellbeing.

Evidently, both relative and absolute achievements are important predictors of wellbeing. However, it remains unclear which explanation is most powerful due to the need for research with both internal and external validity. Experiments can measure expectations and achievements objectively, and often have control groups, but require artificial laboratory scenarios (e.g. McBride, 2010). Conversely, studies in meaningful real-world scenarios are usually correlational and experience difficulty measuring expectations and achievements directly and objectively, often reliant on self-reports (e.g. Ruvolo & Veroff, 1997).

Our research addresses both limitations by using a rigorous quasi-experiment to examine expectation-achievement gaps in FWC2018 supporters. This constitutes an ideal domain in which to study expectation-achievement gaps and untangle the relative and absolute nature of wellbeing, for these methodological reasons. First, FWC2018 is a highly meaningful real-world event for supporters. Second, we can use FWC2018 as an effective experimental intervention. Third, unlike other meaningful events, expectations and achievements in sports like football are simpler to conceptualise and measure objectively. Furthermore, given football's popularity, it is an intrinsically interesting research context, as well as being methodologically advantageous.

While a few previous quasi-experiments have been conducted around sporting championships, there are differences. For instance, Dolan et al. (2016) examined life satisfaction, happiness, and anxiety one year before, during, and one year after the 2012 Olympic Games in residents of host city London, with residents of Paris and Berlin as control groups. Similarly, Mutz (2019) also examined life satisfaction one month before, during, and two months after the 2016 European Football Championship. While both studies found short-term increases in life satisfaction during the events, they focused on participants' experiences of the championships generally, hence the long data collection intervals.

While similar, our quasi-experiment focuses on participants' satisfaction and happiness specifically in response to their team's performance—expected and achieved—and will therefore measure wellbeing immediately before and after FWC2018. Nevertheless, by using a similarly rigorous quasi-experiment, in this methodologically-optimal research context, we can examine expectation-achievement gaps and the size of their true effect on wellbeing accurately, both uniquely and in combination. We first examine satisfaction specifically, as this has been the primary focus of expectation disconfirmation theory (Lankton et al., 2014; Oliver, 1980), our main theoretical framework here. Accordingly, our first hypotheses (H) state:

H1: Changes from expectations to achievements will be positively related to satisfaction.

H2: Relative changes from expectations to achievements will be: (a) more highly positively related to satisfaction than absolute achievements, and (b) perceived as distinct from absolute achievements in relation to satisfaction.

Most previous studies of expectation-achievement gaps have focused on satisfaction, in specific domains without wider psychological significance for participants (e.g. von Meyer-Höfer et al., 2015). However, with FWC2018 evidently being so central to supporters' lives, the performance of their national football team would likely have broader implications for their psychological wellbeing. Indeed, Kavetsos and Szymanski (2010) investigated the relationship between mega-events and people's life satisfaction—equivalent to happiness—within biannual European surveys and linked this to their nation's performances and hosting of the Olympic Games, and football's World Cup and European Championships. Results indicated that winning more Olympic medals than before and hosting an international football tournament led to a temporary 'feelgood' increase in life satisfaction in those countries. Accordingly, we explore this here by also examining happiness as a broader wellbeing outcome of expectation-achievement gaps, with the equivalent hypotheses as satisfaction:

H3: Changes from expectations to achievements will be positively related to happiness.

H4: Relative changes from expectations to achievements will be: (a) more highly positively related to happiness than absolute achievements, and (b) perceived as distinct from absolute achievements in relation to happiness.

Note that in Hypotheses 2 and 4, we predict that relative achievements will influence wellbeing more than absolute achievements, as most literature reviewed earlier supports this (e.g. expectation disconfirmation theory), and hypotheses must specify a testable difference or relationship (i.e. non-null). However, as discussed earlier, we recognise the ongoing debate about absolute and relative explanations and test this here.

Curvilinear changes in psychological wellbeing

Recent psychological wellbeing research has started integrating expectancy-achievement theories with prospect theory (Kahneman & Tversky, 1979), suggesting curvilinear effects. Prospect theory suggests that people generally demonstrate 'loss aversion', where their value judgements of forecasted outcomes are much more sensitive to losses than to gains, with a resultant S-shaped relationship (see Barberis, 2013, for a review). Research examining salaries has also found loss aversion effects, where people's wellbeing decreases more due to a salary decrease than it increases relative to an equal-sized salary increase (Boyce et al., 2013, 2016). These studies only examined linear effects with a disjointed gradient change between losses and gains, though. However, another study of income found a curvilinear effect on life satisfaction, but with losses concave and gains largely linear, rather than convex and concave as prospect theory suggests (Vendrik & Woltjer, 2007).

It therefore seems likely that the positive relationship between expectation-achievement gaps and wellbeing will also be curvilinear here. However, most existing research has focused on salaries (Boyce et al., 2013, 2016; Vendrik & Woltjer, 2007), where losses may have serious life implications such as poverty. Accordingly, loss aversion would be expected, as research has shown that while lack of money may cause unhappiness, happiness does not increase once life becomes financially comfortable (Jebb et al., 2018). However, for football supporters, however painful a team's loss, there is always the next game or next competition to inspire hope. Indeed, in FWC2018, all but one team

was eliminated ultimately, so losses are essentially normalised. By contrast, only one team won FWC2018—and only eight of 211 national teams have ever won the World Cup (FIFA, 2018)—so winning is extremely rare and celebrated nationally. We therefore anticipate—contrary to prospect theory's loss aversion—that football supporters' well-being (i.e. satisfaction and happiness) will be more sensitive to wins (i.e. gains) than losses, a situation we name 'gain preference'. Indeed, a recent review of prospect theory studies concluded that loss aversion often does not occur, with gains and losses often perceived equally, or gains even disproportionately influential (Gal & Rucker, 2018). Our research would therefore contribute to this new theoretical debate and, as discussed above, we strengthen this contribution by also examining whether this positive effect is curvilinear, and convex, indicating gain preference. To do so, we will use polynomial regression and three-dimensional surface modelling (Edwards, 2002), for what we believe is the first time in research investigating expectation-achievement gaps in a wellbeing context. We therefore hypothesise:

H5: Relative changes from expectations to achievements will demonstrate a positive, convex, curvilinear relationship with satisfaction.

H6: Relative changes from expectations to achievements will demonstrate a positive, convex, curvilinear relationship with happiness.

Moderation effect of social identity

The extent to which expectation-achievement gaps influence both satisfaction and happiness is likely influenced by whether that domain is personally meaningful to those involved, however, with higher meaningfulness strengthening the positive relationship. Such meaningfulness has frequently been conceptualised through social identity theory (e.g. Hornsey, 2008; Turner et al., 1979). Social identities are fluid, multiple, and maximise both perceived intra-group similarity and inter-group differences (e.g. Hogg, 2018). Strong social identities therefore indicate the centrality of that construct to people's lives (Hornsey, 2008). This is particularly relevant here, as team identification is known to interact with expectations to determine spectators' loyalty (Trail et al., 2005).

Within football, a previous quasi-experiment found that three related types of football affinity—social identity, interest, and watching televised games live—moderated the positive effect of the 2016 European Football Championship on supporters' life satisfaction (Mutz, 2019). Furthermore, social identity has also been found to moderate the relationship between football teams' performance and their supporters' emotional reactions (Crisp et al., 2007). It therefore seems likely that it may also influence the relationship between expectation-achievement gaps and both satisfaction and happiness. Consequently, we hypothesise the following:

H7: The relationship between changes from expectations to achievements and satisfaction will be moderated by social identity, such that the positive relationship is stronger when identity is higher.

H8: The relationship between changes from expectations to achievements and happiness will be moderated by social identity, such that the positive relationship is stronger when identity is higher.

Integrating these themes, Figure 1 illustrates our theoretical model including the hypotheses (H) and measurement time-points (T).

Method

Participants

We recruited participants using micro-payments through the website *Prolific Academic* (n.d.), immediately before (time-point 1, T1) and after (time-point 3, T3) FWC2018 (which constituted time-point 2, T2; see *Design* sub-section), to complete an online questionnaire via *Qualtrics* (n.d.) software. Relative to alternative online research platforms, *Prolific Academic* participant samples are more diverse (Newman et al., 2021).

