Modelling longitudinal directional associations between self-regulation, physical activity, and habit: results from a cross-lagged panel model

Abstract

**Background.** The directionality of associations between self-regulatory variables, behavior, and automaticity is seldomly tested. In this study, we aimed to examine a volitional, self-regulatory sequence of variables proposed in the HAPA framework (Intention → Action Plans → Action Control → Behavior) and its relationship with the construct of automaticity of the physical activity habit.

**Methods.** Longitudinal data was collected from high school students (N = 203, M\_age = 15.39 (SD = 1.43), 52% women) at three measurement points. First, a CFA measurement model was used to examine the study variables across measurement points. Next, a cross-lagged panel model was used to test directionality between variables.

**Results.** After adequate fit of the measurement model was confirmed, a mechanism integrating self-regulation with behavior and automaticity was examined. The hypothesized directionality between variables was verified overall by cross-lagged analysis. However, for the intention-action plans association, the inverse relationship was found: plans were associated with subsequent intentions, but intentions did not predict plans. Moreover, automaticity was not associated with subsequent physical activity behavior.

**Conclusions.** In general, our findings supported the hypothesized longitudinal direction of the associations, confirming that self-regulation may lead to behavior performance and automaticity. Unexpected findings and implications for intervention and future research are discussed.

**Keywords:** Physical Activity, Automaticity, Habit, Self-Regulation, HAPA model.
Introduction

Physical activity has numerous health benefits, from disease prevention to treatment of several chronic conditions [1,2]. In youth, vigorous physical activity brings added health benefits over and above those of moderate physical activity [3,4], many of which extend into adulthood [5]. However, recent global data confirms that the prevalence of insufficient activity among adolescents is alarming [6]. Some reports suggest that activity levels may decrease through adolescence and the transition to adulthood [7–10] and remain stable afterward [11]. Hence, a better comprehension of the mechanisms associated with physical activity initiation and maintenance during adolescence is crucial for understanding the causal paths involved and informing future intervention efforts.

Many theoretical models have been proposed to account for behaviors such as physical activity [12,13]. However, some aspects of these models have been assumed rather than tested. That is the case of the assumed directional longitudinal associations among self-regulatory variables. In a recent meta-analysis on the relationship between motivational and self-regulatory variables and health behaviors, Zhang, Zhang, Schwarzer, and Hagger (2019) recognized that the direction of associations between variables was not directly tested but only assumed on the basis of theoretical propositions. Even when correlational longitudinal data is collected, panel data is often not used, nor are adequate statistical analyses performed [14]. A cross-lagged model is required to examine the directionality of associations between variables [14–16]. Many studies rely on data that is either wholly or partially cross-sectional, and use analytical approaches (for instance, serial mediations [17,18]) that assume but do not test the directionality of associations among variables.
Moreover, most models define behavior as the final outcome, without further examination of the relationship between behavior and the construct of habit, despite habit being considered one of several key constructs related to behavioral maintenance [21].

Recent theoretical developments have provided promising conceptual distinctions, such as the one between frequent behavior and habit. In the past, researchers took them as synonyms [22], obscuring the complexity of the phenomena under study. More recently, Gardner defined habit as “a process by which a stimulus automatically generates an impulse towards action” [23]. In this regard, habit is not behavior or even frequent behavior. Habit is a learned automatic impulse to act, which depends on the presence of contextual cues to be prompted. Automaticity is at the core of this definition and has been described as the “active ingredient” of habits [24]. Frequent behavior might be the result of either goal-directed, intentional, deliberate processes, or automatic (habitual), context-driven processes [13, 23, 25–28]. Thus, a distinction should be made between intentional and habitual behaviors based not on their frequency but on the mechanisms controlling them.

High school settings may present opportunities for the promotion of health behaviors [29]. A better understanding of the longitudinal associations between habit, behavior, and other related constructs may help design intervention actions that promote adequate and sustained levels of physical activity in adolescents.

1. From intentions to habits

Although intentional and non-intentional mechanisms that control behaviors are distinct, they could be sequentially connected. In other words, there could be longitudinal associations between them, either directional or reciprocal. Lally and Gardner [13] have summarized conceptual and empirical work from several health behavior models to
describe the habit formation process, such as the Theory of Planned Behavior (TPB) [26],
the Transtheoretical Model (TTM) [25,30], and the Health Action Process Approach
(HAPA) [27], as well as contributions from studies on habit [31,32]. First, an intention
(behavioral goal) is developed; second, the intention must be translated into action; third,
the behavior must be repeated over time, which usually requires sustained motivation and
self-regulatory strategies; and fourth, the new action must be repeated in a fashion
conducive to the development of automaticity. Although research on the ‘fashion conducive
to automaticity’ is still scarce, the literature suggests that it involves the repeated
performance of behavior in a consistent context [13,31]. A more detailed presentation of
the factors facilitating the transition from behavior repetition to habit formation can be
found in Gardner and Rebar [28].

