Received: 2 June 2021 | Revised: 6 September 2021 | Accepted: 4 October 2021

DOI: 10.1111/sms.14072

ORIGINAL ARTICLE

WILEY

The effect of descriptive norms and social identification on performance and exertion during a physical fitness task

Mark Stevens | **Áine Fitzpatrick** | **Tegan Cruwys**

Research School of Psychology, The Australian National University, Canberra, ACT, Australia

Correspondence
Mark Stevens, Research School of Psychology, The Australian National University, Canberra ACT 2601, Australia.
Email: mark.stevens@anu.edu.au

Funding information
National Health and Medical Research Council

Research suggests that people's perceptions of the typical physical activity behaviors of others—descriptive norms—shape their own physical activity. However, prior work has primarily used cross-sectional designs, focused on people's physical activity intentions or self-reported overall physical activity, and failed to attend to how the normative referent affects the norm-behavior relationship. In a pre-registered experiment, we manipulated the descriptive norm to explore its effect on perceived exertion during, and objectively assessed performance on, an exercise (running) task. We also measured the strength of participants' social identification as a member of the norm reference group as a potential moderator. Following a baseline trial, participants ($N = 78$) were either told that their baseline score on the running task was below average (high norm condition) or above average (low norm condition) compared to members of the reference group. Participants in both conditions tended to perform better in Trial 2 than Trial 1. However, participants in the high norm condition improved their performance significantly (2.5 times) more than participants in the low norm condition. Social identification moderated the effect of the norm manipulation on perceived exertion. High identifiers tended to exert themselves less than low identifiers during Trial 2 if they were told they were above average, whereas if told they were below average, high identifiers exerted themselves slightly more than low identifiers. Results provide evidence that descriptive norm messages can improve people's performance on physically demanding tasks, and indicate that descriptive physical activity norms may be more influential when the norm reference group is subjectively meaningful. Opportunities to fruitfully deploy norm messages in applied contexts are discussed.

KEYWORDS
behavior change, exertion, normative influence, performance, social identity

1 | INTRODUCTION

Identifying effective ways to improve people's physical activity behaviors is a priority for both sport and exercise psychology research. In sports contexts, the goal of behavior change attempts is often to improve performance. For example, researchers have examined how leaders can best motivate team members to increase their effort, and thus go further or faster in the interests of the team.[1,2] In exercise contexts, the ultimate goal of behavior

change attempts is typically improved health. A common proximal objective is to identify effective ways to increase people's overall volume of physical activity.[3,4] However, motivating greater effort within sessions is also valuable in these contexts because it can, for example, enhance the health and fitness benefits that people derive from exercise,[5] and help people to achieve their exercise goals (eg, to improve their personal best deadlift or 5 km run time).

For several decades, researchers have sought to ascertain whether social norms are one factor that can be harnessed to achieve physical activity behavior change. Recently, promising evidence has indicated that *descriptive physical activity norms*—perceptions of what is common and typical among others[6]—are related to people's own physical activity behaviors, and might be a fruitful target for efforts to improve people's physical activity behaviors.[7-12] However, this research has primarily used cross-sectional designs, and focused on people's physical activity intentions or self-reported overall physical activity, rather than on their observed, objectively measured behavior during specific physically demanding exercise tasks. Moreover, research to date has failed to pay sufficient attention to how the normative referent (ie, the source of the norm) affects the influence of norms on behavior. We sought to address these shortcomings and advance understanding of how descriptive norms can affect physical activity behaviors, generating knowledge with potential practical application in both sport and exercise contexts.

1.1 | Descriptive norms and physical activity behaviors

In the last decade, there has been growing interest in the role descriptive norms might play in shaping physical activity behaviors, with several observational studies reporting promising findings. For example, Priebe and Spink[8] found evidence for a positive relationship between descriptive physical activity norms of one's friends and participants' own self-reported total physical activity in a small sample ($N = 75$) comprising university students and office workers. Firestone et al.[12] replicated these findings in a large population sample ($N = 3806$) of United States residents, while Ball et al.[11] found positive relationships between Australian women's ($N = 3610$) perceptions of the descriptive exercising and walking norms of their acquaintances and neighbors and their own self-reported leisure time physical activity and volitional walking. However, these studies share the weaknesses of relying on cross-sectional designs and self-report methods to examine the relationship between norms and physical activity.

A handful of studies have involved experimental manipulations of descriptive physical activity norms, allowing their effect to be examined. In a large experimental survey study of young adults ($N = 1200$), van Bavel et al.[13] told participants in the experimental conditions either that 79% of people of the same age, gender, and country of residence as them do more than (high norm condition), or do not do more than (low norm condition), 30 minutes of physical activity on most days of the week. Again, results were promising; compared to participants in a control group (who received no norm manipulation), participants in *both* the high and low norm conditions subsequently reported stronger intentions to engage in regular physical activity the following week.

However, studies testing whether descriptive norm messages can improve people's actual physical activity behaviors have found mixed results. Priebe and Spink[9] found that emails encouraging office workers to be more active because other office workers were active (ie, because of descriptive norms) led to a greater increase in participants' self-reported mild physical activity than emails that encouraged participants to be more active for health reasons, appearance reasons, or no reason. However, participants in the descriptive norm condition did not report increasing their moderate, strenuous, or overall physical activity over the two-week study period to a greater extent than participants in the other conditions. Moreover, in a second study—where the researchers repeated the intervention with a sample of university students—there were no differences between conditions in pre- to post-manipulation changes for any form of physical activity. Similarly, Priebe and Spink[7] found that descriptive norm messages informing office workers ($N = 96$) that other office workers stand up from their desk, walk around the office, and use the stairs more often than oneself (ostensibly based on information derived from a questionnaire) reduced participants' self-reported longest period of sitting time and increased the number of times participants reported walking around the office and using the stairs during the day. However, the messages had no effect on participants' standing time.