At T1, we recruited participants via two streams: (1) an experimental group, whose participants' national football teams would be competing in FWC2018; and (2) a control group, whose participants' national football teams would not be competing in FWC2018. Overall, we funded the recruitment of 500 participants, specifying 400 in the experimental group and 100 in the control group.

At T1, we received completed questionnaires from 501 participants overall: 402 in the experimental group and 99 in the control group. However, 29 participants selected the wrong stream, so we reallocated them to the correct group for the T3 data collection. At T3, we contacted the T1 participants again, via anonymous identification codes, and asked them to complete a second questionnaire. At T3, we received completed questionnaires from 391 participants overall, a response rate of 78.04%, comprising 297 in the experimental group and 94 in the control group.

We then screened these questionnaires. First, we matched participants across T1 and T3, and removed any non-matched T3 participants. Second, to ensure data integrity, we removed any participants with illogically inconsistent demographic data across time-points. Third, we removed any participants with T1 data only. Finally, we removed three participants with outlier data (see *Results* section). Following this screening, there were 341 participants remaining in the final sample: 278 participants in the experimental group and 63 in the control group, each matched across T1 and T3.

To analyse the quasi-experiment, in the first analysis phase, we divided these 278 participants in the experimental group into two experimental conditions: (1) 138 in

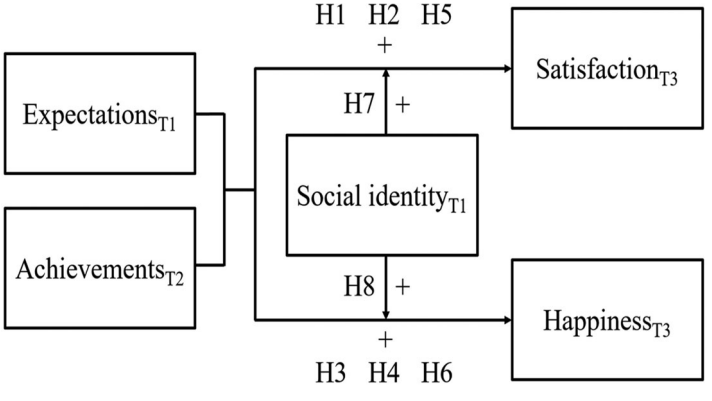

Figure 1. Theoretical framework and hypotheses.

the negative expectation-achievement gap condition, where expectations > achievements; and (2) 125 in the positive expectation-achievement gap condition, where expectations < achievements. A further 15 participants had no expectation-achievement gap (i.e. expectations = achievements) and were not therefore included in either experimental condition. Aside from the quasi-experiment, in the second analysis phase we further analysed the data from all 278 participants in the experimental group (see *Design* sub-section). Table 1 shows the detailed demographic data for all participants.

Design

We collected data longitudinally at three time-points: before (T1) and after (T3) FWC2018 for the questionnaires, and immediately after the championship stage at which a participant's national team had been eliminated from FWC2018 and their results were therefore known (T2). Participants completed the T1 questionnaire on 14 June 2018 before the first game of FWC2018, and the T3 questionnaire between 16 July and 18 July 2018 after the final game. The FWC2018 championship was 32 days in duration, from 14 June to 15 July 2018 inclusive, with the initial Group Stages finishing after the first 15 days (FIFA, 2018); so, T2 varied from day 15 to day 32 of FWC2018.[1]

The research was a 3 × 2 mixed-design quasi-experiment (i.e. without random participant allocation) with the results of FWC2018 constituting a naturally occurring real-world intervention (Grant & Wall, 2009). The first between-participants independent variable was the valence of the expectation-achievement gap (EA Gap, in figures and tables), with three conditions: (1) *negative*, where expectations > achievements; (2) *positive*, where expectations < achievements; and (3) *control*, for participants whose national team did not compete in FWC2018. The manipulation checks for these conditions were successful (see *Results* section). The second within-participants independent variable was time, with two conditions: (1) pre-FWC2018, and (2) post-FWC2018. The dependent variables were satisfaction and happiness.

Table 1. Demographic data for participants.

		Age			
Participant group	*N*	*M*	*SD*	Gender	National football team
Experimental (overall)	278	31.60	9.99	132 female 146 male	2 Argentina, 11 Australia, 4 Belgium, 5 Brazil, 2 Colombia, 5 Denmark, 126 England, 7 France, 8 Germany, 1 Iceland, 1 Japan, 4 Mexico, 28 Poland, 43 Portugal, 1 Russia, 27 Spain, 3 Sweden
Experimental (−EA Gap)	138	29.77	9.10	55 female 83 male	2 Argentina, 11 Australia, 1 Belgium, 5 Brazil, 2 Colombia, 1 Denmark, 12 England, 8 Germany, 1 Iceland, 1 Japan, 3 Mexico, 28 Poland, 42 Portugal, 21 Spain
Experimental (+EA Gap)	125	33.53	10.77	69 female 56 male	3 Belgium, 4 Denmark, 104 England, 6 France, 1 Mexico, 1 Russia, 4 Spain, 2 Sweden
Control	63	31.58	10.22	25 female 38 male	1 Albania, 3 Canada, 1 Czech Republic, 1 Finland, 3 Greece, 2 Hungary, 1 Indonesia, 16 Italy, 1 Latvia, 1 Netherlands, 2 Northern Ireland, 1 Philippines, 2 Republic of Ireland, 7 Scotland, 1 Singapore, 1 Slovenia, 1 Thailand, 1 Turkey, 13 USA, 2 Venezuela, 1 Wales, 1 Zimbabwe

Notes: EA Gap = expectation-achievement gap

As the data for valence of the expectation-achievement gap were continuous, we dichotomised this independent variable (see *Participants* sub-section) to create the experimental conditions and enable comparisons with the control condition in the first analysis phase. However, we also examined the full variance of this continuous independent variable in the second analysis phase with the longitudinal data from the entire experimental group. Figure 1 shows the measurement time-points for the variables examined in each hypothesis.

We avoided common method bias by collecting data on independent and dependent variables from separate sources (Podsakoff et al., 2003). We reinforced this methodological separation by not reminding participants of their national team's T2 performance in FWC2018 in the T3 questionnaire; rather, participants drew only on their memories of this real-world event.

Measures

We measured all variables using an online questionnaire (see *Participants* sub-section), at both T1 and T3 unless specified, except achievements which were determined at T2 by the results of FWC2018, an independent real-world event. We now detail the measures used.[2]

Demographics

Participants were asked *How old are you?* and *What is your gender?* (coded 0 = female; 1 = male) and selected the appropriate response categories.

National football team

Participants were asked *Which is your national football team?* and selected the appropriate response from a list of all recognised 211 national football teams (FIFA, 2018).

Expectations

Expectations were inevitably measured at T1 only (i.e. before FWC2018). Participants were asked *How do you think your national football team will perform in the 2018 FIFA World Cup?* and selected one of 32 rankings from a descending list ranging from *Final: Winning Team (1st rank overall)* to *Group Stage: 4th Place (32nd rank overall)*, reverse-coded 32 through 1, respectively, so that higher scores indicate higher performance, with an additional option for *Not competing in the 2018 FIFA World Cup*. These 32 rankings were nested within seven broad ranking levels officially recognised by FIFA (2018), namely: (1) *Final: Winning Team*, (2) *Final: Losing Team*, (3) *Semi-Final: Losing Team*, (4) *Quarter-Final: Losing Team*, (5) *Round of 16: Losing Team*, (6) *Group Stage: 3rd Place*, and (7) *Group Stage: 4th Place*. The full list of response rankings is provided in the Supplemental Online Material, for brevity.

Although, strictly, the data for expectations and achievements (below) were ordinal level, the distances between consecutive rankings were essentially equal intervals.[3] Furthermore, the corresponding Pearson and Spearman correlations from the analyses of these data were highly similar (see below). So, we treated these data as interval level to enable the use of parametric analyses, which are preferable to non-parametric

analyses—even of ordinal data—as the results correspond but afford more detail (Mircioiu & Atkinson, 2017).

