Original Paper

The Feasibility of Using Instagram Data to Predict Exercise Identity and Physical Activity Levels: Cross-sectional Observational Study

Sam Liu[1*], PhD; Megan Perdew[1*], MSc; Alexander Lithopoulos[1*], PhD; Ryan E Rhodes[1,2*], PhD

[1]School of Exercise Science, Physical and Health Education, University of Victoria, Victoria, BC, Canada
[2]Department of Psychology, University of Victoria, Victoria, BC, Canada
[*]all authors contributed equally

Corresponding Author:
Sam Liu, PhD
School of Exercise Science, Physical and Health Education
University of Victoria
PO Box 3010 STN CSC
Victoria, BC, V8W 3N4
Canada
Phone: 1 2507218392
Email: samliu@uvic.ca

Abstract

Background: Exercise identity is an important predictor for regular physical activity (PA). There is a lack of research on the potential mechanisms or antecedents of identity development. Theories of exercise identity have proposed that investment, commitment and self-referential (eg, I am an exerciser) statements, and social activation (comparison, support) may be crucial to identity development. Social media may be a potential mechanism to shape identity.

Objective: The objectives of this study were to (1) explore whether participants were willing to share their Instagram data with researchers to predict their lifestyle behaviors; (2) examine whether PA-related Instagram uses (ie, the percentage of PA-related Instagram posts, fitness-related followings, and the number of likes received on PA-related posts) were positively associated with exercise identity; and (3) evaluate whether exercise identity mediates the relationship between PA-related Instagram use and weekly PA minutes.

Methods: Participants (18-30 years old) were asked to complete a questionnaire to evaluate their current levels of exercise identity and PA. Participants' Instagram data for the past 12 months before the completion of the questionnaire were extracted and analyzed with their permission. Instagram posts related to PA in the 12 months before their assessment, the number of likes received for each PA-related post, and verified fitness- or PA-related followings by the participants were extracted and analyzed. Pearson correlation analyses were used to evaluate the relationship among exercise identity, PA, and Instagram uses. We conducted mediation analyses using the PROCESS macro modeling tool to examine whether exercise identity mediated the relationship between Instagram use variables and PA. Descriptive statistical analyses were used to compare the number of willing participants versus those who were not willing to share their Instagram data.

Results: Of the 76 participants recruited to participate, 54% (n=41) shared their Instagram data. The percentage of PA-related Instagram posts ($r=0.38$; $P=.01$) and fitness-related Instagram followings ($r=0.39$; $P=.01$) were significantly associated with exercise identity. The average number of "likes" received ($r=0.05$, $P=.75$) was not significantly associated with exercise identity. Exercise identity significantly influenced the relationship between Instagram usage metrics (ie, the percentage of PA-related Instagram posts [$P=.01$] and verified fitness-related Instagram accounts [$P=.01$]) and PA level. Exercise identity did not significantly influence the relationship between the average number of "likes" received for the PA-related Instagram posts and PA level.

Conclusions: Our results suggest that an increase in PA-related Instagram posts and fitness-related followings were associated with a greater sense of exercise identity. Higher exercise identity led to higher PA levels. Exercise identity significantly influenced the relationship between PA-related Instagram posts ($P=.01$) and fitness-related followings on PA levels ($P=.01$). These results suggest that Instagram may influence a person's exercise identity and PA levels. Future intervention studies are warranted.

XSL•FO
RenderX

Introduction

Regular physical activity (PA) is critical to prevent chronic diseases and maintain overall well-being [1]. However, physical inactivity (ie, not meeting recommended guidelines) continues to be a major public health concern worldwide [2]. In Canada, approximately 79% of adults aged 18-39 did not meet the PA guidelines (150 minutes of moderate-to-vigorous PA per week) in 2015 [3]. Thus, there is an urgent need to find innovative solutions to better target and tailor PA interventions based on a person's needs to maximize public health impact [4]. Much of the PA promotion research literature has been based in the social cognitive tradition, which is focused on increasing expectations about the benefits of PA, improving perceptions of capability, and developing goals and self-regulation tactics to enable behavior change [5-8]. The social cognitive–based behavior change theories suggest that intentions are proximal antecedents of behavior; thus, supporting positive PA intentions can lead to subsequent behavioral enactment [9]. This approach has shown some effectiveness in terms of behavior change, albeit with modest effects [5]. There is clearly a need to continue to investigate additional mechanisms that can assist in PA continuance.

One of these mechanisms may be identity. Identity represents a person's self-categorization into a particular role [10]. This process assists in developing personal rules and standards of conduct with a particular behavior, such as PA [11]. As such, the central mechanism that sustains behavior becomes regulation around the identity standard [12]. Exercise identity refers to the degree to which an individual identifies himself/herself as an "exerciser" or someone who values and regularly engages in PA [13]. Discrepancies between the person's exercise identity and the present situation (eg, being physically inactive) can create negative emotions, which are proposed to motivate the individual to close the gap between the present behavior and the identity standard [12,14]. In essence, identity is a reflexive self-regulating system of behavioral continuance/stability [10]. Exercise identity has emerged as an important construct for PA promotion because of the reciprocal relationship between role identities and behavior [15]. Self-report surveys, such as the Exercise Identity Scale, have been developed and rigorously tested for assessing an individual's exercise identity, which has allowed for further investigation of the association between exercise identity and PA behavior [16].