Finally, Wally and Cameron[10] examined the effect of descriptive norm messages on university students' pedometer-assessed step counts. The researchers found that participants who received daily information about the average number of steps other participants had taken achieved marginally, but not significantly, greater step counts than participants in a control group (who received no feedback) in the final two days of the eight-day testing period. In this study, the norm information was accurate such that participants were informed of the number of steps their fellow participants had actually averaged (even if this figure was less than their step count). Interestingly,

there was no evidence that participants who reported above average physical activity levels at baseline (and were thus more likely to receive messages informing them that most other participants had taken fewer steps than them) decreased their own activity as a result. It is possible, however, that this aspect of the study's design led to an underestimation of the degree of influence of the descriptive norm on physical activity.

1.2 | The important role of the normative referent

More generally, one potential reason for the inconsistent findings from previous studies is their insufficient attention to the normative referent. Specifically, it is possible that the influence of descriptive norms on physical activity behavior has, at times, been underestimated because studies have failed to measure or target norms pertaining to individuals or groups that are most relevant, or that exert the greatest influence. Along these lines, research in other contexts (eg, healthy eating, alcohol consumption) has found evidence that normative influence is greater (and in some instances only occurs) when the normative referent is a group that the target categorizes themselves—and possesses a *strong social identity*—as a member of.[14,15] This evidence for a norm × social identification interaction is consistent with social identity theorizing. Specifically, the social identity approach[16,17] recognizes people's capacity to define themselves in terms of both their personal identity (as "I" and "me") and their social identities (as a member of the groups to which they belong). Following on from this, one of the approach's key propositions is that, to the extent that a person possesses a strong social identity as a member of a particular group, they will be motivated to align their personal behaviors with those of representative group members.[17]

In physical activity contexts, researchers have found consistent evidence that the strength of people's social identification as a member of their physical activity groups is associated with their greater participation in group sessions.[18-20] This finding aligns with our argument because engaging in regular physical activity is normative in many physical activity groups. However, only one study to date has directly examined whether norms and social identification *interact* to shape physical activity. In a sample of 133 university students, Terry and Hogg[21] found that the (cross-sectional) relationship between friendship group physical activity norms and participants' intentions to regularly exercise was moderated by the strength of participants' social identification as a member of their friendship group, such that it was only significant among strong identifiers. However, this interaction did not predict participants' self-reported exercise behaviors two weeks later. Terry and Hogg's[21] findings thus provide only limited evidence that descriptive norms and social identification interact to shape physical activity behaviors.

1.3 | The present research

The present study extended previous research in several ways. Notably, using an experimental design, it sought to provide (a) a rigorous first test of whether descriptive norm information can have an immediate effect on people's physical activity behaviors, and (b) the most rigorous test to date of whether descriptive norms and social identification interact to affect people's physical activity behaviors. Moreover, in contrast to previous physical activity descriptive norms research, we focused on participants' behavior during a physically demanding exercise task (ie, rather than on their overall physical activity), and assessed this objectively. Specifically, participants completed a baseline trial of the task before we randomly allocated them to either a high or low norm condition (where they were told that their Trial 1 performance was either below or above the average of ingroup members). In line with evidence that descriptive norms are related to, and may shape, people's physical activity behaviors during daily life,[7-12] we expected that participants who were told that their performance on the first trial of our task was below (compared to above) the norm would be more motivated to positively change their behavior during their second trial to conform to group norms. On this basis, our first hypothesis was that participants who were told that their baseline performance was below the norm would demonstrate greater improvements in their performance (H1a) and report greater increases in their exertion (H1b) from Trial 1 to Trial 2 than participants who were told that their baseline performance was above the norm. Our second hypothesis was that the effect of the descriptive norm manipulation on participants' performance (H2a) and exertion (H2b) would be moderated by participants' social identification as a member of the reference group, such that those who identified more strongly as a member of the group would demonstrate greater norm convergence.

2 | METHODS

2.1 | Participants

We recruited a sample of 82 recreationally active students from an Australian university to complete an experiment with a mixed (within-and-between) 2 × 2 design. Participants were recruited through online forums

TABLE 1 Participant anthropometric data

	High norm condition ($N = 38$)	Low norm condition ($N = 40$)	Total ($N = 78$)
Age (years)	20.71 ± 2.25 (18–30)	21.42 ± 2.80 (18–31)	21.08 ± 2.56 (18–31)
Height (cm)	167.19 ± 9.19 (145–190)	168.82 ± 9.58 (150–188)	168.04 ± 9.37 (145–190)
Weight (kg)	62.13 ± 10.25 (40–89)	62.60 ± 12.36 (40–95)	62.37 ± 11.31 (40–95)
Body Mass Index	21.97 ± 2.66 (15.57–28.73)	21.79 ± 3.82 (15.43–37.58)	21.88 ± 3.30 (15.43–37.58)

Note: Age, height, weight, and BMI data are presented as Mean ± SD (range).

and media distributions (eg, emails, university student Facebook pages; $N = 48$), and via the psychology school's research participation scheme ($N = 34$). The former group received AUD$40 for taking part, while the latter group received course credit. Four participants failed the manipulation check (see below), leaving a final sample of 78 participants (48 females, 29 males, and 1 who identified as other; $M_{age} = 21.08$, $SD = 2.56$). Further characteristics of the final sample are presented in Table 1.