Achievements

Achievements were inevitably measured at T2 only using the same 32 rankings as expectations, again coded so that higher scores indicate higher performance. After FWC2018, the official results for all 32 national teams were obtained from FIFA's (2018) website. These results listed only the seven broad ranking levels, so the more precise rankings were calculated within each level using the official FIFA (2018) criteria of *points* (for FWC2018 Group Stages only), then *goal difference*, and then, if national teams are still tied in rank, *goals scored*. Officially, FIFA only apply these criteria to determine positions within groups during the initial Group Phase of FWC2018; however, we also used the criteria here to determine rankings across groups and games within each of the official seven broad ranking levels.[4] However, the 32 rankings and seven broad ranking levels were extremely highly correlated for both expectations ($r = .89$; $\rho = .97$; both $p < .001$) and achievements ($r = .96$; $\rho = .98$; both $p < .001$).[5]

To determine whether participants were allocated to the negative or positive expectation-achievement gap experimental condition (see *Design* sub-section), we subtracted the T1 expected ranking from the T2 achieved ranking for each participant's national team. As rankings with higher values indicated higher performance (32 = 1st rank overall; 1 = 32nd rank overall), this calculation yielded positive values (i.e. expectation-achievement gaps) if achievements exceeded expectations and negative values if achievements failed to exceed expectations, and participants were allocated to experimental conditions accordingly. For instance, if a participant's national team was expected to finish 4th in FWC2018, but actually finished 1st, then the expectation-achievement gap would be +3 (i.e. 32–29), while if the team actually finished 10th, then the expectation-achievement gap would be −6 (i.e. 23–29).

Multi-item scales

The variables satisfaction, social identity, and happiness were each measured by multi-item scales, selected or adapted from the research literature for their psychometric properties and conceptual alignment. In each case, a 7-point Likert rating scale was used for optimal discriminatory power (Preston & Colman, 2000; Robinson, 2018), with the following standard rating scale anchors (see e.g. Baltes et al., 2007): *strongly disagree* (0), *disagree* (1), *slightly disagree* (2), *neutral* (3), *slightly agree* (4), *agree* (5), *strongly agree* (6). Scale scores were calculated as the mean of the constituent items and higher scores indicated higher levels of each variable.

Satisfaction

Satisfaction was measured with four items adapted from a stem ('I am satisfied with … ') that has been used to measure satisfaction both as a single-item (e.g. de Jonge et al., 2010) and within multi-item scales (e.g. Lee et al., 2014). At T1, the items read *I am satisfied with my national football team's performance in the last year*, *I am satisfied with the way my national football team have played in the last year*, *I am satisfied with my national football team's recent results*, and *I am satisfied with the current effectiveness of my national football team*. At T3, the control group received the same four items. While

the experimental group also received the same last two items at T3, for the first two items we changed the focal period from *the last year* to *the FIFA World Cup 2018* with the remaining text unchanged from T1.

For the first analysis phase, we created a satisfaction scale from the last two items only ($satisfaction_{short}$) to ensure that participants in the experimental and control conditions responded to identical items at T1 and T3. However, for the second analysis phase, where we only examined data from the experimental group, we created a satisfaction scale from all four items ($satisfaction_{full}$) at T1 and T3.

Social identity

Social identity was measured with three items drawn from different sources. First, we used Postmes et al.'s (2013) general single-item measure, namely *I identify with my national football team*. We combined this item with two sport-focused items from Robinson and Trail's (2005) attachment to the team scale, namely *I consider myself to be a 'real' fan of my national football team* and *Being a fan of my national football team is very important to me*. We inserted the focal term *my national football team* into these items.

Happiness

Happiness was measured with three items from Hills and Argyle's (2002) happiness scale, namely *I am very happy at the moment*, *Life is good at the moment*, and *I often experience joy and elation at the moment*. We appended the words *at the moment* to the original items to distinguish clearly between the two measurement time-points (i.e. so that the second rating period did not overlap with the first in participants' perceptions).

Stressful events

Finally, to control for other life events that could have affected wellbeing during the study, we asked participants *Have you experienced any especially stressful events in the last month?*, with *yes*/*no* response options, and *If you have answered 'yes' to the previous question, please describe the stressful event(s) generally and briefly (in just one sentence or a few words)*. Where possible, we coded the qualitative responses to this second question into the relevant 'life event' categories of Holmes and Rahe's (1967) social readjustment scale. We then allocated each coded life event its updated numerical weight from Scully et al.'s (2000) research and summed these at each time-point for each participant as a composite measure. The lead author coded the qualitative responses and these were then verified by a research assistant (95.65% inter-rater agreement), with the few disagreements resolved.

Results

Data screening

Initial data screening eliminated 157 participants with non-valid questionnaires or incomplete longitudinal data (see *Method* section). We screened the data further for impossible values, missing values, skew, kurtosis, and outliers outside $M \pm 3SD$ within the experimental variables (see Osborne & Overbay, 2004). The variables' distributions demonstrated little skew or kurtosis (most $<|1.00|$; all $<|1.70|$), but three outliers were

identified in the variable $expectations_{T1}$ and excluded. Following all screening, there were 278 participants in the experimental group—comprising 138 in the negative expectation-achievement gap condition, 125 in the positive expectation-achievement gap condition, and 15 with no expectation-achievement gap—and 63 in the control group (see Table 1).

Manipulation checks

We conducted manipulation checks on the first independent variable—valence of the expectation-achievement gap—which successfully confirmed that: (a) the expectation-achievement gap in the negative condition was negative ($M = -10.28$, $SD = 6.33$) and different from zero, $t(137) = -19.08$, $p < .001$, Cohen's $d = -1.62$; (b) the expectation-achievement gap in the positive condition was positive ($M = 6.42$, $SD = 5.05$) and different from zero, $t(124) = 14.20$, $p < .001$, $d = 1.27$; and (c) the expectation-achievement gap in each of these conditions differed from each other, $t(261) = -23.49$, $p < .001$, $d = -2.92$.

Next, to verify the construct validity of the multi-item scales, we conducted an exploratory factor analysis of the T1 data for the 341 participants in the experimental and control groups for the four satisfaction items, the three social identity items, and the three happiness items. The hypothesised three-factor solution was clearly identified from the scree-plot and extracted using principal axis factoring with varimax rotation, accounting for 78.67% of variance. All items loaded highly ($\geq .75$) on their intended scale's factor with correspondingly high Cronbach's alphas ($\geq .88$; see Table 2). Furthermore, there were no cross-loadings $\geq .30$, thereby verifying the construct validity of the three scales (Robinson, 2018).

Finally, to test the conceptual distinctiveness of the expectation-achievement gap from achievement alone, we conducted hierarchical analyses using both analysis of variance and regression, as reported below. Descriptive statistics and Pearson correlations for all variables are shown in Table 2 for the experimental group of participants ($n = 278$).

Analysis phase 1: quasi-experiment

To test Hypotheses 1–4, we conducted a 2×3 mixed-design multiple analysis of covariance (MANCOVA) of the quasi-experiment data. Time (pre-FWC2018, post-FWC2018) was the within-participants independent variable and expectation-achievement gap (negative, positive, control) was the between-participants independent variable, while satisfaction and happiness were the dependent variables, and age, gender, and stressful events were the covariates we controlled.

The multivariate interaction between time and expectation-achievement gap for the two dependent variables collectively was highly significant, Wilks's lambda (Λ) = .57, $F(4, 636) = 51.53$, $p < .001$, $\eta_p^2 = .24$, justifying further examination.

The subsequent univariate interaction between time and expectation-achievement gap was highly significant for satisfaction, $F(2, 319) = 120.00$, $p < .001$, $\eta_p^2 = .43$, supporting Hypothesis 1. These trends are clearly visible in Figure 2, on which data we also conducted the following post-hoc pairwise comparisons with Bonferroni adjustments. First, satisfaction decreased from before ($M = 4.10$, $SD = 1.47$) to after ($M = 2.11$, $SD =$

Table 2. Descriptive statistics and Pearson correlations of variables.