2. From self-regulation to behavior: the HAPA volitional proposal

Before the formation of a habit, a behavior must be initiated, involving a set of self-
regulation variables. Self-regulation is supposed to be under intentional and conscious
control. The Health Action Process Approach (HAPA) provides a framework to account for
the intentional initiation of behavior. This process is described as having two phases or
stages [27]. In the first one (motivational stage), an intention, or behavioral goal, is
elaborated (e.g., “I intend to practice physical activity regularly”). In the second stage
(volitional), the individual translates his/her intention into action using a sequence of self-
regulatory strategies, namely, the elaboration of action and coping plans and the setup of
action control strategies.

Action plans are detailed instructions on how, where, and when or how often the
desired behavior should be performed. For instance, after developing the intention of
running physical activity regularly, one might elaborate plans of going jogging (how),
early in the morning (when), from Monday to Friday (how often) at one specific park or
facility in the neighborhood (where). Coping plans are alternative actions to overcome
anticipated barriers to act [27]. For instance, if it is raining, one might have predefined
alternative plans (doing specific exercises indoors). Action control has been proposed as a
self-regulatory strategy consisting of three facets: awareness of standards (i.e., staying
aware of the desired end states), self-monitoring (i.e., monitoring one’s current behavior
and continuously comparing it to the envisioned standards), and self-regulatory effort (i.e.,
reducing the distance between current behavior and envisioned standards) [33]. Thus, after
having elaborated plans of going jogging (e.g., early in the morning, from Monday to
Friday, and so on), one may keep one’s plans in mind (awareness of standards), monitor
one’s behavior for the past seven days (self-monitoring), and when there is a gap between
current behavior and predefined plans, one may make an effort to reduce it (self-regulatory
effort).

Self-regulatory variables are supposed to bridge the intention-behavior gap [18,33–
35]. The literature often assumes that these sets of variables work in a sequence of
longitudinal and directional associations [18,36]; that is, intentions predict action plans,
action plans predict action control, and action control predicts behavior. However, Zhang,
Zhang, Schwarzer, and Hagger have summarized the findings of earlier studies and
supplied valuable meta-analytic evidence on the associations between HAPA variables. The
authors recognize that the studies’, most of them following cross-sectional and longitudinal
designs assume, but do not test, directionality [14], and recommend using panel data and
crossed-lagged models for testing specific effects [14,16].
The HAPA framework has also been used previously, not only as a theoretical model to account for the initiation of behavior but also for understanding how intentions can be translated into a learned *automatic* impulse to act (a habit) [20]. Therefore, a better understanding of the mechanisms proposed by the HAPA might provide some insights into the self-regulatory antecedents of habits.

4. This study: Examining hypothesized directional and reciprocal associations between self-regulation variables, physical activity, and automaticity

When a set of variables is longitudinally associated, there may be *directional* or *reciprocal* associations [37]. A directional association is when variable X is associated at one point in time with variable Y at a later point; however, Y is not associated at one point in time with X at a later point. A reciprocal association is when X is associated at one point in time with Y at a later point, and Y is associated at one point in time with X at a later point. Studies from models such as HAPA usually assume a directional sequence of variables and associations [18,36], but directionality needs to be tested [14].

To overcome that limitation, this study aimed to examine longitudinal data on physical activity from a sample of high school students and employ a cross-lagged model [15] to inspect the directionality and reciprocity of the associations in a volitional sequence of variables based on the HAPA framework and habit literature [18,27,28,38]. Considering that adolescence is a critical period for the development of physical activity habits [7–10, 39], a sample of high school students will be used to shed some light on the relationships between self-regulation, physical activity behavior, and habit in this age group.

From the theoretical background presented, we assume that intentions lead to action plans, which lead to action control. Action control leads then to physical activity, which
leads to the development of automaticity (although automaticity also leads to physical activity behavior). This rationale translates to the following set of hypotheses that this study aimed to examine:

1. There is a directional association between intentions and action plans: Intentions from an earlier time will be longitudinally associated with action plans at a later time (suggesting that intentions precede action plans). The opposite association is not hypothesized: earlier action plans are not expected to be related to later intentions (suggesting plans do not precede intentions).

2. There is a directional association between action plans and action control: Action plans from an earlier time will be longitudinally associated with action control at a later time (suggesting that plans precede action control). The opposite association is not hypothesized: earlier action control is not expected to be related to later action plans.