2.2 | Power analysis

The lack of experimental research assessing the effect of descriptive norms on physical activity behavior meant there was limited guidance available to determine the appropriate effect size for our power analysis. Most notably, however, Priebe and Spink[7] reported medium to large effect sizes equivalent to f's ranging from 0.25 to 0.60 for the main effect of their descriptive norm manipulation on physical activity outcomes (eg, walking, using the stairs). With regard to power, the range of values researchers use also varies considerably: 80% power is a common benchmark[22] yet lower values and values up to 95% are sometimes used. The wide range of possible effect sizes and power values that we could have used meant that the range of sample sizes that power analyses indicated were required (depending on the values we inputted) was also considerable. For example, using the largest effect size from Priebe and Spink's[7] study (equivalent to f = 0.60) and power of 0.8, an *a priori* power analysis (G*Power[23]) indicated that a sample of just 26 would be required to analyze the two-way interaction between time (Trial 1 – Trial 2) and condition (high norm versus low norm) that would be the focus for analyses pertaining to our primary hypothesis (H1). On the contrary, using the smallest effect size from Priebe and Spink's[7] study (equivalent to f = 0.25) and power of 0.95, power analyses indicated that we would require 212 participants. Practical limitations on available resources meant that obtaining a sample toward the upper end of this range was not possible; we thus targeted a sample of approximately 80—the maximum our available resources permitted.[24] Given the possibility that the "true" effects may be smaller than those that this sample size is able to detect with a high level of power, it is important to note that results should be interpreted with a degree of caution, and that replication studies with larger sample sizes would be valuable.

2.3 | Procedure

Ethical approval for the study procedures was obtained from the Human Research Ethics Committee at the first author's institution (#2019/139). The study's hypotheses, design, and analysis strategy were pre-registered prior to data collection—see: http://aspredicted.org/blind.php?x=ks95k6[1]

Advertisements stated that the study was examining healthy lifestyles in young adults and noted that it involved completing a running task and some related questions twice, on two separate days. Upon registering their interest in the study, potential participants were sent a health-screening questionnaire (an adapted version of the Adult Exercise Pre-Screening System[25]) via email. Individuals who noted any contraindications to strenuous exercise on this questionnaire (eg, a disease, or signs of a disease, which may have increased their risk of harm during the testing sessions) were not eligible to participate.

All testing sessions took place in the same sports hall, which participants attended individually on both occasions (ie, the experimenter was the only other person present). To (a) ensure participants had fully recovered from their first session, and (b) minimize potential changes in participants' fitness levels between sessions, participants completed their two sessions one to two weeks apart. To further reduce the potential impact of extraneous factors on participants' task performance, participants were also asked to refrain from (a) consuming caffeine in the two hours prior to the sessions and (b) engaging in physical activity outside of their normal routine in the two days prior to the sessions.[1] Upon arriving for their first session, the experimenter explained the specific running task that participants would be required to complete (see below).

[1]We made an error in our original power calculations and thus the power analyses presented in the pre-registration document should be disregarded.

Participants gave written informed consent based on this information.

The structure of the two sessions was identical, except participants (a) completed a demographic questionnaire at the beginning of Session 1 only, (b) had their height and weight measured at the end of Session 1 only, and (c) completed the identity salience tasks and were exposed to the experimental manipulation at the beginning of Session 2 only (see details below). Otherwise, in both sessions participants were given time to warm up, then started the test when they indicated that they were ready. After a brief period of recovery (ie, 30–60 seconds), they were asked to complete a short questionnaire.

2.4 | The running task

We used a task widely used to assess aerobic fitness—the Multi-Stage Fitness or "beep" Test.[26] This involves running between two points in space 20 meters apart, in time with electronic beeps played through a speaker. In the present instance, cones placed along two lines on the ground indicated these points. The time between beeps decreases as the test progresses (ie, each level), meaning participants are required to run at an accelerating pace. The test ends when the participant either (a) decides to stop, or (b) fails to reach the endpoints prior to the corresponding beep on two consecutive occasions. The primary instruction participants were given prior to commencing their first trial was to continue the test for "as long as you can." No verbal encouragement was given during the test.

This task was chosen for three primary reasons. First, improving one's performance on the Multi-Stage Fitness Test requires motivation (ie, to more wholeheartedly exert oneself). Examining changes in performance, and in the extent to which participants perceived they had exerted themselves, from Trial 1 to Trial 2 thus provided both objective and subjective indicators of how the norm manipulation affected participants' motivation to positively change their physical activity behavior (a key aim of the study). At the same time, second, this task afforded a highly conservative test of our hypotheses because performance on the Multistage-Fitness Test is limited by physical fitness. Improving from a baseline (where, in this instance, participants were explicitly instructed to do the best they could) thus requires considerable motivation and effort, and improvements are likely to be small even when a person is highly motivated to improve. Finally, the task's simplicity (eg, in comparison with skill-based tasks or even a time trial) mitigated the potential for improvements from Trial 1 to Trial 2 to occur as a result of practice or learning effects.

2.5 | Identity salience tasks and experimental manipulation

Participants were randomly allocated to either the high or low norm experimental condition using a random number generator. The first part of Session 2 was standardized across the two conditions such that all participants began by completing tasks designed to make their social identity as a member of the norm reference group salient. Participants' social identity as a student at their university was chosen because all participants shared this social identity. First, participants were asked to respond to three negative and three positive statements about being a student at their university by selecting statements they agreed with and leaving blank statements they disagreed with. The negative statements (eg, "I feel no affiliation with X university students") were designed to be easy to disagree with, and the positive statements (eg, "In general, I like being an X university student") were designed to be easy to agree with. Previous studies have shown that that this task is an effective way to prime a given social identity (eg, such that it increases social identification with the group in question[27]). Second, participants completed a 14-item measure of their social identification as a student at their university.[28] This also provided the focal measure of social identification for the analyses pertaining to H2a and H2b (see below).