Observational evidence has shown that PA behavior and identity have a reliable association in the medium to large effect size range [17]. Furthermore, the relationship is often independent of more traditional social–cognitive determinants (eg, intention) [17] and the mechanisms of identity dissonance and negative affect have been well-validated in hypothetical PA abstinence situations [18-20]. What has seen less attention, however, is the potential mechanisms or antecedents of identity development. Theories of exercise identity have proposed that investment, commitment and self-referential statements (eg, I am an exerciser), and social activation (comparison, support) may be

crucial to identity development [13,21-23]. Another possible mechanism of identity is self-expression. Bem's self-perception theory [24], for example, suggests that identity may be formed largely by observing one's behavior and inferring self-categorization. This premise has some support in PA interventions from 2 small feasibility trials where participants who were encouraged to present PA-related imagery and self-expression (eg, a picture on the mantle) showed corresponding increases in identity [25,26]. However, the large focus on social media and self-expression across much of the population may be an even more effective mechanism of shaping identity.

Social media websites, such as Instagram, allow users to easily and freely communicate and express themselves by posting photos, status messages, or links that become instantly available to the public [27]. Infodemiology is a new area of study where researchers examine ways to use data generated on social media to better understand and inform public health problems in real time [28]. Infoveillance has referred to applications where infodemiology methods are utilized, particularly for surveillance purposes [29]. PA promotion interventions can potentially extract useful information from user-generated social data for infoveillance and to tailor health promotion efforts to improve PA levels. In fact, recent studies have shown the possibility of using social media and digital interventions to promote PA [30-32].

Researchers have already shown that behavior characteristics extracted from social media sites, such as Twitter, can be used to model alongside other biomedical data sets to inform PA on a population level [30-34]. We believe that it is critical to explore the use of social data from other platforms, such as Instagram, to expand the field of infodemiology. Instagram is a highly unique and user-friendly platform that allows users to share photos and videos on their mobile devices anywhere at any time. This platform has evolved into a visual-oriented culture, promoting photos first and text second. Instagram enables users to express their personalities and lifestyles with others and form relationships with users that share similar interests and values. In contrast to Twitter and Facebook, Instagram's "photo-first" culture may contribute to different user behavior and motivation [35]. Approximately 71% of online adults in America have an Instagram account [36] compared with 38% of adults (18-29 years of age) that use Twitter [37]. In Canada, 37% of adults have a social networking account, and among adults aged 18-34, about 65% have an Instagram account [38]. An Instagram study suggested that analysis of the content of Instagram pictures, color analysis (amount of saturation brightness, photo filters), frequency of posts, patterns of followings, algorithmic face detection, and the number of likes are associated with symptoms of depression [39].

However, it remains unclear whether PA-related markers extracted from Instagram are associated with a person's exercise identity and subsequent PA level. Furthermore, Twitter data are designed to be publicly available, whereas Instagram data are

not publicly available by default [40]. Users may perceive their Instagram data as being more private and sensitive [41]. The percentage of participants who are willing to share their Instagram data for health behavior monitoring is currently unclear. Thus, the objectives of this study were to (1) explore whether participants were willing to share their Instagram data with researchers to predict their lifestyle behaviors; (2) examine whether PA-related Instagram uses (ie, the percentage of PA-related Instagram posts, fitness-related followings, and the number of likes received on PA-related posts) were positively associated with exercise identity; and (3) evaluate whether exercise identity mediates the relationship between PA-related Instagram use and weekly PA minutes. Commensurate with self-perception theory [24] and past research [25,26], we hypothesized (1) more than 50% of the participants would be willing to share their Instagram data for health infoveillance, (2) PA-related Instagram use was positively associated with exercise identity, and (3) exercise identity would mediate the relationship between PA-related Instagram use (ie, number of likes) and PA minutes. Figure 1 presents a visual representation of the hypothesized mediation model.

Figure 1. Hypothesized mediation model including Instagram use, exercise identity, and physical activity level.
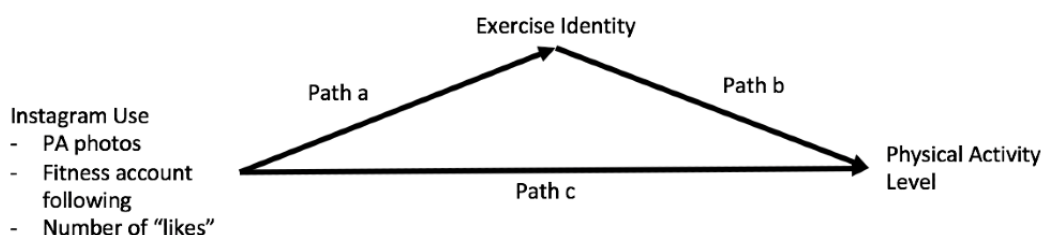
Methods

Participants and Procedure

This cross-sectional study recruited university students (n=76) from the University of Victoria, British Columbia, Canada, using research posters and online advertisements. The rolling recruitment strategy took place between June 30, 2018, and July 1, 2019. Inclusion criteria required that participants be Instagram users, and between the ages of 18 and 29. Exclusion criteria included the inability to comprehend English. Interested participants were asked to contact the research team. The research team obtained consent from all eligible participants prior to data collection. During the consent process, participants were notified of the study objectives and that the sharing of their Instagram data was voluntary. All data obtained in connection with the study remained confidential. Confidentiality was maintained by means of encrypting on a secure server. All survey data and ID numbers were used on all databases. Researchers had access to participant Instagram data for up to 1 year from the start of the study.

Eligible participants who gave consent to participate in the study were asked to complete a questionnaire to record their demographic characteristics, daily Instagram use duration, and evaluate their current exercise identity and PA levels. Participants were also asked if they are willing to share their Instagram data. Participants that refused to share their Instagram data were given the option to write down reasons for not sharing. Those participants who agreed to share their Instagram data were asked to follow an Instagram research account in order for researchers to extract their Instagram data (photos, likes, followings). Researchers analyzed the previous 12 months of Instagram data from the completion date of the exercise identity and PA questionnaires. This study received research ethics approval from the research ethics board at the University of Victoria (protocol number: 17-294).