3. There is a directional association between action control and physical activity: Action control from an earlier time will be longitudinally associated with physical activity at a later time (suggesting that action control precedes behavior). The opposite association is not hypothesized: earlier physical activity is not expected to be related to later action control.

4. There are reciprocal associations between physical activity and automaticity: Physical activity from an earlier time is expected to be longitudinally associated with automaticity at a later time (suggesting that behavior precedes automaticity development). Also, automaticity from an earlier time is expected to be longitudinally associated with physical activity at a later time (suggesting that automaticity may lead to or be a cause of, physical activity).
Methods

Participants

The study included a longitudinal convenience sample of high school students, \( n = 203 \) (aged between 12 and 19 years; \( M_{\text{age}} = 15.39 \) (SD = 1.43), 52% women), living in the Great Metropolitan Area of Costa Rica. Students were recruited from two urban high schools. Written parental informed consent and adolescent informed assent to participate in the study were mandatory, per the Costa Rican legislation for research involving human subjects.

Procedure

Each participant was asked to complete a self-administered questionnaire on socio-demographic characteristics and intentions, action plans, action control, physical activity, and automaticity. Questionnaires were filled out in the classroom after coordinating with the corresponding teachers. All the instruments were applied in Spanish.

Students were instructed to fill out the same questionnaire at three separate times. At Time 1 (T1), 376 participants filled out the questionnaire; at Time 2 (T2, approximately two months later), there were 267 participants, and at time 3 (T3, about four months later after T2), there were 203 participants. Time gaps between measurement points were adjusted to the school calendar.

A local ethics committee approved the study. No monetary incentives or reimbursements were provided to participants, per the Costa Rican legislation for research involving human subjects.
Instruments

A Spanish 5-point Likert scale (1: “not at all true,” 5: “completely true”), was used to measure intentions, action plans, and action control related to physical activity. The items for each measure were conceptually consistent with the HAPA framework [27] and have been previously applied in adolescents [40]. The reliability and structure of these measures in Spanish have been previously reported [18]. Moreover, these general HAPA measures have been found to present medium-to-large correlations with vigorous physical activity [41].

Intentions were measured by three items, introduced by the stem “For the following weeks...”. An example item is “…I intend to practice physical activity regularly”. In our data, Cronbach’s alpha reliability coefficients were .91, .92, and .85 at T1, T2, and T3, respectively.

Action Plans were measured by three items. An example item is “I have already planned where to perform physical activity.” In this study, Cronbach’s alpha reliability coefficients were .85, .85, and .89 at T1, T2, and T3, respectively.

The measure of Action Control includes six items, two for each of its three facets: awareness of standards, self-monitoring, and self-regulatory effort [18, 33]. For the awareness-of-standards facet, the items were: “I have often had the intention to be physically active in my mind” and “I have always been aware of my plans to practice physical activity.” For the self-monitoring facet, the items were “I have monitored myself continuously to determine whether I perform physical activity frequently enough” and “I have carefully ensured that I practice physical activity for at least 30 minutes daily with the recommended strain per bout.” For the self-regulatory-effort facet, the items were “I have
tried my best to act according to my standards of physical activity” and “I have tried to practice physical activity frequently.”

Considering that there are more than five observed indicators, we created three distinct parcels [42] based on the item averages in each facet. Little [43] has recommended creating parcels based on averages instead of total scores because the average will be in the same metric as the original items. The parceling method we used is consistent with the similar content and higher bivariate correlation criteria described by Landis, Beal, and Tesluk [44]. In our data, alpha coefficients of the parcelled measures were .84, .86, and .91 at T1, T2, and T3, respectively.

Physical activity was measured by the item “How many days during the last week did you practice vigorous physical activity?” adapted from the International Physical Activity Questionnaire [45]. We used the version frequently used in Latin America[46], which focuses on vigorous physical activity within the past week, as in Barz et al. [41] and van Bree et al. [47]. Responses ranged from 0 to 7. The item was preceded by a short statement describing physical activity: “Physical activity produces sweating and rapid heartbeat. It makes you breathe harder than normal (e.g., lifting heavy objects, digging, aerobics, or cycling fast).” This definition was placed at the top of the questionnaire so that respondents could use it as a reference when answering items on intentions, action plans, action control, and habit. Among adolescents, vigorous physical activity brings added health benefits over and above those of moderate physical activity [3,4], and as such, it was the focus of this study.