After completing these tasks, participants were told that they would be given feedback on their Trial 1 performance and presented with a feedback sheet (see Figure 1). This comprised supposed normative information about "the results of the study so far." Specifically, this information was presented in the form of a bell curve on which the "average" score for other students of the participant's gender (a detail included to enhance believability) and the participant's own score from the first session were marked with a cross and indicated with a numeric value. The placement of the cross varied by condition. In the high norm condition, the cross was placed to the left of (ie, below) the average, while in the low norm condition it was placed to the right of (ie, above) the average. The number indicated for the average was half a level higher or lower than the participant's own score from the first session. This half a level difference was chosen because (following a pilot) this was considered to be a challenging but attainable degree of improvement for participants in the high norm condition. To reinforce the manipulation, and ensure participants understood the information that had been presented, the feedback was also verbally communicated by the experimenter, who explained that the average they were looking at was for male/female students, stated the average, and told the participant that they were "around half a level" above or below the average. The session then proceeded

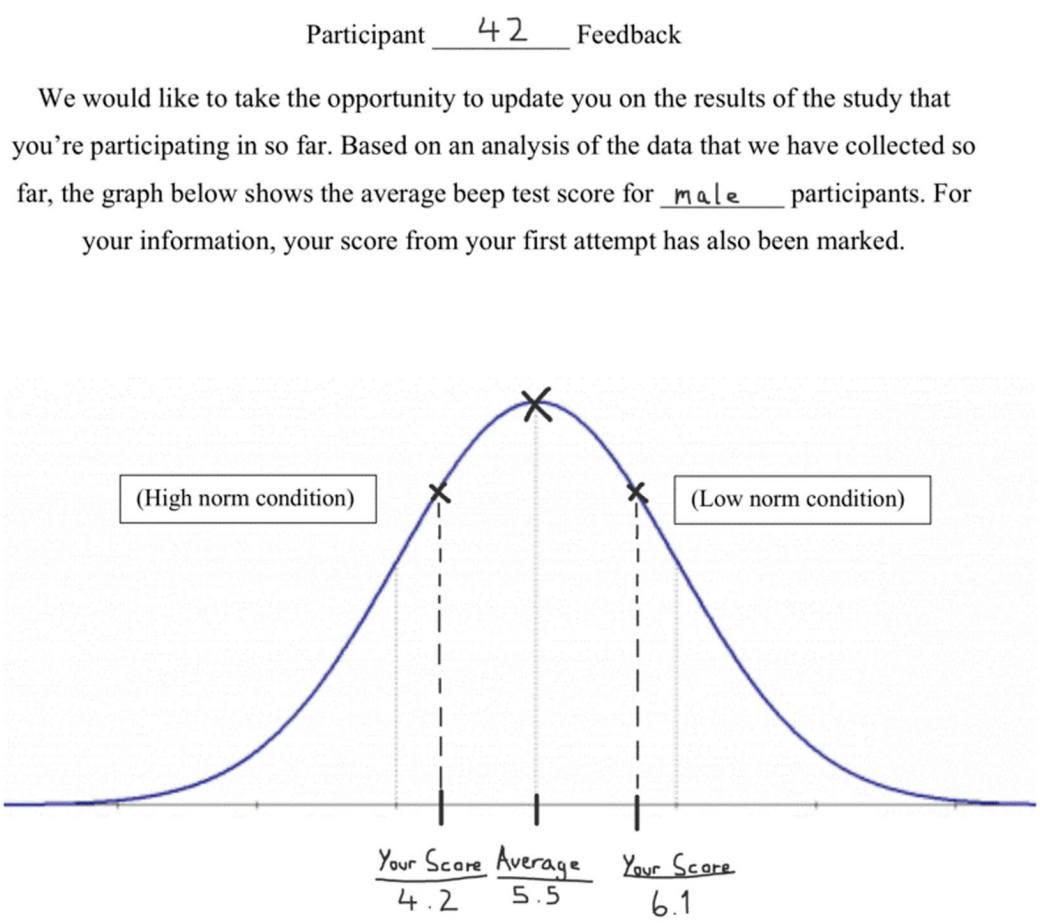

FIGURE 1 Example of the feedback sheet used for the experimental manipulation

identically to Session 1, such that participants next proceeded to warm up before undertaking their second trial of the running task and completing the short post-trial questionnaire. At the end of Session, 2 participants were fully debriefed verbally and in writing.

2.6 | Measures

2.6.1 | Social identification

Participants' social identification as a student at their university was measured using Leach et al.'s[28] 14-item scale (eg, "I feel a bond with X university students"). This measure was chosen because, in addition to being well-validated, it includes a greater number of items than many other widely used scales.[29,30] As such, it required participants to consider their university student social identity in more detail, and for longer, than would have been the case had they completed an alternative scale (thus aiding our efforts to make this identity salient, see above). The scale demonstrated good internal reliability in the present sample ($\alpha = 0.91$).

2.6.2 | Performance

Participants' Multi-Stage Fitness Test scores provided an objective measure of their task performance (and thus behavior). In line with recommendations, scores (ie, the level participants reached) were converted to estimated VO_2 max values using standard conversion tables.[31] This conversion is necessary because it helps account for the fact that the difference between Multi-Stage Fitness test levels is not equal: a greater degree of effort and fitness is required to progress through additional levels as the test progresses (and thus, eg, the performance difference between two participants who reached levels one and two would not have been equivalent to the performance difference between two participants who reached levels seven and eight).

2.6.3 | Exertion

As noted above, greater motivation to exert oneself during the Multi-Stage Fitness Test may not always translate into an (considerably) improved score because

performance on the test is limited by physical fitness. Participants were therefore asked to complete the Rating of Perceived Exertion (RPE) scale[32] following each trial. This provided a measure of the extent to which participants perceived they had exerted themselves during the two trials (ie, how they had behaved), and thus how this changed from pre- to post-manipulation. The RPE scale is typically used to measure perceived exertion during physical tasks. However, it can also provide a valid measure of exertion when used retrospectively.[33,34] Participants were asked to indicate their maximal level of perceived exertion during the task immediately following its completion (ie, after a recovery period of 30–60 seconds). The 15-point scale ranges from 6 "no exertion at all" to 20 "maximal exertion."