Measures

Demographic information (age, sex, and ethnicity) and daily Instagram usage time were collected using self-report instruments. PA levels were assessed using the Recent Physical Activity Questionnaire (RPAQ) [42]. Minutes of PA at low (metabolic equivalent of task [MET] < 6/week), moderate (MET 3-6/week), and high (MET > 6/week) intensities were calculated based on the RPAQ scoring instruction [42].

Exercise identity was assessed using the Exercise Identity Scale [16], which is a 9-item questionnaire on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree) ($\alpha=.92$). Participants' exercise identities were determined by calculating a total score from the 9-item questionnaire; thus, scores ranged from 9 to 45 whereby a higher score represented a stronger exercise identity.

Participants' Instagram data were extracted using a customized program [43], which enabled the researchers to export users' posts, likes received for each post, and followings into separate Microsoft Excel files. Instagram posts extracted during the past 12 months before the completion of the study questionnaire were manually coded by 2 trained research assistants using standardized criteria to determine whether posts were PA related or not (eg, binary coded). PA-related posts were defined as those posts suggesting participants performed physical activities defined by the guidelines for exercise testing published by the American College of Sports Medicine [44]. Specifically, Instagram picture and video posts were identified as PA related if they included the person in sporting attire or the person in a location to perform PA (eg, kayaking, skiing, swimsuit, hiking) individually or in a group setting. Any discrepancies between the coders were resolved through discussion. The percentage of the number of PA-related posts relative to the total number of Instagram posts was calculated for each participant. The average number of likes for PA-related posts for each participant was also calculated.

An analysis of participants' Instagram "followings" was completed to identify the number of verified fitness or exercise-related Instagram accounts each participant followed. The verified fitness- or exercise-related Instagram accounts means Instagram has confirmed that an account is the authentic presence of the public figure, celebrity, or global brand it represents. A list of the most popular 300 verified fitness- or exercise-related Instagram accounts was compiled from publicly available Instagram rankings [45]. The percentage of exercise-related Instagram accounts following relative to the total number of Instagram followings was computed for each participant.

Statistical Analysis

All analyses were performed using the standard SPSS version 26.0 for Mac (SPSS, Inc./IBM), with a significance level set at $P<.05$. Descriptive statistical analyses were used to compare the number of willing participants with those who were not willing to share their Instagram data. Participants that were not willing to share their Instagram data were asked to provide reasons for not sharing in an open-ended question. Responses from these questions were grouped to create common themes. Independent t tests were used to compare differences between those groups (Instagram data shared versus Instagram data not shared) on demographics, PA, and exercise identity. A chi-square test was used to compare differences between categorical variables.

We examined normality (eg, skewness > 2 and kurtosis > 3) of all variables to determine whether any transformations were required by conversion to z-scores [46]. Pearson correlation analyses were used to evaluate the relationship among exercise identity, PA, and Instagram use metrics (percentage of PA-related Instagram post, percentage of fitness-related Instagram followings, and average number of "likes" on PA-related posts). Mediation analyses were based on 5000 bootstrapped samples using Hayes PROCESS Macro version 3.5 [47]. Multiple mediation analyses was used to examine whether exercise identity mediated the relationship between Instagram use variables and PA level [47]. This process involved examining path a, the association between Instagram use (independent variable) and exercise identity; path b, the impact of exercise identity (mediator variable) on PA level; and path c, the effect of Instagram use on PA levels (outcome variable). All analyses were controlled for age and sex. The 95% CIs must not cross 0 to satisfy the criteria for mediation.

Results

Participant Characteristics

A total of 76 participants were recruited and completed the study. Participant characteristics are presented in Table 1. The mean age was 19.7 (SD 0.31) years and the sample consisted of 72% female (55/76). Overall, 46% (35/76) of participants did not agree to share their Instagram data. Of these participants, 14 participants provided reasons for not sharing. Privacy was a major concern; specifically, there were concerns with (1) data oversight to prevent misusage of data and technology (n=9), (2) limitations to remain anonymous using Instagram photos (n=4), and (3) data ownership (n=1). Participants that were not willing to share their Instagram data reported significantly fewer weekly PA minutes at moderate to high intensity (MET ≥3/week) relative to participants that agreed to share their Instagram data (P=.04); however, there were no significant differences between groups for demographic characteristics, exercise identity, weekly PA minutes at low intensity (MET <3/week), and daily Instagram usage (Table 1).

A total of 41 participants agreed to share their Instagram data. The mean number of posts during the 12-month study period was 29.6 (SD 4.7). Examination of shared Instagram data revealed that, on average 7.3 (SD 1.1) or 25% of Instagram posts shared during the previous 12 months were related to PA. The mean number of Instagram following per person was 439.3 (SD 258) accounts. The mean number of verified PA- or fitness-related Instagram accounts following per user was 3 (SD 3.14). The mean number of "likes" received on all posts was 102.2 (SD 72.4) and the mean number of "likes" received on PA-related Instagram posts was 99 (SD 73).

Table 1. Baseline characteristics.