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Expectation$_{T1}$	24.96	5.70	—													
2. Achievement$_{T2}$	22.74	8.31	.04	—												
3. Satisfaction$_{T1(full)}$	3.55	1.39	.44**	−.21**	(.92)											
4. Satisfaction$_{T3(full)}$	3.33	1.97	−.05	.62**	−.08	(.97)										
5. Satisfaction$_{T1(short)}$	3.60	1.45	.39**	−.17**	.95**	−.03	(.88)									
6. Satisfaction$_{T3(short)}$	3.22	1.98	−.04	.60**	−.06	.98**	−.01	(.94)								
7. Social identity$_{T1}$	3.14	1.82	.36**	−.09	.50**	−.05	.44**	−.05	(.91)							
8. Social identity$_{T3}$	3.14	1.70	.28**	.18**	.35**	.24**	.32**	.22**	.69**	(.89)						
9. Happiness$_{T1}$	3.91	1.32	.03	−.03	.14*	.00	.14*	.02	.14*	.16**	(.89)					
10. Happiness$_{T3}$	3.85	1.30	.07	.04	.10	.05	.07	.07	.08	.13*	.68**	(.91)				
11. Age$_{T1}$	31.60	9.99	.04	.23**	−.01	.12*	−.02	.12	.03	.08	−.07	.00	—			
12. Gender$_{T1}$	—	—	.07	−.19**	.10	−.23**	.08	−.21**	.13*	−.05	−.14*	−.17**	−.11	—		
13. Stressful events$_{T1}$	15.50	22.85	−.01	.04	−.09	.10	−.06	.09	−.05	−.02	−.18**	−.10	−.03	−.10	—	
14. Stressful events$_{T3}$	10.48	19.94	−.10	.05	−.11	.02	−.08	.01	−.09	−.05	−.08	−.12*	.01	−.10	.35**	—

$N = 278$ (experimental group). * $p < .05$. ** $p < .01$. Cronbach's alphas are shown in parentheses on the leading diagonal.

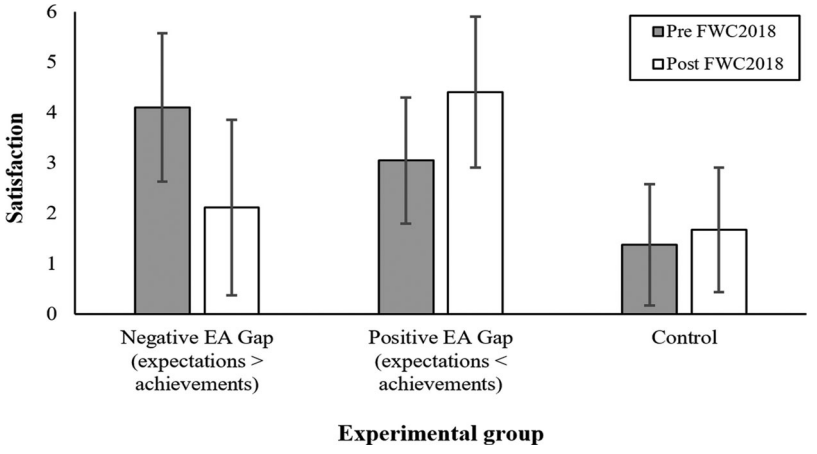

Figure 2. Ratings of satisfaction with the national team for the experimental groups.

1.74) FWC2018 for the experimental group whose national team's achievements failed to exceed expectations (i.e. the negative expectation-achievement gap condition), $t(137) = 11.54$, $p < .001$, $d = 1.23$. Second, satisfaction increased from before ($M = 3.05$, $SD = 1.25$) to after ($M = 4.40$, $SD = 1.50$) FWC2018 for the experimental group whose national team's achievements exceeded expectations (i.e. the positive expectation-achievement gap condition), $t(124) = -9.18$, $p < .001$, $d = -0.98$. Finally, satisfaction remained stable from before ($M = 1.37$, $SD = 1.20$) to after ($M = 1.67$, $SD = 1.24$) FWC2018 for the control group with no national team in FWC2018, $t(62) = -2.22$, $p = .03$ (non-significant with Bonferroni adjustment), $d = -0.25$.

However, the equivalent univariate interaction for happiness was not significant, $F(2, 319) = 0.45$, $p = .64$, $\eta_p^2 = .00$, not supporting Hypothesis 3. Happiness was unchanged from before to after FWC2018 for the negative expectation-achievement gap condition ($M = 4.02$, $SD = 1.32$, versus $M = 3.88$, $SD = 1.20$, respectively), the positive expectation-achievement gap condition ($M = 3.76$, $SD = 1.35$, versus $M = 3.76$, $SD = 1.42$, respectively), and the control condition ($M = 3.59$, $SD = 1.42$, versus $M = 3.51$, $SD = 1.42$, respectively).

Finally, to verify the conceptual distinctiveness of the expectation-achievement gap from achievement alone, we conducted another identical MANCOVA but controlling for achievement as a covariate. As the national teams of those in the control condition did not compete in FWC2018, they had no achievement data and it was therefore only possible to conduct this second MANCOVA with participants from the positive and negative experimental conditions.

As with the first MANCOVA, the multivariate interaction between time and expectation-achievement gap for the two dependent variables, satisfaction and happiness, was highly significant, $\Lambda = .80$, $F(2, 255) = 32.68$, $p < .001$, $\eta_p^2 = .20$, justifying further examination. The subsequent univariate interaction was highly significant again for satisfaction, $F(1, 256) = 65.31$, $p < .001$, $\eta_p^2 = .20$, verifying the conceptual distinctiveness of the expectation-achievement gap over and above achievement alone in relation to satisfaction, thereby supporting Hypothesis 2b. However, the equivalent univariate interaction for happiness was not significant, $F(1, 256) = 0.01$, $p = .94$, $\eta_p^2 = .00$, not

supporting Hypothesis 4b. In both cases, trend directions aligned with those in the original MANCOVA.

Analysis phase 2: polynomial regression

To test Hypotheses 1–4 further, and to test Hypotheses 5–8, we analysed data from the experimental group of participants ($n = 278$) only. Table 3 shows the results of these hierarchical polynomial regression analyses predicting satisfaction and happiness.

We used the moderated polynomial regression procedures advocated by Edwards (2002), as described here and below. This enabled us to examine the three-dimensional curvilinear surface graph of the relationships between predictors and outcomes. Our outcome variable was T3 satisfaction (Z1) or T3 happiness (Z2) (i.e. post-FWC2018) for the analyses reported in Table 3. This difference aside, the remaining procedures were identical for each set of analyses, as described now. In Step 1, we included our control variables and the T1 measurement of the outcome variable to control for baseline levels. In Step 2, we included the predictors T1 expectations (X) and T2 achievements (Y), both mean-centred on the joint mean of both variables to ensure the zero-points of each represented identical FWC2018 performance rankings which was conceptually important here. In Step 3, we included the X^2, XY, and Y^2 terms, with a significant R^2 change (Δ) indicating a curvilinear three-dimensional relationship explained collectively by these three terms. In Step 4, we included the mean-centred potential moderator, social identity (V). Finally, in Step 5, we included the XV, YV, X^2V, XYV, and Y^2V terms, with a significant R^2 change indicating a moderation effect.

Table 3 reports the results of the hierarchical polynomial regression predicting satisfaction. After entering the control variables at Step 1, expectations and achievements collectively predicted a highly significant and substantial additional 35.22% of variance (adjusted $R^2 \times 100$) in satisfaction at Step 2, with both variables significant predictors, supporting Hypothesis 1. However, against predictions, achievements (.64, $p < .001$) were a more powerful predictor than expectations (−.13, $p = .02$), not supporting Hypothesis 2a. Figure 3a shows the accompanying linear interaction surface graph[6] at Step 2.

At Step 3, the polynomial interaction between expectations and achievements predicted a highly significant incremental 5.53% of variance in satisfaction, supporting Hypothesis 5. Figure 3b shows the accompanying curvilinear polynomial interaction surface at Step 3. However, while social identity was not a significant predictor (Step 4), its moderating effect (Step 5) was approaching significance ($p = .08$) in partial support of Hypothesis 7. Figure 4 shows this moderating effect at (a) low (0.00–1.99), (b) medium (2.00–3.99), and (c) high (4.00–6.00) levels of social identity.