2.6.4 | Manipulation and suspicion checks

After completing their second running task and the perceived exertion measure, participants were asked to indicate how their score from the first session compared to the average (from four response options: "I was below average"; "I was at the average"; "I was above average"; "I don't remember"). A suspicion check—which consisted of a free response box in which participants were asked to indicate their thoughts on the main purpose of the experiment—was also included.

3 | RESULTS

3.1 | Preliminary analyses

All analyses were conducted in IBM SPSS version 27. As noted above, four participants failed the manipulation check (ie, they did not correctly identify how their Trial 1 score compared to the average) and were thus excluded from the analysis in accordance with our preregistered analysis plan. No participants correctly identified the purpose of the study, and thus, none were excluded on the basis of the suspicion check. Of the remaining 78 participants, five had missing values on at least one of our key variables. To retain these participants in the dataset, and thus preserve the power of our sample, we conducted multiple imputation (a popular approach to handling missing data that has been shown to outperform several alternative approaches[35,36]). The Bar Procedure[37] was subsequently used to create pooled estimates suitable for traditional analytic methods. Independent *t* tests revealed no significant differences between participants in the high and low norm conditions for Trial 1 performance ($t(76) = -0.34$, $p = 0.74$, $d = 0.08$) or Trial 1 exertion ($t(76) = -0.59$, $p = 0.55$, $d = 0.14$).

3.2 | Main analyses

Participants' performance scores demonstrated high stability ($r = 0.94$, $p < 0.001$) between Trial 1 ($M = 32.21$, $SD = 8.59$, range = 20.4–60.9) and Trial 2 ($M = 34.03$, $SD = 8.82$, range = 20.0–64.55). Participants' exertion scores also demonstrated high stability ($r = 0.74$, $p < 0.001$) between Trial 1 ($M = 15.27$, $SD = 2.19$, range = 6–19) and Trial 2 ($M = 15.50$, $SD = 2.30$, range = 6–19.32). On average, participants identified relatively strongly as a student at their university (on seven-point scales, the mean per item was 4.86, $SD = 0.82$, range = 1–6.57).

3.2.1 | H1—The effect of the norm manipulation on performance and exertion

To assess H1a and H1b, we conducted two-way repeated measures ANOVAs, with condition (high norm versus low norm) as a between-subjects factor and time (Trial 1 – Trial 2) as a within-subjects factor. For performance, there was a significant main effect for time such that participants in both conditions tended to improve their performance from Trial 1 to Trial 2: $F(1,76) = 28.00$, $p < 0.001$, $\eta_p^2 = 0.269$ ($M_{improvement} = 1.83$, $SD = 3.14$). Supporting H1a, there was also a significant condition × time interaction: $F(1,76) = 4.47$, $p = 0.038$, $\eta_p^2 = 0.056$, such that the performance scores of participants in the high norm condition improved from Trial 1 to Trial 2 to a significantly greater extent than the performance scores of participants in the low norm condition (high norm condition $M_{improvement} = 2.58$, $SD = 2.54$; low norm condition $M_{improvement} = 1.11$, $SD = 3.51$; see Figure 2). Participants in the high norm condition improved their performance 2.5 times as much as participants in the low norm condition. For exertion, the main effect for time was not significant: $F(1,76) = 1.63$, $p = 0.206$, $\eta_p^2 = 0.021$. The condition × time interaction was also not significant ($F(1,76) = 0.58$, $p = 0.449$, $\eta_p^2 = 0.008$), and thus, H1b was not supported.

3.2.2 | H2—The moderating role of social identification

To assess H2a and H2b, we first repeated the ANOVA analyses with the addition of social identification as a between-subjects continuous moderator. For performance, the condition × social identification interaction

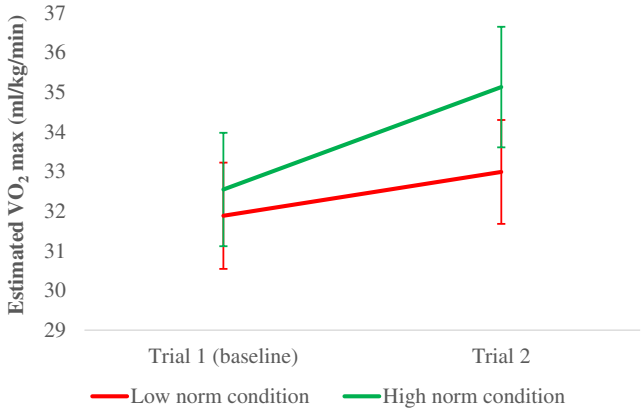

FIGURE 2 Participants who were told that their running task performance in Trial 1 was below average (high norm condition) improved their performance to a significantly greater extent than participants who were told that their Trial 1 performance was above average (low norm condition)

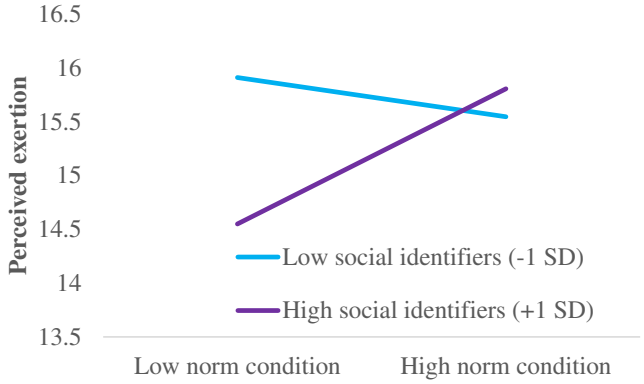

FIGURE 3 Participants' perceived exertion during the running task was affected by the feedback they received about their group's normative performance, but only among high identifiers with the group. Note that this figure shows estimated marginal means from a regression analysis using PROCESS[38]

was non-significant, $F(1,74) = 2.07$, $p = 0.155$, $\eta_p^2 = 0.037$, and thus, H2a was not supported. However, for exertion the condition × social identification interaction was significant: $F(1,74) = 6.43$, $p = 0.013$, $\eta_p^2 = 0.080$ (supporting H2b).