Characteristics	Instagram data shared (n=41)	Instagram data not shared (n=35)	P value
Age, mean (SD)	19.66 (2.08)	20.03 (1.81)	.42
Sex (female), n (%)	27 (66)	28 (80)	.20
Ethnicity, n (%)			.70
White	29 (71)	24 (69)	
East Asian	6 (15)	6 (17)	
Black	0 (0)	1 (3)	
Other (mixed, Spanish, South Asian)	6 (15)	4 (11)	
Exercise identity, mean (SD)	34.00 (8.20)	34.79 (7.16)	.67
Weekly PA[a] minutes at low intensity (MET[b] <3/week), mean (SD)	38.51 (135.94)	22.92 (59.94)	.50
Weekly PA min at moderate to high intensity (MET ≥3/week), mean (SD)	476.37 (542.40)	263.82 (238.27)	.04
Instagram usage (hours/day), mean (SD)	1.87 (1.04)	2.32 (1.91)	.29

[a]PA: physical activity.

[b]MET: metabolic equivalent of task.

The Relationship Among Exercise Identity, Physical Activity, and Instagram Use Metrics

Normality analysis revealed that weekly PA minutes at moderate to high intensity (MET ≥3/week) and fitness-related Instagram followings percentage were kurtotic (ie, values ≥3). These variables were transformed by conversion to z-scores. The Pearson correlation analyses are shown in Table 2. The results indicated that the percentage of PA-related Instagram posts ($r=0.38$; $P=.01$) and fitness-related Instagram ($r=0.39$; $P=.01$) followings were significantly associated with exercise identity. The strengths of the associates were moderate. We did not find that the average number of "likes" received ($r=0.05$, $P=.75$) to be significantly associated with exercise identity.

Table 2. Correlation analyses among exercise identity, physical activity, and Instagram use metrics.

Metrics	Variables				
	1	2	3	4	5
Pearson *r*					
Exercise identity	—				
Weekly PA[a] minutes at moderate to high intensity (MET[b] ≥3/week)	0.49	—			
Percentage of physical activity–related Instagram posts	0.38	0.03	—		
Percentage of verified fitness-related Instagram following	0.39	0.23	0.19	—	
Average number of "likes" received for the physical activity–related Instagram posts	0.05	0.01	–0.04	–0.13	—
***P* value**					
Exercise identity	—				
Weekly PA[a] minutes at moderate to high intensity (MET[b] ≥3/week)	.001	—			
Percentage of physical activity–related Instagram posts	.01	.86	—		
Percentage of verified fitness-related Instagram following	.01	.14	.23	—	
Average number of "likes" received for the physical activity–related Instagram posts	.75	.95	.78	.43	—

[a]PA: physical activity.

[b]MET: metabolic equivalent of task.

Mediation Analysis: Exercise Identity, Percentage of Physical Activity–Related Instagram Post, and Physical Activity

Exercise identity significantly influenced the relationship between the percentage of PA-related Instagram posts and PA levels (indirect effect: 3.28, 95% CI 0.54 to 8.11; direct effect: –3.65, 95% CI –9.70 to 2.39). As indicated in Table 3, a unit increase in percentage of PA-related Instagram posts was associated with a 0.10-unit increase in exercise identity; a unit increase in exercise identity was associated with 34 minutes of PA (MET ≥3/week) increase. Approximately 29% of the variance was accounted for by the predictors (R^2=0.29).

Table 3. Mediation analyses of the effects of exercise identity on Instagram use and physical activity levels.[a]

Effects	B	95% CI	β	*P* value
Instagram PA[b]photos—PA levels (MET[c] ≥3/week)[d]				
Path a: Instagram PA photos—exercise identity	0.10	0.01 to 0.18	.33	.02
Path b: Exercise identity—PA levels	34.0	11.66 to 56.36	.51	<.01
Path c: Instagram PA photos—PA levels	–3.65	–9.70 to 2.39	–.18	.22
Instagram PA following—PA levels (MET ≥3/week)[e]				
Path a: Instagram PA following—exercise identity	3.18	1.02 to 5.33	.48	<.001
Path b: Exercise identity—PA levels	29.34	7.16 to 51.52	.44	.01
Path c: Instagram PA following—PA levels	32.75	–129.17 to 194.68	.06	.68
Instagram PA "likes"—PA levels (MET ≥3/week)[f]				
Path a: Instagram PA "likes"—exercise identity	0.02	–0.01 to 0.06	.21	.17
Path b: Exercise identity—PA levels	28.6	6.75 to 50.47	.43	.01
Path c: Instagram PA "likes"—PA levels	0.28	–2.06 to 2.61	.04	.81

[a]Model covariates include age, gender, B (unstandardized beta), and β (standardized beta).

[b]PA: physical activity.

[c]MET: metabolic equivalent of task.

[d]Indirect effect: 3.28 (95% CI 0.54 to 8.11); direct effect: –3.65 (95% CI –9.70 to 2.39).

[e]Indirect effect: 91.9 (95% CI 36.93 to 174.82); direct effect: 60.53 (95% CI –115.15 to 236.21).

[f]Indirect effect: 0.69 (95% CI –0.10 to 2.28); direct effect: 0.28 (95% CI –2.06 to 2.61).

Mediation Analysis: Exercise Identity, Fitness-Related Instagram Followings, and Physical Activity

Exercise identity significantly influenced the relationship between the percentage of verified fitness-related Instagram following and PA level (MET ≥3/week; indirect effect: 91.9, 95% CI 36.93 to 174.82; direct effect: 60.53, 95% CI –115.15 to 236.21). A unit increase in percentage of fitness-related Instagram following leads to a 3.18-unit increase in exercise identity; a unit increase in exercise identity was associated with an increase of 29.34 minutes of PA (MET ≥3/week; Table 3). Approximately 26% of the variance was accounted for by the predictors (R^2=0.26).