Table 3 also reports the results of the equivalent hierarchical polynomial regression predicting happiness. However, neither Steps 2, 3, 4, or 5 predicted any incremental variance in happiness over the control variables, not supporting Hypotheses 3, 4a, 6, or 8.[7]

Discussion

Broadly, the current study investigated expectation-achievement gaps and their effects on satisfaction and happiness in the context of people's national football team's performance

Table 3. Moderated polynomial regression analyses predicting satisfaction and happiness.

	$Satisfaction_{T3}$					$Happiness_{T3}$				
Variable	Step 1	Step 2	Step 3	Step 4	Step 5	Step 1	Step 2	Step 3	Step 4	Step 5
Age_{T1}	.10†	−.03	−.04	−.05	−.04	.04	.03	.04	.04	.03
$Gender_{T1}$	−.21***	−.12*	−.11*	−.11*	−.12*	−.07	−.07	−.08	−.07	−.09†
$Stressful\ events_{T1}$	.09	.09†	.06	.06	.07	.05	.05	.05	.05	.04
$Stressful\ events_{T3}$	−.04	−.06	−.03	−.03	−.05	−.09†	−.09†	−.08†	−.08†	−.07
$Satisfaction_{T1}$ (or) $Happiness_{T1}$	−.05	.12*	.20***	.18**	.20***	.67***	.67***	.66***	.66***	.66***
$Expectations_{T1}$		−.13*	−.13*	−.14*	−.18**		.04	.02	.03	.07
$Achievements_{T2}$		.64***	1.05***	1.07***	1.11***		.04	−.11	−.14	−.16
$Expectations^2_{T1}$			−.11*	−.11*	−.08			−.01	−.01	−.02
$Expectations_{T1} \times Achievements_{T2}$			−.02	−.02	−.08			.06	.06	.06
$Achievements^2_{T2}$			.47***	.49***	.51***			−.19†	−.21*	−.24*
$Identity_{T1}$				.06	−.03				−.05	−.05
$Expectations_{T1} \times Identity_{T1}$					−.01					.00
$Achievements_{T2} \times Identity_{T1}$					.13					.15†
$Expectations^2_{T1} \times Identity_{T1}$					.12†					−.13†
$Expectations_{T1} \times Achievements_{T2} \times Identity_{T1}$					−.09					.01
$Achievements^2_{T2} \times Identity_{T1}$					.07					.16
R^2 *(adjusted)*	.06***	.41***	.46***	.47***	.48***	.47***	.47***	.47***	.47***	.48***
ΔR^2		.35***	.06***	.00	.01†		.00	.00	.00	.01

$N = 278$ (experimental group). † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

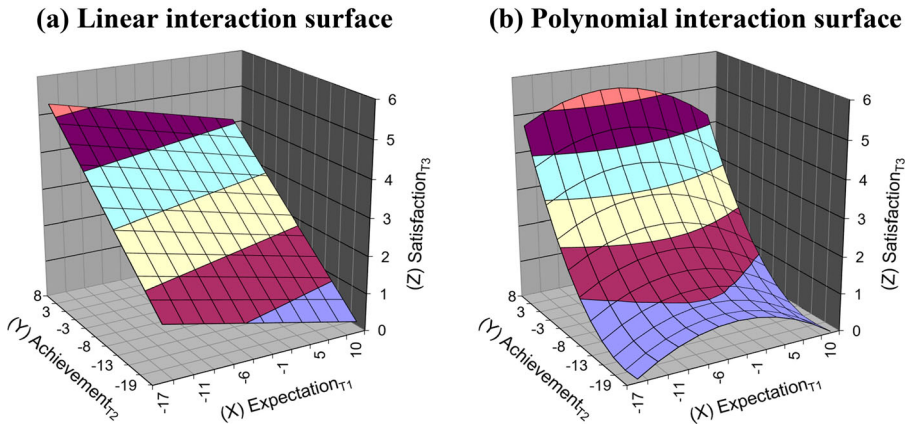

Figure 3. Linear and polynomial interaction surfaces between expectation and achievement predicting satisfaction.

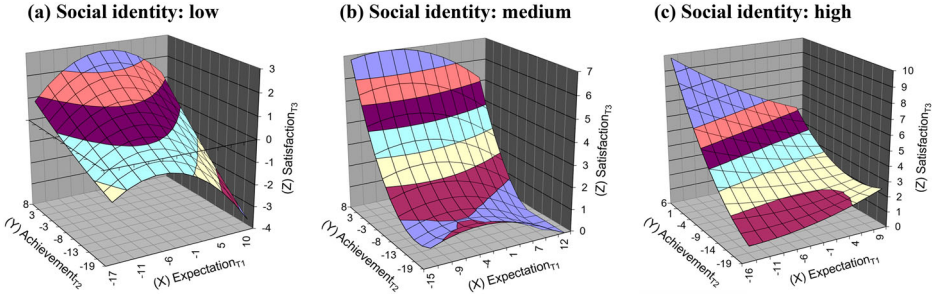

Figure 4. Moderating effect of social identity on the polynomial interaction surface between expectation and achievement predicting satisfaction.

in FWC2018. As well as being of interest as the world's most popular spectator sport (Palacios-Huerta, 2004), this research offered key methodological advantages for well-being research. First, FWC2018 is a real-world scenario of major psychological significance (e.g. Encrenaz et al., 2012). Second, FWC2018 provided accurate, objective data concerning expectations and achievements, which are rare outside of laboratory experiments. Third, by conducting a quasi-experiment around FWC2018 with a control group, we balanced the higher internal validity of experiments and the higher external validity of real-world research (Fiske, 2016). Consequently, this quadrennial event enabled us to examine satisfaction and happiness through the lens of expectation-achievement gaps in an accurate and meaningful way.

Linear effects on satisfaction

The key significant results related to satisfaction. Hypothesis 1, that changes from expectations to achievements would be positively related to satisfaction, was strongly supported by both analysis phases. This aligns with previous empirical research (e.g. Li & Fung, 2012; Shepperd & McNulty, 2002) and expectation disconfirmation theory (Lankton et al., 2014; Oliver, 1980). The MANCOVA and post-hoc analyses of the quasi-experiment data showed highly significant effects—with satisfaction increasing

when achievements exceeded expectations, decreasing when achievements failed to exceed expectations, and remaining stable in the control condition—and provide strong evidence of causality. Step 2 of the subsequent hierarchical regression in Table 3 found that expectations and achievements explained a substantial 35% of incremental variance in satisfaction, an effect of considerable size (Cohen, 1992).

Contrary to Hypothesis 2a, though, achievements explained more variance in satisfaction than expectations did in the regression, although both were significant predictors (see Table 3). This illuminates ongoing theoretical debates, by suggesting that absolute achievements determine satisfaction more than relative achievements (i.e. expectation-achievement gaps). These findings align with earlier research showing higher, absolute salaries are related to higher wellbeing (Diener et al., 1993; Veenhoven, 1991), but contradict recent research findings that salary relative to immediate reference groups predicts satisfaction more (Boyce et al., 2010; D'Ambrosio & Frick, 2007). The latter explanation has become more accepted, aligned with both expectation disconfirmation theory (Lankton et al., 2014; Oliver, 1980) and decision affect theory (Mellers et al., 1997). However, the very latest research on income (Killingsworth, 2021), and our research on football supporters' satisfaction here, suggest that absolute achievements explain wellbeing more, which warrants further research after neglect.

Nevertheless, expectations did still explain incremental variance in satisfaction over and above achievements, verifying the conceptual distinctiveness of the expectation-achievement gap and supporting Hypothesis 2b. These trends are visible in Figure 3a, illustrating the highly significant Step 2 result from the regression in Table 3, with two clear slopes to the straight three-dimensional surface. The first, along the y-axis, shows a highly significant steep increase in satisfaction as achievements increase, and the second, along the x-axis, shows a less steep but still significant increase in satisfaction as expectations decrease. Although expectations and achievements have been examined in football before (Biscaia et al., 2012; González-Gómez & Picazo-Tadeo, 2010), this is the first study to measure them, and their relationship with wellbeing, directly.