We subsequently used PROCESS[38] (Model 1) to model the simple slope analyses for exertion. Specifically, we tested whether participants' social identification as a member of the norm reference group moderated the effect of the norm manipulation on participants' Trial 2 exertion. Trial 1 exertion was included in the model as a covariate. The condition × social identification interaction was significant: $t(78) = 2.13$, $p = 0.036$, B = 0.99, SE = 0.465 (supporting H2b). This was such that the effect of the norm manipulation on perceived exertion was non-significant among those who were relatively low in their social identification with the referent group (−1 SD: B = −0.37, SE = 0.50, $p = 0.466$), but significant among those who strongly identified as a member of the referent group (+1 SD: B = 1.25, SE = 0.53, $p = 0.022$). As Figure 3 shows, strong identifiers tended to exert themselves less than weak identifiers during Trial 2 if they were in the low norm condition, and slightly more than weak identifiers if they were in the high norm condition.

4 | DISCUSSION

This study examined the effect of a descriptive norm manipulation on participants' performance and exertion during an exercise task, and whether the strength of participants' social identification as a member of the norm reference group moderated these effects. We found evidence that participants' task performance (an objective operationalisation of their physical activity behavior) tended to improve from Trial 1 to Trial 2. Supporting H1a, this improvement was 2.5 times greater among participants who were told that their performance was below the norm, compared to above the norm. The effect of the norm manipulation on performance was not moderated by social identification (inconsistent with H2a). There was no main effect of norm condition on participants' perceived exertion (a subjective indicator of their physical activity behavior), contrary to H1b. However, providing initial support for the process of norm convergence specified in H2b, the effect of the norm manipulation on exertion was significant among those who strongly identified as a member of the norm reference group but non-significant among those whose identification as a member of the norm reference group was (relatively) weak. The significant moderation effect observed was such that strong identifiers reported exerting themselves less than weak identifiers during Trial 2 if they were in the low norm condition, and slightly more than weak identifiers if they were in the high norm condition.

These findings have a range of implications and extend previous research in important ways. First, findings extend our understanding of the effects that descriptive norms can have in physical activity contexts. Previous research has pointed to a positive relationship between descriptive physical activity norms and physical activity engagement,[8,11,12] and tentatively indicated that norm-based messaging might be an effective way to increase people's physical activity during daily life.[7,10] Our findings are the first to show that descriptive norm messages can have an immediate positive effect on people's behavior during a physically demanding exercise task. Indeed, our confidence in the robustness of this effect is strengthened

because the task on which it was demonstrated—the Multi-Stage Fitness Test—provides an (objective) assessment of aerobic fitness. The capacity for participants to improve from a proximal baseline was therefore limited (even if they were strongly motivated to do so).

It is also valuable to consider our findings in conjunction with those from previous research in which manipulations of different forms have been used as part of efforts to increase or reduce people's self-efficacy and assess the downstream effects of this (eg, during handgrip and treadmill running tasks). Here, studies have found that telling participants their performance on an initial task placed them in the bottom 10th or 20th percentiles relative to norms (eg, for people of the same age and gender) reduced their self-efficacy and led to negative outcomes (eg, reduced enjoyment, less sustained effort, and increased state anxiety) before, during, and after similar subsequent tasks.[39,40] By contrast, telling participants their initial task performance placed them in the top 10th or 20th percentile relative to norms increased participants' self-efficacy and facilitated positive outcomes in these studies. This perhaps helps explain why participants in our low norm condition (who were told that their Trial 1 performance was above average) tended to perform marginally better in Trial 2 than Trial 1; the norm manipulation may have increased their self-efficacy—a "boost" that was independent of norm convergence effects. Viewing our findings in conjunction with those from self-efficacy research also provides suggestive evidence that the way descriptive norm messages are *framed* and the perceived *attainability* of the norm that they convey, may play a key role in determining whether such messages positively or negatively impact that person's subsequent behavior. The "low self-efficacy" manipulations used in previous research[39,40] did not give participants a "goal" to aim at; they simply emphasized that the participant's performance was among the worst for members of the norm reference group. On the contrary, we ensured that the normative score was clear to participants and deliberately set this at a challenging but attainable level, relative to participants' own performance. The seemingly contrasting effects of these two approaches suggest that nuances surrounding how a norm is presented, and the content it includes, should be carefully considered by those seeking to positively harness the power of normative influence.

Also on a practical note, our findings point to the potential value of using descriptive norm messages in contexts where motivating immediate improvements in people's physical activity behaviors is a goal. For example, fitness instructors and personal trainers could include aspirational (but achievable) norms as part of their communications with exercisers (eg, "most people do X repetitions of this exercise in a minute"). This could help increase the amount of effort people exert during exercise sessions—a key determinant of the physiological benefits they derive from the exercise, and whether they ultimately achieve their exercise goals.[5] Similarly, in performance-focused contexts (eg, elite team sport) coaches might be able to increase players' motivation to work harder through messages such as "most other players in the team have covered more distance than you" (eg, during half-time breaks). Indeed, at elite level there is now often a wealth of player data readily available (eg, pertaining to kilometers run and number of sprints completed) that could be shown to players to support such messages. Of course, normative influence is likely to already be prevalent in many team sport environments (eg, whereby players' effort is shaped by their perceptions of other team members' effort). Our findings indicate the potential value of overt attempts to harness the power of this normative influence.