Mediation Analysis: Exercise Identity, Physical Activity–Related Instagram "Likes," and Physical Activity

Exercise identity did not significantly influence the relationship between the average number of "likes" received for the PA-related Instagram posts and PA level (MET ≥3/week; indirect effect: 0.69, 95% CI –0.10 to 2.28). We did not observe a significant direct effect between the number of "likes" received on PA-related Instagram posts and PA minutes (direct effect: 0.28, 95% CI –2.06 to 2.61; Table 3).

Discussion

Principal Findings

The objectives of this study were to explore whether participants were willing to share their Instagram data with researchers to predict their lifestyle behaviors and examine the relationship among PA-related Instagram metrics, exercise identity, and PA. We found that the majority of participants were willing to share their Instagram data for monitoring lifestyle behaviors. We also found that only PA-related Instagram posts and fitness-related followings were significantly associated with an increase in exercise identity. Furthermore, we found that exercise identity significantly influenced the relationship between certain Instagram use metrics (eg, PA-related Instagram posts and fitness-related followings) and PA level. To our knowledge, this is one of the first studies to examine the relationship among Instagram use, exercise identity, and PA levels. The methods and the findings from this study can be used to inform future infodemiology PA-related studies.

Our results suggest that over 50% (41/76, 54%) of the users were willing to share their Instagram data for monitoring their PA and health-related behaviors. This observation was supported by our hypothesis. It was worth noting that participants that

agreed to share their data reported significantly higher weekly PA minutes at moderate to high intensity (MET ≥3/week) compared with participants who were not willing to share their data. A potential explanation is that those who were less active felt guilty for not showing consistent behavior with their social media posts. Individuals may not always portray their real selves on social media sites due to factors such as lower self-esteem, self-reflection, and self-concept clarity [48]. Alternatively, those participants may feel being monitored or judged for what they were posting, given the purpose of this study [27,49]. Concerns over privacy were the main reasons for not sharing their Instagram data. Unlike Twitter data, Instagram data are not publicly available by default, and the users mainly share pictures instead of text. Thus, this poses a limitation to remain anonymous when analyzing Instagram photos. Our findings reinforced the ethical challenges faced in using social data for public health monitoring. There is a need to establish "best-practice" standards for analyzing social media data while simultaneously acknowledging peoples' individual rights and respecting their privacy [50,51]. Previous studies have shown that users are more likely to accept the use of social media data for research if there is an oversight body that protects user data to ensure that their data are not being used beyond the intended purpose and that user data can remain anonymous to protect the identity of the user [52,53]. Overall, these privacy considerations need to be addressed for the future development of social data monitoring technology.

Based on the self-perception theory [24] and past research [25,26], we hypothesized that PA-related Instagram use would be positively associated with exercise identity and that exercise identity would mediate the relationship between PA-related Instagram use and PA minutes. Interestingly, we observed that only the percentage of PA-related Instagram posts and fitness-related followings were positively associated with exercise identity. Exercise identity also significantly influenced the relationship between these Instagram usage metrics and PA level (MET ≥3/week). Exercise identity did not influence the relationship between the number of "likes" and PA level. Previous research has demonstrated that social media "likes" can contribute to PA engagement [54,55]. The Instagram "likes" metric is unlike the number of followings and posts metrics because users do not have direct control over the number of "likes" received. The lack of consistency of our findings compared with previous literature may be attributed to the way individuals are using their Instagram in our sample. "Surveillance users" (eg, users focused on monitoring other users) may post lesser often and have a lesser number of followers, leading to a lower number of "likes" compared with "cool-ness users" who are motivated to become popular [56,57]. Thus, the types of users need to be taken into consideration for future studies.

Findings from this study have several implications for PA promotion research. First, Instagram may have the potential to target identity antecedents through the reception of both inspirational and informative posts about exercise from respected verified fitness accounts (ie, enhancing self-efficacy and congruency with the self) [58-60]. Based on the self-perception theory [24] and self-definition model [21], future PA promotion

interventions using Instagram may target personal investment, perceived commitment (ie, engaging with relevant fitness groups/users), social activation (ie, contributing and receiving support from fitness groups/users), and self-expression (ie, share photos of their PA or fitness-related experiences or write about these activities using social media), as these may play a role in strengthening an individual's exercise identity and subsequent PA behaviors [22,24]. However, it is worth noting that Instagram as a medium for expressing one's PA behavior with the end goal being to strengthen individuals' exercise identities may only be appealing to certain segments of the population. Second, results from this study extend previous Infodemiology research that Instagram data may be used to provide insights into health-related outcomes [39]. Future studies can build upon the methods used in this study and may also explore other Instagram metrics (eg, frequency of post) in PA infodemiology studies. Finally, PA researchers can leverage these social media analysis techniques to build prediction tools to monitor PA on a population level in real time. These tools may in the future aid public health agencies in identifying particular PA-related trends in various geographical areas on which to focus health and wellness initiatives.

There are potential negative aspects associated with using Instagram for the purpose of promoting PA and forming or strengthening one's exercise identity that must be considered. For example, recent movements known as "fitspiration" and "fitspro" have become popular on Instagram to inspire and motivate others to eat a healthy diet and exercise regularly. Fitspiration provides many positive attributes, such as accountability for exercising and education about fitness [61]. However, some women sharing fitspiration posts on Instagram may indicate disordered eating and exercise behaviors [62]. Despite the potential of using Instagram to motivate adherence to a healthy lifestyle, it is important for future interventions to consider the potential negative effect of social media on mental and physical health. Future studies need to examine ways to create a positive environment where participants feel comfortable posting pictures of their exercise experience and journey to fitness.