Curvilinear effects on satisfaction

Next, we examined the associated curvilinear relationships. Figure 3b shows the curvilinear interaction effects of expectations and achievements on satisfaction, representing the highly significant incremental effect of the Step 3 polynomial regression terms in Table 3, supporting Hypothesis 5. Detecting these curvilinear effects advances previous theory that has focused mostly on linear effects, as we discuss below. Indeed, we believe this is the first examination of the effects of expectation-achievement gaps on wellbeing using polynomial regression and three-dimensional surface modelling.

In Figure 3b, two clear slopes are again visible in the three-dimensional surface, both curvilinear this time and in opposite directions yielding a saddle-like surface (Edwards, 2002). The first curved slope, along the y-axis, is convex, and shows a highly significant increase in satisfaction as achievements increase, aligned with the equivalent results in Figure 3a and expectation disconfirmation theory (Lankton et al., 2014; Oliver, 1980), but curvilinear. However, for low achievements, this satisfaction slope is almost flat before steepening sharply for high achievements, yielding the convex curve observed. This is theoretically and practically interesting, suggesting that satisfaction increases at

a disproportionately high rate at higher levels of achievement. As this curve represents the effect of achievements, having controlled for expectations (i.e. relative to expected reference points), it is directly comparable to the equivalent prospect theory curve (Barberis, 2013; Kahneman & Tversky, 1979); although the three-dimensional surface modelling (Edwards, 2002) is richer. Here, however, we have identified 'gain preference'—where satisfaction is more sensitive to wins (gains) than losses—in contrast to prospect theory's 'loss aversion'. This illuminates current theoretical debates questioning loss aversion, and recommending identification of its preconditions (Gal & Rucker, 2018). Indeed, we believe gain preference occurs in situations where losing is much more likely than winning (only one of 32 national teams won FWC2018), where winning elicits euphoria, and where losses are disappointing but endurable (unlike poverty). Perhaps, in football and other sports, to quote Abba's famous lyrics, 'the winner takes it all', or at least a large share of the available satisfaction.

The second slope in Figure 3b, along the x-axis, differs somewhat from the equivalent slope in Figure 3a, however. Here, the relationship between expectations and satisfaction has a shallow, concave shape, where satisfaction is higher for moderate expectations than for low or high expectations. The right side of this curve aligns with predictions, as satisfaction is higher for moderate than high expectations. However, the left side of this curve is unexpected and counter-intuitive, as satisfaction is higher for moderate than low expectations. This concave pattern is challenging to explain, but of substantial theoretical and practical interest. As expected, when moving from high to moderate levels of expectations, satisfaction increases for any given level of achievement. This aligns with expectation-disconfirmation theory (Lankton et al., 2014; Oliver, 1980), as any given level of achievement appears relatively higher (i.e. more satisfying) with lower expectations. Furthermore, with moderate expectations, it is still possible to perceive even high achievements as both realistic and deserved, even if unlikely. However, for low expectations—where satisfaction was lower than for moderate expectations—even modest achievements may seem unrealistic, but particularly very high achievements which may be perceived as lucky. This is particularly likely in football, where improbable results can occur, such as the 'lottery' of penalty shootout wins (Wood et al., 2015). So, in these circumstances, while supporters may be initially delighted with unexpected success, this is perhaps tempered by sadness that it will probably not happen again. Indeed, attribution theory suggests that surprise success may be perceived as lucky and unstable, leading to uncertainty and anxiety about the future (Weiner, 2014). Such pessimism would therefore perhaps manifest in slightly lower satisfaction for supporters with low expectations than for those with moderate expectations.

Moderating effect of social identity

The moderating effect of social identity on the relationship between changes from expectations to achievements and satisfaction approached significance ($p = .08$), partially supporting Hypothesis 7. This aligns with previous research suggesting social identity moderates the relationships between viewing championship football and life satisfaction (Mutz, 2019), and between a team's performance and supporters' emotional reactions (Crisp et al., 2007). However, unlike previous studies, we examined a three-way interaction between expectations, achievements, and social identity as predictors of satisfaction (see Figure 4), providing novel theoretical insights about these effects and when they occur.

At medium levels of social identity, the interaction surface in Figure 4b resembles the overall curvilinear surface in Figure 3b. Furthermore, the increase in satisfaction along the y-axis resulting from increasing achievements also remains similar at low and high levels of social identity, in Figures 4a and 4c respectively, although largely linear and steeper in the latter case. However, the concave relationship between expectations and satisfaction occurs mainly at low levels of social identity, in Figure 4a. At medium levels of social identity, in Figure 4b, this concave curve is shallower and the relationship between expectations and satisfaction is slightly negative overall due to the slight clockwise rotation of the surface relative to the x-axis. At high levels of social identity, in Figure 4c, this predicted negative relationship between expectations and satisfaction becomes strongly negative overall and straighter, except the slight positive slope for low achievements.

Happiness

The results for happiness were non-significant in both analysis phases. Counter to Hypotheses 3, 4a, and 4b, expectations and achievements had no effect on happiness in the quasi-experiment and explained no incremental variance in happiness in the regression. Neither did the polynomial terms explain any incremental variance in happiness, counter to Hypothesis 6. This is surprising, given the extremely large effect on satisfaction, and suggests that happiness—as a broader and more general construct (Hills & Argyle, 2002)—is further removed affectively from sporting events than event-specific satisfaction is.

Accordingly, we expected that supporters who identify more strongly with their national team would be more likely to experience happiness, aligned with previous research (Crisp et al., 2007; Mutz, 2019). However, contrary to Hypothesis 8, the moderation effect of social identity on the relationship between changes from expectations to achievements and happiness was also non-significant. It is possible that our social identity measure was not sufficiently broad conceptually to capture the extent of supporters' engagement with both their national team and FWC2018, though. For instance, alongside social identity, Mutz (2019) also measured supporters' interest in football generally and whether they watched televised games live. While these three measures of football affinity overlapped, statistically and theoretically, and each moderated the effect of the championship on life satisfaction (essentially happiness), they were conceptually distinct. So, including these broader affinity measures here may have captured this moderation effect.

Conclusions

The current study harnessed FWC2018 to conduct a quasi-experiment examining the effect of expectation-achievement gaps on supporters' satisfaction. FWC2018 provided the methodologically optimal conditions of a psychologically meaningful real-world context, accurate measurement of expectations and achievements, and a quasi-experiment with a control group. Collectively, expectations and achievements explained a substantial 41% of variance in satisfaction, including newly discovered curvilinear effects of theoretical and practical importance. While achievements predicted satisfaction more strongly, expectations still explained incremental variance thereby demonstrating the

conceptual distinctiveness of the expectation-achievement gap. Finally, there is promising evidence that social identity moderates the curvilinear relationship between changes from expectations to achievements and satisfaction.

Notes

1. Although, for a few low-performing national teams, their elimination from FWC2018 at the Group Stages may have been apparent before day 15.
2. We included four further items in this questionnaire, which we report here for transparency. First, we developed one item to measure participants' engagement with their national team, namely *I can name the players in my national football team*. However, when factor analysed, this item loaded on the social identity scale, albeit lower than the three social identity items, so we excluded it. Second, we included a 3-item measure of optimism (Scheier et al., 1994) to test whether optimism moderated the relationship between changes from expectations to achievements and (a) satisfaction, and (b) happiness. However, we found no significant moderation effects, so for brevity we only report this here.
3. The absolute goal difference in the FWC2018 Knockout Phase ($M = 0.60$, $SD = 0.74$) or points difference in the FWC2018 Group Phase ($M = 0.27$, $SD = 0.59$) between each pair of national teams listed in descending rank order was broadly uniform.
4. Except the rankings of third and fourth, which were determined by the official third place playoff game between the losing teams from each semi-final (FIFA, 2018).
5. We used the 32 rankings for expectations and achievements here, rather than the seven broad ranking levels, for two reasons. First, the 32 rankings were what participants responded to in the questionnaire. Second, the 32 rankings offer more detail about participants' perceptions. We are not suggesting that supporters are specifically interested in points and goal difference *per se*. Rather, these statistics mathematically capture supporters' psychological perceptions of the winning or losing margins in games and stages (e.g. a resounding 4–0 win will be perceived differently to a narrow 1–0 win), and comparisons with rivals (e.g. 'We lost, but at least we didn't get hammered 3–0 like [Team X]!'). Nevertheless, the two sets of rankings are very similar, as demonstrated by the correlations reported here and the comparison of analyses reported later in the *Results* section.
6. We plotted Figures 3 and 4 using response surface Excel macros from Edwards (n.d.) based on Edwards (2002).
7. Finally, for precaution, we reran the analyses using the seven broad ranking levels for expectations and achievements. The significance of most results was the same as found for the 32 rankings, as reported above—with Hypotheses 1 and 2(b) also supported, and Hypotheses 2 (a), 3, 4(a), 4(b), 6, and 8 also rejected—indicating high correspondence. However, Hypothesis 5 no longer reached significance (the curvilinear relationship between the expectation-achievement gap and satisfaction) and Hypothesis 7 changed from approaching significance to highly significant (the curvilinear moderating effect of social identity). Nevertheless, the analyses reported above using the 32 rankings take precedence for methodological reasons (see *Method* section).