Finally, with regard to the findings related to H2, we found some evidence that normative influence was affected by social identification. The interaction between the manipulation and group identification did not significantly predict performance but did significantly predict exertion. Here, it was particularly noticeable that high identifiers in the low norm condition exerted themselves less than their low identifying counterparts, with smaller differences as a function of identification apparent in the high norm condition (see Figure 3). This finding may, at least in part, be because the demanding nature of the task created a ceiling effect. That is, in the low norm condition highly identified participants could readily exert themselves less than lowly identified participants. However, in the high norm condition, highly identified participants may have been more limited in their capacity to exert themselves more than their lowly identified counterparts because all participants, including low identifiers, exerted themselves to a high degree. It would thus be valuable for future research to explore whether social identification is a stronger moderator when participants' capacity to improve on the focal outcome is greater (eg, on non-exhaustive exercise tasks). This would shed further light on the extent to which social identification affects the relationship between descriptive norms and exercise behavior. More broadly, it is also important to note that our sample size means caution should be applied in interpreting the findings of our moderation analyses in particular. It is possible that the "true" effects may be smaller than those observed here; attempts to replicate these effects would therefore be valuable.

4.1 | Limitations and future research

This study had several strengths. Notably, its repeated measures experimental design afforded a rigorous test

of the potential for descriptive norm messages to cause *changes* in physical activity behaviors—that is, while controlling for other potential influences on behavior (eg, physical fitness, gender). Moreover, the task we used enabled us to (a) assess these changes objectively, and (b) mitigate the potential influence of practice or learning effects. At the same time, however, these features of the study also limited its ecological validity. Further studies testing the effects of descriptive norm messages on a wider range of objectively assessed physical activity behaviors in real-world contexts are therefore needed. As alluded to above, there are both sport and exercise contexts in which such research could be conducted. For example, research could assess the impact of descriptive norm messages on sports team players' effort and performance during training or matches, or on exercisers' effort during workouts. It would also be valuable for such research to examine how long the effects of these messages persist. This would help shed further light on the magnitude of the associated benefits. For instance, if descriptive norm messages—either delivered once or reinforced on multiple occasions—were found to have a lasting effect on people's motivations and behaviors (eg, such that people tend to maintain their motivation to exert themselves more or perform better until they reach the norm), then this would provide particularly compelling evidence for their use in sport and exercise contexts.

5 | CONCLUSION

This study provides initial evidence that descriptive norm messages can have an immediate effect on people's behavior during a physically demanding exercise task. Findings further suggest that the effect of such messages might be greater if the target possesses a strong social identity as a member of the norm reference group. Together, findings therefore indicate that descriptive norm messages could be fruitfully deployed in sport and exercise contexts with a view to improving people's performance- and health-related behavior, and offer guidance to maximize the effect, and thus benefit, of such efforts.

6 | PERSPECTIVE

Previous research has found a positive relationship between descriptive norms and people's physical activity behaviors,[8,11,12] and provided some evidence that descriptive norm messages can be effective in improving people's physical activity behaviors during daily life.[7,10] This study provides the first experimental evidence that descriptive norm messages can have an immediate positive effect on people's physical activity behavior during a demanding exercise task. More specifically, findings suggest that using such messages might be an effective way for sport and exercise practitioners (eg, sports team coaches, fitness instructors, and personal trainers) to increase people's motivation to exert more effort and perform better. We also found evidence that the strength of people's social identification as a member of the norm reference group may affect the extent to which norm messages influence their physical activity behaviors, suggesting that this could be an important factor to consider for those seeking to harness normative influence.

ACKNOWLEDGEMENTS

An Australian National University Futures Scheme grant supported this research. The third author's contribution was supported by an NHMRC Emerging Leadership Fellowship (#1173270). The funders had no role in the development of the research questions, analysis, or decision to publish.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Mark Stevens https://orcid.org/0000-0002-7908-3624

REFERENCES

1. Stevens M, Rees T, Steffens NK, Haslam SA, Coffee P, Polman R. Leaders' creation of shared identity impacts group members' effort and performance: evidence from an exercise task. *PLoS One*. 2019;14(7):e0218984.
2. De Cuyper B, Boen F, Van Beirendonck C, Vanbeselaere N, Fransen K. When do elite cyclists go the extra mile? Team identification mediates the relationship between perceived leadership qualities of the captain and social laboring. *Int J Sport Psychol*. 2016;47(4):355-372.
3. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. *Am J Public Health*. 2011;101(4):751-758.
4. Kahn EB, Ramsey LT, Brownson RC, et al. The effectiveness of interventions to increase physical activity: a systematic review. *Am J Prev Med*. 2002;22(4):73-107.
5. Fisher J, Steele J. Questioning the resistance/aerobic training dichotomy: a commentary on physiological adaptations determined by effort rather than exercise modality. *J Hum Kinet*. 2014;44(1):137-142.
6. Cialdini RB, Reno RR, Kallgren CA. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *J Pers Soc Psychol*. 1990;58(6):1015-1026.
7. Priebe CS, Spink KS. Less sitting and more moving in the office: using descriptive norm messages to decrease sedentary behavior and increase light physical activity at work. *Psychol Sport Exerc*. 2015;19:76-84.
8. Priebe CS, Spink KS. When in Rome: descriptive norms and physical activity. *Psychol of Sport Exerc*. 2011;12(2):93-98.