In terms of the limitations of the study, our sample only included university students which may limit the ability to generalize our findings beyond our sample. Future studies should examine the relationship between social media posts and PA-related behavior in middle-aged and older adults. Another limitation is that we used a list of verified fitness-related Instagram account in our analysis. It is possible that users may follow fitness-related Instagram accounts that are not verified and did not make it onto our list. Thus, this means that the observed relationship between exercise identity and the percentage of the fitness-related Instagram following needs to be interpreted with caution. Self-report questionnaires may result in confirmation bias. Moreover, it may be possible that some users' may have deleted photos, and thus affecting the number of PA-related photos in our analysis. Finally, the lack of longitudinal data for exercise identity and PA levels was another limitation when conducting mediation analysis. Thus, the mediation analysis result needs to be interpreted with caution. Future study with

longitudinal data is needed to build more complex multivariate models to better examine mediation effects.

Conclusion

This study examined the relationships among Instagram use, exercise identity, and PA levels. Findings from this study demonstrated that over 50% (41/76, 54%) of participants were willing to share their Instagram data with researchers for the purpose of monitoring and predicting PA behaviors. Exercise identity significantly influenced the relationship between Instagram use (eg, PA-related Instagram posts and fitness-related followings) and PA level. Results from this study suggest that there is an association between Instagram posts and PA-related outcomes (ie, exercise identity and PA levels). Future studies should examine whether Instagram may be used for a future intervention to help form or strengthen an individual's exercise identity, particularly if the individual can be supported by their "social world."

Conflicts of Interest

None declared.

References

1. Fletcher GF, Landolfo C, Niebauer J, Ozemek C, Arena R, Lavie CJ. Promoting Physical Activity and Exercise: JACC Health Promotion Series. J Am Coll Cardiol 2018 Oct 02;72(14):1622-1639 [FREE Full text] [doi: 10.1016/j.jacc.2018.08.2141] [Medline: 30261965]
2. Ozemek C, Lavie CJ, Rognmo Ø. Global physical activity levels - Need for intervention. Prog Cardiovasc Dis 2019 Mar;62(2):102-107. [doi: 10.1016/j.pcad.2019.02.004] [Medline: 30802461]
3. Statistics Canada. Distribution of the household population meeting/not meeting the Canadian physical activity guidelines, by sex and age group, occasional (percentage). 2015. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310033701 [accessed 2021-04-12]
4. Liu S, Brooks D, Thomas SG, Eysenbach G, Nolan RP. Effectiveness of User- and Expert-Driven Web-based Hypertension Programs: an RCT. Am J Prev Med 2018 Apr;54(4):576-583. [doi: 10.1016/j.amepre.2018.01.009] [Medline: 29456025]
5. Rhodes R, McEwan D, Rebar A. Theories of physical activity behaviour change: A history and synthesis of approaches. Psychology of Sport and Exercise 2019 May;42:100-109. [doi: 10.1016/j.psychsport.2018.11.010]
6. Hagger M, Weed M. DEBATE: Do interventions based on behavioral theory work in the real world? Int J Behav Nutr Phys Act 2019 Apr 25;16(1):36 [FREE Full text] [doi: 10.1186/s12966-019-0795-4] [Medline: 31023328]
7. McEwan D, Beauchamp M, Kouvousis C, Ray C, Wyrough A, Rhodes R. Examining the active ingredients of physical activity interventions underpinned by theory versus no stated theory: a meta-analysis. Health Psychol Rev 2019 Mar;13(1):1-17. [doi: 10.1080/17437199.2018.1547120] [Medline: 30412685]
8. Brand R, Cheval B. Theories to Explain Exercise Motivation and Physical Inactivity: Ways of Expanding Our Current Theoretical Perspective. Front Psychol 2019;10:1147 [FREE Full text] [doi: 10.3389/fpsyg.2019.01147] [Medline: 31164856]
9. Rhodes RE, de BG. What predicts intention-behavior discordance? A review of the action control framework. Exerc Sport Sci Rev 2013 Oct;41(4):201-207. [doi: 10.1097/JES.0b013e3182a4e6ed] [Medline: 23873134]
10. Burke P, Stets J. Identity Theory. New York, NY: Oxford University Press; 2009.
11. Stryker S, Burke PJ. The Past, Present, and Future of an Identity Theory. Social Psychology Quarterly 2000 Dec;63(4):284. [doi: 10.2307/2695840]
12. Burke PJ. Identity Change. Soc Psychol Q 2016 Jun 23;69(1):81-96. [doi: 10.1177/019027250606900106]
13. Strachan SM, Fortier MS, Perras MG, Lugg C. Understanding variations in exercise-identity strength through identity theory and self-determination theory. International Journal of Sport and Exercise Psychology 2013 Sep;11(3):273-285. [doi: 10.1080/1612197x.2013.749005]
14. Stryker S, Burke PJ. The Past, Present, and Future of an Identity Theory. Social Psychology Quarterly 2000 Dec;63(4):284. [doi: 10.2307/2695840]
15. Biddle B, Bank B, Slavings R. Norms, Preferences, Identities and Retention Decisions. Social Psychology Quarterly 1987 Dec;50(4):322 [FREE Full text] [doi: 10.2307/2786817]
16. Anderson DF, Cychosz CM. Development of An Exercise Identity Scale. Percept Mot Skills 2017 May 30;78(3):747-751. [doi: 10.1177/003151259407800313]
17. Rhodes RE, Kaushal N, Quinlan A. Is physical activity a part of who I am? A review and meta-analysis of identity, schema and physical activity. Health Psychol Rev 2016 Jun 02;10(2):204-225. [doi: 10.1080/17437199.2016.1143334] [Medline: 26805431]
18. Strachan SM, Brawley LR. Reactions to a perceived challenge to identity: a focus on exercise and healthy eating. J Health Psychol 2008 Jul 01;13(5):575-588. [doi: 10.1177/1359105308090930] [Medline: 18519432]