Acknowledgments

We thank Emma Fairclough for verifying the coding of the qualitative data from the stressful events item. The data collection was funded by the University of Leeds.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The data collection was funded by the University of Leeds.

ORCID

Mark A. Robinson http://orcid.org/0000-0001-5535-8737
Matthew C. Davis http://orcid.org/0000-0003-1577-7544
Kerrie L. Unsworth http://orcid.org/0000-0002-0826-7565

References

Baltes, B. B., Bauer, C. B., & Frensch, P. A. (2007). Does a structured free recall intervention reduce the effect of stereotypes on performance ratings and by what cognitive mechanism? *Journal of Applied Psychology*, *92*(1), 151–164. https://doi.org/10.1037/0021-9010.92.1.151

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, *27*(1), 173–196. https://doi.org/10.1257/jep.27.1.173

Biscaia, R., Correia, A., Rosado, A., Maroco, J., & Ross, S. (2012). The effects of emotions on football spectators' satisfaction and behavioural intentions. *European Sport Management Quarterly*, *12*(3), 227–242. https://doi.org/10.1080/16184742.2012.679949

Boyce, C. J., Brown, G. D. A., & Moore, S. C. (2010). Money and happiness: Rank of income, not income, affects life satisfaction. *Psychological Science*, *21*(4), 471–475. https://doi.org/10.1177/0956797610362671

Boyce, C. J., Wood, A. M., Banks, J., Clark, A. E., & Brown, G. D. A. (2013). Money, well-being, and loss aversion: Does an income loss have a greater effect on well-being than an equivalent income gain? *Psychological Science*, *24*(12), 2557–2562. https://doi.org/10.1177/0956797613496436

Boyce, C. J., Wood, A. M., & Ferguson, E. (2016). Individual differences in loss aversion: Conscientiousness predicts how life satisfaction responds to losses versus gains in income. *Personality and Social Psychology Bulletin*, *42*(4), 471–484. https://doi.org/10.1177/0146167216634060

Bücker, S., Nuraydin, S., Simonsmeier, B. A., Schneider, M., & Luhmann, M. (2018). Subjective well-being and academic achievement: A meta-analysis. *Journal of Research in Personality*, *74*, 83–94. https://doi.org/10.1016/j.jrp.2018.02.007

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Crisp, R. J., Heuston, S., Farr, M. J., & Turner, R. N. (2007). Seeing red or feeling blue: Differentiated intergroup emotions and ingroup identification in soccer fans. *Group Processes & Intergroup Relations*, *10*(1), 9–26. https://doi.org/10.1177/1368430207071337

D'Ambrosio, C., & Frick, J. R. (2007). Income satisfaction and relative deprivation: An empirical link. *Social Indicators Research*, *81*(3), 497–519. https://doi.org/10.1007/s11205-006-0020-0

de Jonge, J., van Vegchel, N., Shimazu, A., Schaufeli, W., & Dormann, C. (2010). A longitudinal test of the demand-control model using specific job demands and specific job control. *International Journal of Behavioral Medicine*, *17*(2), 125–133. https://doi.org/10.1007/s12529-010-9081-1

Deloitte. (2020). *Deloitte football money league 2020*. Deloitte Sports Business Group.

Derksen, F. A. W. M., Olde Hartman, T., Bensing, J., & Lagro-Janssen, A. (2018). Empathy in general practice—The gap between wishes and reality: Comparing the views of patients and physicians. *Family Practice*, *35*(2), 203–208. https://doi.org/10.1093/fampra/cmx080

Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1993). The relationship between income and subjective well-being: Relative or absolute? *Social Indicators Research*, *28*(3), 195–223. https://doi.org/10.1007/BF01079018

Dolan, P., Kavetsos, G., Krekel, C., Mavridis, D., Metcalfe, R., Senik, C., Szymanski, S., & Ziebarth, N. R. (2016). The host with the most? The effects of the Olympic Games on happiness. *DIW, Berlin, Discussion Paper*, *1599*, 1–48.

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analysing behavior in organizations: Advances in measurement and data analysis* (pp. 350–400). Wiley.
Edwards, J. R. (n.d.). Response surface Excel macro. http://public.kenan-flagler.unc.edu/faculty/edwardsj/downloads.htm
Encrenaz, G., Contrand, B., Leffondré, K., Queinec, R., Aouba, A., Jougla, E., Miras, A., & Lagarde, E. (2012). Impact of the 1998 Football World Cup on suicide rates in France: Results from the National Death Registry. *Suicide and Life-Threatening Behavior*, *42*(2), 129–135. https://doi.org/10.1111/j.1943-278X.2011.00076.x
Eurosport. (2011). Huge TV figures for PL. *Eurosport*. https://www.eurosport.com/football/premier-league/2011-2012/huge-tv-figures-for-pl_sto2987062/story.shtml
FIFA. (2018). Fédération Internationale de Football Association (FIFA). https://www.fifa.com/
Fiske, S. T. (2016). How to publish rigorous experiments in the 21st century. *Journal of Experimental Social Psychology*, *66*, 145–147. https://doi.org/10.1016/j.jesp.2016.01.006
Gal, D., & Rucker, D. D. (2018). The loss of loss aversion: Will it loom larger than its gains? *Journal of Consumer Psychology*, *28*(3), 497–516. https://doi.org/10.1002/jcpy.1047
González-Gómez, F., & Picazo-Tadeo, A. J. (2010). Can we be satisfied with our football team? Evidence from Spanish professional football. *Journal of Sports Economics*, *11*(4), 418–442. https://doi.org/10.1177/1527002509341020
Grant, A. M., & Wall, T. D. (2009). The neglected science and art of quasi-experimentation: Why-to, when-to, and how-to advice for organizational researchers. *Organizational Research Methods*, *12*(4), 653–686. https://doi.org/10.1177/1094428108320737
Green, C. P., Heywood, J. S., Kler, P., & Leeves, G. (2018). Paradox lost: The disappearing female job satisfaction premium. *British Journal of Industrial Relations*, *56*(3), 484–502. https://doi.org/10.1111/bjir.12291
Hills, P., & Argyle, M. (2002). The Oxford Happiness Questionnaire: A compact scale for the measurement of psychological well-being. *Personality and Individual Differences*, *33*(7), 1073–1082. https://doi.org/10.1016/S0191-8869(01)00213-6
Hogg, M. A. (2018). Social identity theory. In P. J. Burke (Ed.), *Contemporary social psychological theories* (pp. 112–138). Stanford University Press.
Holmes, T. H., & Rahe, R. H. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research*, *11*(2), 213–218. https://doi.org/10.1016/0022-3999(67)90010-4
Hornsey, M. J. (2008). Social identity theory and self-categorization theory: A historical review. *Social and Personality Psychology Compass*, *2*(1), 204–222. https://doi.org/10.1111/j.1751-9004.2007.00066.x
Hult, G. T. M., Sharma, P. N., Morgeson III, F. V., & Zhang, Y. (2019). Antecedents and consequences of customer satisfaction: Do they differ across online and offline purchases? *Journal of Retailing*, *95*(1), 10–23. https://doi.org/10.1016/j.jretai.2018.10.003
Jebb, A. T., Tay, L., Diener, E., & Oishi, S. (2018). Happiness, income satiation and turning points around the world. *Nature Human Behaviour*, *2*(1), 33–38. https://doi.org/10.1038/s41562-017-0277-0
Jones, M. V., Coffee, P., Sheffield, D., Yangüez, M., & Barker, J. B. (2012). Just a game? Changes in English and Spanish soccer fans' emotions in the 2010 World Cup. *Psychology of Sport and Exercise*, *13*(2), 162–169. https://doi.org/10.1016/j.psychsport.2011.10.008
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292. https://doi.org/10.2307/1914185
Kavetsos, G., & Szymanski, S. (2010). National well-being and international sports events. *Journal of Economic Psychology*, *31*(2), 158–171. https://doi.org/10.1016/j.joep.2009.11.005
Killingsworth, M. A. (2021). Experienced well-being rises with income, even above $75,000 per year. *Proceedings of the National Academy of Sciences*, *118*(4), e2016976118. https://doi.org/10.1073/pnas.2016976118
Lankton, N., McKnight, D. H., & Thatcher, J. B. (2014). Incorporating trust-in-technology into expectation disconfirmation theory. *The Journal of Strategic Information Systems*, *23*(2), 128–145. https://doi.org/10.1016/j.jsis.2013.09.001