9. Priebe CS, Spink KS. Using messages promoting descriptive norms to increase physical activity. *Health Commun*. 2012;27(3):284-291.
10. Wally CM, Cameron LD. A randomized-controlled trial of social norm interventions to increase physical activity. *Ann Behav Med*. 2017;51(5):642-651.
11. Ball K, Jeffery RW, Abbott G, McNaughton SA, Crawford D. Is healthy behavior contagious: associations of social norms with physical activity and healthy eating. *Int J Behav Nutri Phys Act*. 2010;7(1):86.
12. Firestone MJ, Yi SS, Bartley KF, Eisenhower DL. Perceptions and the role of group exercise among New York City adults, 2010–2011: an examination of interpersonal factors and leisure-time physical activity. *Prev Med*. 2015;72:50-55.
13. Van Bavel R, Esposito G, Baranowski T. Is anybody doing it? An experimental study of the effect of normative messages on intention to do physical activity. *BMC Pub Health*. 2014;14(1):778.
14. Liu J, Thomas JM, Higgs S. The relationship between social identity, descriptive social norms and eating intentions and behaviors. *J Exp Soc Psychol*. 2019;82:217-230.
15. Neighbors C, LaBrie JW, Hummer JF, et al. Group identification as a moderator of the relationship between perceived social norms and alcohol consumption. *Psychol Addict Behav*. 2010;24(3):522-528.
16. Tajfel H, Turner JC. An integrative theory of intergroup conflict. In: Austin WG, Worchel S, eds. *The Social Psychology of Intergroup Relations*. Brooks/Cole; 1979:33-47.
17. Turner JC, Hogg MA, Oakes PJ, Reicher SD, Wetherell MS. *Rediscovering the Social Group: A Self-Categorization Theory*. Blackwell; 1987.
18. Stevens M, Rees T, Polman R. Social identification, exercise participation, and positive exercise experiences: Evidence from parkrun. *J Sports Sci*. 2019;37(2):221-228.
19. Stevens M, Rees T, Coffee P, Steffens NK, Haslam SA, Polman R. Leading 'us' to be active: A two-wave test of relationships between identity leadership, group identification, and attendance. *Sport Exerc Perform Psychol*. 2020;9(1):128-142.
20. Stevens M, Rees T, Coffee P, Haslam SA, Steffens NK, Polman R. Leaders promote attendance in sport and exercise sessions by fostering social identity. *Scand J Med Sci Sports*. 2018;28(9):2100-2108.
21. Terry DJ, Hogg MA. Group norms and the attitude-behavior relationship: A role for group identification. *Article. Pers Soc Psychol Bull*. 1996;22(8):776-793.
22. Perugini M, Gallucci M, Costantini G. A practical primer to power analysis for simple experimental designs. *Int Rev Soc Psychol*. 2018;31(1):1-23.
23. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 2009;41(4):1149-1160.
24. Lakens D. Sample size justification. *PsyArXiv*. 2021; https://doi.org/10.31234/osf.io/9d3yf.
25. Fitness Australia. Adult pre-exercise screening system (APSS). Accessed 6 April 2021, https://bp-fitnessaustralia-production.s3.amazonaws.com/uploads/uploaded_file/file/386441/ADULT_PRE-EXERCISE_SCREENING_SYSTEM__APSS__2019FINALv2.pdf
26. Léger LA, Mercier D, Gadoury C, Lambert J. The multistage 20 metre shuttle run test for aerobic fitness. *J Sports Sci*. 1988;6(2):93-101.
27. Greenaway KH, Haslam SA, Cruwys T, Branscombe NR, Ysseldyk R, Heldreth C. From, "we" to "me": Group identification enhances perceived personal control with consequences for health and well-being. *J Pers Soc Psychol*. 2015;109(1):53-74.
28. Leach CW, van Zomeren M, Zebel S, et al. Group-level self-definition and self-investment: a hierarchical (multicomponent) model of in-group identification. *J Pers Soc Psychol*. 2008;95(1):144-165.
29. Postmes T, Haslam SA, Jans L. A single-item measure of social identification: reliability, validity, and utility. *Brit J Soc Psychol*. 2013;52(4):597-617.
30. Doosje B, Ellemers N, Spears R. Perceived intragroup variability as a function of group status and identification. *J Exp Soc Psychol*. 1995;31(5):410-436.
31. Ramsbottom R, Brewer J, Williams C. A progressive shuttle run test to estimate maximal oxygen uptake. *Brit J Sports Med*. 1988;22(4):141-144.
32. Borg G. Perceived exertion as an indicator of somatic stress. *Scand J Rehabil Med*. 1970;2(2):92-98.
33. Graupensperger S, Gottschall JS, Benson AJ, Eys M, Hastings B, Evans MB. Perceptions of groupness during fitness classes positively predict recalled perceptions of exertion, enjoyment, and affective valence: an intensive longitudinal investigation. *Sport Exerc Perform Psychol*. 2019;8(3):290-304.
34. Foster C, Florhaug JA, Franklin J, et al. A new approach to monitoring exercise training. *J Strength Cond Res*. 2001;15(1):109-115.
35. Hayati Rezvan P, Lee KJ, Simpson JA. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. *BMC Med Res Methodol*. 2015;15(1). https://doi.org/10.1186/s12874-015-0022-1
36. Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC Med Res Methodol*. 2006;6(1):57.
37. Baranzini D The "Bar Procedure": single data-file compressing SPSS multiply imputed files. 2018; https://doi.org/10.13140/RG.2.2.33750.70722
38. Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-based Approach*, 2nd ed. The Guildford Press; 2017.
39. Marquez DX, Jerome GJ, McAuley E, Snook EM, Canaklisova S. Self-efficacy manipulation and state anxiety responses to exercise in low active women. *Psychol Health*. 2002;17(6):783-791.
40. Hutchinson JC, Sherman T, Martinovic N, Tenenbaum G. The effect of manipulated self-efficacy on perceived and sustained effort. *J Appl Sport Psychol*. 2008;20(4):457-472.

How to cite this article: Stevens M, Fitzpatrick Á, Cruwys T. The effect of descriptive norms and social identification on performance and exertion during a physical fitness task. *Scand J Med Sci Sports*. 2022;32:313–323. https://doi.org/10.1111/sms.14072