19. Strachan SM, Brawley LR, Spink KS, Jung ME. Strength of exercise identity and identity-exercise consistency: affective and social cognitive relationships. J Health Psychol 2009 Nov 26;14(8):1196-1206. [doi: 10.1177/1359105309346340] [Medline: 19858339]

20. Strachan SM, Flora PK, Brawley LR, Spink KS. Varying the cause of a challenge to exercise identity behaviour: reactions of individuals of differing identity strength. J Health Psychol 2011 May 23;16(4):572-583. [doi: 10.1177/1359105310383602] [Medline: 21346019]

21. Kendzierski D, Morganstein M. Test, revision, and cross-validation of the Physical Activity Self-Definition Model. J Sport Exerc Psychol 2009 Aug;31(4):484-504. [doi: 10.1123/jsep.31.4.484] [Medline: 19842544]

22. Kendzierski D, Furr R, Schiavoni J. Physical activity self-definitions: Correlates and perceived criteria. J Sport Exerc Psychol 1998;20:193. [doi: 10.1123/jsep.20.2.176]

23. Rhodes R. Chapter Five - The Evolving Understanding of Physical Activity Behavior: A Multi-Process Action Control Approach. In: Inlliot AJ, editor. Advances in Motivation Science. Cambridge, MA: MAlsevier Academic Press; 2017:171-205.

24. Bem DJ. In: Berkowitz L, editor. Self-perception theory. New York, NY: Academic Press; 1972.

25. Husband CJ, Wharf-Higgins J, Rhodes RE. A feasibility randomized trial of an identity-based physical activity intervention among university students. Health Psychology and Behavioral Medicine 2019 Apr 10;7(1):128-146. [doi: 10.1080/21642850.2019.1600407]

26. Lim C. Working Out with F.I.D.O. (Frequency, Intensity, Duration, & Outcomes) - A Feasibility Randomized Controlled Trial (MSc Thesis). 2017. URL: https://www.researchgate.net/publication/319177116 [accessed 2021-04-14]

27. Alhabash S, Ma M. A Tale of Four Platforms: Motivations and Uses of Facebook, Twitter, Instagram, and Snapchat Among College Students? Social Media + Society 2017 Feb 01;3(1):205630511769154. [doi: 10.1177/2056305117691544]

28. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009 Mar 27;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]

29. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]

30. Liu S, Zhu M, Yu DJ, Rasin A, Young SD. Using Real-Time Social Media Technologies to Monitor Levels of Perceived Stress and Emotional State in College Students: A Web-Based Questionnaire Study. JMIR Ment Health 2017 Jan 10;4(1):e2 [FREE Full text] [doi: 10.2196/mental.5626] [Medline: 28073737]

31. Liu S, Young SD. A survey of social media data analysis for physical activity surveillance. J Forensic Leg Med 2018 Jul;57:33-36 [FREE Full text] [doi: 10.1016/j.jflm.2016.10.019] [Medline: 29801949]

32. Liu S, Chen B, Kuo A. Monitoring Physical Activity Levels Using Twitter Data: Infodemiology Study. J Med Internet Res 2019 Jun 03;21(6):e12394 [FREE Full text] [doi: 10.2196/12394] [Medline: 31162126]

33. Liu S, Zhu M, Young SD. Monitoring Freshman College Experience Through Content Analysis of Tweets: Observational Study. JMIR Public Health Surveill 2018 Jan 11;4(1):e5 [FREE Full text] [doi: 10.2196/publichealth.7444] [Medline: 29326096]

34. Paul M, Dredze M. You are what you tweet: Analyzing twitter for public health. 2011 Presented at: Fifth International AAAI Conference on Weblogs and Social Media; July 17–21, 2011; Baltimore, MD.

35. Lee E, Lee J, Moon JH, Sung Y. Pictures Speak Louder than Words: Motivations for Using Instagram. Cyberpsychol Behav Soc Netw 2015 Sep;18(9):552-556. [doi: 10.1089/cyber.2015.0157] [Medline: 26348817]

36. Greenwood S, Perrin A, Duggan M. Social Media Update 2016. Washington, D.C: Pew Research Center; 2016. URL: https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2016/11/PI_2016.11.11_Social-Media-Update_FINAL.pdf [accessed 2020-08-01]

37. Perrin A, Anderson M. Share of U.S. adults using social media, including Facebook, is mostly unchanged since 2018. Washington, D.C: Pew Research Center; 2019. URL: https://www.pewresearch.org/fact-tank/2019/04/10/share-of-u-s-adults-using-social-media-including-facebook-is-mostly-unchanged-since-2018/ [accessed 2020-05-01]

38. Gruzd A, Jacobson J, Mai P, Dubois E. The State of Social Media in Canada 2017. SSRN Journal 2018:1-18. [doi: 10.2139/ssrn.3158771]

39. Reece AG, Danforth CM. Instagram photos reveal predictive markers of depression. EPJ Data Sci 2017 Aug 8;6(1):s13688. [doi: 10.1140/epjds/s13688-017-0110-z]

40. Tenkanen H, Di Minin E, Heikinheimo V, Hausmann A, Herbst M, Kajala L, et al. Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. Sci Rep 2017 Dec 14;7(1):17615 [FREE Full text] [doi: 10.1038/s41598-017-18007-4] [Medline: 29242619]