Lee, J. S., & Kang, J. H. (2015). Effects of sport event satisfaction on team identification and revisit intention. *Sport Marketing Quarterly*, *24*(4), 225–234. https://fitpublishing.com/articles/effects-sport-event-satisfaction-team-identification-and-revisit-intent

Lee, S. H. M., Qureshi, I., Konrad, A. M., & Bhardwaj, A. (2014). Proactive personality heterophily and the moderating role of proactive personality on network centrality and psychological outcomes: A longitudinal study. *Journal of Business and Psychology*, *29*(3), 381–395. https://doi.org/10.1007/s10869-013-9320-y

Li, T., & Fung, H. H. (2012). Partner discrepancies and age differences in marital quality: Distinguishing the ideal, expected, and perceived partner. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *67*(4), 417–422. https://doi.org/10.1093/geronb/gbr119

McBride, M. (2010). Money, happiness, and aspirations: An experimental study. *Journal of Economic Behavior & Organization*, *74*(3), 262–276. https://doi.org/10.1016/j.jebo.2010.03.002

McLean, J., Brennan, D., Wyper, D., Condon, B., Hadley, D., & Cavanagh, J. (2009). Localisation of regions of intense pleasure response evoked by soccer goals. *Psychiatry Research: Neuroimaging*, *171*(1), 33–43. https://doi.org/10.1016/j.pscychresns.2008.02.005

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, *8*(6), 423–429. https://doi.org/10.1111/j.1467-9280.1997.tb00455.x

Mircioiu, C., & Atkinson, J. (2017). A comparison of parametric and non-parametric methods applied to a Likert scale. *Pharmacy*, *5*(2), 1–12. https://doi.org/10.3390/pharmacy5020026

Mutz, M. (2019). Life satisfaction and the UEFA EURO 2016: Findings from a nation-wide longitudinal study in Germany. *Applied Research in Quality of Life*, *14*(2), 375–391. https://doi.org/10.1007/s11482-018-9599-y

Newman, A., Bavik, Y. L., Mount, M., & Shao, B. (2021). Data collection via online platforms: Challenges and recommendations for future research. *Applied Psychology*, *70*(3), 1380–1402. https://doi.org/10.1111/apps.12302

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460–469. https://doi.org/10.1177/002224378001700405

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*, *9*, article 6. https://doi.org/10.7275/qf69-7k43

Palacios-Huerta, I. (2004). Structural changes during a century of the world's most popular sport. *Statistical Methods and Applications*, *13*(2), 241–258. https://doi.org/10.1007/s10260-004-0093-3

Park, H. J., Li, R. X., Kim, J., Kim, S. W., Moon, D. H., Kwon, M. H., & Kim, W. J. (2009). Neural correlates of winning and losing while watching soccer matches. *International Journal of Neuroscience*, *119*(1), 76–87. https://doi.org/10.1080/00207450802480069

Pavot, W., & Diener, E. (2008). The Satisfaction With Life Scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology*, *3*(2), 137–152. https://doi.org/10.1080/17439760701756946

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Postmes, T., Haslam, S. A., & Jans, L. (2013). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, *52*(4), 597–617. https://doi.org/10.1111/bjso.12006

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, *104*(1), 1–15. https://doi.org/10.1016/S0001-6918(99)00050-5

Prolific Academic. (n.d.). Prolific academic research participant recruitment. https://www.prolific.co/

Qualtrics. (n.d.). Qualtrics online questionnaire hosting. https://www.qualtrics.com/uk/

Robinson, M. A. (2018). Using multi-item psychometric scales for research and practice in human resource management. *Human Resource Management*, *57*(3), 739–750. https://doi.org/10.1002/hrm.21852

Robinson, M. J., & Trail, G. T. (2005). Relationships among spectator gender, motives, points of attachment, and sport preference. *Journal of Sport Management*, *19*(1), 58–80. https://doi.org/10.1123/jsm.19.1.58

Ruvolo, A. P., & Veroff, J. (1997). For better or for worse: Real-ideal discrepancies and the marital well-being of newlyweds. *Journal of Social and Personal Relationships*, *14*(2), 223–242. https://doi.org/10.1177/0265407597142005

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait-anxiety, self-mastery, and self-esteem): A re-evaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*(6), 1063–1078. https://doi.org/10.1037/0022-3514.67.6.1063

Scully, J. A., Tosi, S., & Banning, K. (2000). Life event checklists: Revisiting the Social Readjustment Rating Scale after 30 years. *Educational and Psychological Measurement*, *60*(6), 864–876. https://doi.org/10.1177/00131640021970952

Shepperd, J. A., & McNulty, J. K. (2002). The affective consequences of expected and unexpected outcomes. *Psychological Science*, *13*(1), 85–88. https://doi.org/10.1111/1467-9280.00416

Trail, G. T., Anderson, D. F., & Fink, J. S. (2005). Consumer satisfaction and identity theory: A model of sport spectator conative loyalty. *Sport Marketing Quarterly*, *14*(2), 98–111. https://fitpublishing.com/content/consumer-satisfaction-and-identity-theory-model-sport-spectator-conative-loyalty

Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, *9*(2), 187–204. https://doi.org/10.1002/ejsp.2420090207

Vannier, S. A., & O'Sullivan, L. F. (2017). Passion, connection, and destiny: How romantic expectations help predict satisfaction and commitment in young adults' dating relationships. *Journal of Social and Personal Relationships*, *34*(2), 235–257. https://doi.org/10.1177/0265407516631156

Veenhoven, R. (1991). Is happiness relative? *Social Indicators Research*, *24*(1), 1–34. https://doi.org/10.1007/BF00292648

Vendrik, M. C., & Woltjer, G. B. (2007). Happiness and loss aversion: Is utility concave or convex in relative income? *Journal of Public Economics*, *91*(7–8), 1423–1448. https://doi.org/10.1016/j.jpubeco.2007.02.008

von Meyer-Höfer, M., Nitzko, S., & Spiller, A. (2015). Is there an expectation gap? Consumers' expectations towards organic. *British Food Journal*, *117*(5), 1527–1546. https://doi.org/10.1108/BFJ-07-2014-0252

Weiner, B. (2014). The attribution approach to emotion and motivation: History, hypotheses, home runs, headaches/heartaches. *Emotion Review*, *6*(4), 353–361. https://doi.org/10.1177/1754073914534502

Weiss, H. M. (2002). Deconstructing job satisfaction: Separating evaluations, beliefs and affective experiences. *Human Resource Management Review*, *12*(2), 173–194. https://doi.org/10.1016/S1053-4822(02)00045-1

Wood, G., Jordet, G., & Wilson, M. R. (2015). On winning the "lottery": Psychological preparation for football penalty shoot-outs. *Journal of Sports Sciences*, *33*(17), 1758–1765. https://doi.org/10.1080/02640414.2015.1012103