41. Talib S, Abdul RS, Olowolayemo A, Salependi M, Ahmad N, Kunhamoo S, et al. Perception analysis of social networks? privacy policy: Instagram as a case study. 2014 Presented at: 5th International Conference on Information and Communication Technology for the Muslim World (ICT4M); 17-18 Nov. 2014; Kuching, Malaysia. [doi: 10.1109/ict4m.2014.7020612]

42. InterAct Consortium, Peters T, Brage S, Westgate K, Franks PW, Gradmark A, et al. Validity of a short questionnaire to assess physical activity in 10 European countries. Eur J Epidemiol 2012 Jan;27(1):15-25 [FREE Full text] [doi: 10.1007/s10654-011-9625-y] [Medline: 22089423]

XSL•FO

RenderX

43. Mishra S. Instagram Data Download Internet. 2020. URL: https://codecanyon.net/item/instagram-scrapper/21146630 [accessed 2020-07-05]

44. American College of Sports Medicine. In: Lupash E, editor. ACSM's Guidelines for Exercise Testing Prescription. Ninth Edition. Philadelphia, PA: Wolters Kluwer Health, Lippincott Williams & Wilkins; 2014.

45. 30 Inspiring Fitness Girls To Follow On Instagram. 2020. URL: https://www.harpersbazaar.com/beauty/diet-fitness/g4018/inspiring-fitness-girls-on-instagram/ [accessed 2020-06-27]

46. Kim H. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod 2013 Feb;38(1):52-54. [doi: 10.5395/rde.2013.38.1.52] [Medline: 23495371]

47. Hayes A. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. New York, NY: Guilford Publications; 2017:1.

48. Yang C, Brown BB. Online Self-Presentation on Facebook and Self Development During the College Transition. J Youth Adolesc 2016 Feb 3;45(2):402-416. [doi: 10.1007/s10964-015-0385-y] [Medline: 26534776]

49. Hausenblas HA, Brewer BW, Van Raalte JL. Self-Presentation and Exercise. Journal of Applied Sport Psychology 2004 Jan;16(1):3-18. [doi: 10.1080/10413200490260026]

50. Vayena E, Salathé M, Madoff LC, Brownstein JS. Ethical challenges of big data in public health. PLoS Comput Biol 2015 Feb;11(2):e1003904 [FREE Full text] [doi: 10.1371/journal.pcbi.1003904] [Medline: 25664461]

51. Conway M. Ethical issues in using Twitter for public health surveillance and research: developing a taxonomy of ethical concepts from the research literature. J Med Internet Res 2014 Dec 22;16(12):e290 [FREE Full text] [doi: 10.2196/jmir.3617] [Medline: 25533619]

52. Beninger K, Lepps H. Research using social media: users' views. NatCen Social Research. 2014. URL: https://www.researchgate.net/publication/261551701 [accessed 2021-04-14]

53. Mikal J, Hurst S, Conway M. Ethical issues in using Twitter for population-level depression monitoring: a qualitative study. BMC Med Ethics 2016 Apr 14;17:22 [FREE Full text] [doi: 10.1186/s12910-016-0105-5] [Medline: 27080238]

54. Tu R, Hsieh P, Feng W. Walking for fun or for "likes"? The impacts of different gamification orientations of fitness apps on consumers' physical activities. Sport Management Review 2019 Nov;22(5):682-693. [doi: 10.1016/j.smr.2018.10.005]

55. Hamari J, Koivisto J. "Working out for likes": An empirical study on social influence in exercise gamification. Computers in Human Behavior 2015 Sep;50:333-347. [doi: 10.1016/j.chb.2015.04.018]

56. Sheldon P, Bryant K. Instagram: Motives for its use and relationship to narcissism and contextual age. Computers in Human Behavior 2016 May;58:89-97. [doi: 10.1016/j.chb.2015.12.059]

57. Bakhshi S, Shamma D, Gilbert E. Faces engage us: Photos with faces attract more likes and comments on instagram. New York, NY: ACM; 2014 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 26 to May 1, 2014; Toronto, ON p. 965-974 URL: https://www.researchgate.net/publication/266655817 [doi: 10.1145/2556288.2557403]

58. Al-Eisa E, Al-Rushud A, Alghadir A, Anwer S, Al-Harbi B, Al-Sughaier N, et al. Effect of Motivation by "Instagram" on Adherence to Physical Activity among Female College Students. Biomed Res Int 2016;2016:1546013-1546016 [FREE Full text] [doi: 10.1155/2016/1546013] [Medline: 27034927]

59. Santarossa S, Coyne P, Lisinski C, Woodruff SJ. #fitspo on Instagram: A mixed-methods approach using Netlytic and photo analysis, uncovering the online discussion and author/image characteristics. J Health Psychol 2019 Mar 15;24(3):376-385. [doi: 10.1177/1359105316676334] [Medline: 27852889]

60. Vaterlaus JM, Patten EV, Roche C, Young JA. #Gettinghealthy: The perceived influence of social media on young adult health behaviors. Computers in Human Behavior 2015 Apr;45:151-157. [doi: 10.1016/j.chb.2014.12.013]

61. DiBisceglie S, Arigo D. Perceptions of #fitspiration activity on Instagram: Patterns of use, response, and preferences among fitstagrammers and followers. J Health Psychol 2019 Aug 28:1359105319871656 [FREE Full text] [doi: 10.1177/1359105319871656] [Medline: 31455118]

62. Holland G, Tiggemann M. "Strong beats skinny every time": Disordered eating and compulsive exercise in women who post fitspiration on Instagram. Int J Eat Disord 2017 Jan;50(1):76-79. [doi: 10.1002/eat.22559] [Medline: 27302867]

Abbreviations

MET: metabolic equivalent of task
PA: physical activity
RPAQ: Recent Physical Activity Questionnaire

XSL•FO
RenderX