Habit was measured using the Self-Report Behavioral Automaticity Index (SRBAI), a subscale of the Self-Report Habit Index [38,48,49] that focuses on the automatic
components of habits. This instrument has been previously applied in children and adolescents, and has shown good evidence of validity and reliability [50]. The stem “Physical activity is something...”) is followed by a 4-item scale (“I do automatically,” “I do without having to remember consciously,” “I do without thinking,” and “I start doing before realizing I am doing it”). The response format is a 7-point Likert scale (1: “completely disagree,” 7: “completely agree”). Gardner et al. [38] conducted 45 reliability assessments for SRBAI with data sets from various studies and found that internal consistency ranged from $\alpha = .68$ to $\alpha = .97$. In our data set, Cronbach’s alpha coefficient was .80, .88, and .92 at T1, T2, and T3, respectively.

Data analyses

First, the descriptive information and correlations were computed. Cut-off values suggested in the literature to detect collinearity based on correlations ($r \geq .85$ or $r \geq .90$) [51–54] were considered. Missing values were all < 5%, which has been deemed as inconsequential [55], and therefore they were entered into SPSS with the Expectation Maximization Algorithm [56,57] before analysis in the Structural Equation Models.

To evaluate the cross-lagged panel model, we used confirmatory factor analysis (CFA) with a maximum likelihood estimation in AMOS 18 [58]. Fit indices that minimize the likelihood of Type I and Type II errors [59] were employed: chi-square test ($\chi^2$), chi-square to degrees-of-freedom ratio ($\chi^2$/df), Comparative Fit Index (CFI), Tucker Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Akaike’s Information Criteria (AIC).

Several cut-off values have been proposed as strict criteria to evaluate Goodness-of-fit [59–62]: CFI and TLI: close to .90 or .95 [59,61,62]; RMSEA: close to .06 [59,61,62];
chi-square to df ratio: close to 2.0 [51] or 3.0 [61]. AIC is used when comparing models, and lower values are considered evidence of a better fit. [42]. However, scholars recommend caution against using these criteria as strict cut-offs without further considerations [52,60]. For instance, when there are multiple factors in a model (e.g., 5 to 10) and 50 or more items, conventional rules of thumb about goodness-of-fit are considered too restrictive [60]. Furthermore, previous studies have shown that indices such as RMSEA and CFI present lower values in more complex models and when the sample size is smaller, suggesting that they should be interpreted with caution in those circumstances [63].

The measurement model for the cross-lagged panel model was tested first. A cross-lagged panel model was then specified to examine for the presence of either directional or reciprocal associations between variables [15] (see Figure 1 for specific path details). A model is lagged when it includes auto-regressions, that is, regressions of a construct on itself across measurement points. Models are crossed when they include paths from one variable (“x”) on other variables (“y”). By specifying crossed regression paths, the existence of variance in an outcome not explained by autoregressive paths is examined [16]. In this study, the auto-regressions for each variable (intentions, action plans, action control, physical activity, and automaticity) were combined with specific crossed paths in the expected direction (intention → action plans → action control → physical activity → automaticity) and the opposite direction to what was theoretically expected (physical activity → action control → action plans → intention). We also added automaticity →
physical activity longitudinal paths, which we assumed as theoretically consistent with the
habit construct: learned *automatic* impulses to act (habits) are supposed to be one of the
influences leading to behavior (the other one being the conscious mechanism that involves
the intention $\rightarrow$ action plans $\rightarrow$ action control sequence described earlier). Thus, each of
the four study hypotheses was statistically examined as follows:

Hypothesis 1 – Directional longitudinal association from intentions to action plans:
Crossed paths from T1 intentions to T2 action plans, and from T2 intentions to T3 action
plans. Opposite direction for elimination purposes: Crossed paths from T1 action plans to
T2 intentions, and from T2 action plans to T3 intentions. Auto-regressions (from intentions
to subsequent intentions and from action plans to subsequent action plans).

Hypotheses 2 – Directional longitudinal association from action plans to action
control: Crossed paths from T1 action plans to T2 action control, and from T2 action plans
to T3 action control. Opposite direction for elimination purposes: Crossed paths from T1
action control to T2 action plans, and from T2 action control to T3 action plans. Auto-
regressions (from action plans to subsequent action plans and from action control to
subsequent action control).

Hypothesis 3 – Directional longitudinal association from action control to physical
activity behavior: Crossed paths from T1 action control to T2 physical activity, and from T2
action control to T3 physical activity. Opposite direction for elimination purposes: Crossed
paths from T1 physical activity to T2 action control, and from T2 physical activity to T3
action control. Auto-regressions were also specified.

Hypothesis 4 – Reciprocal longitudinal association between physical activity
behavior and automaticity: Crossed paths from T1 physical activity to T2 automaticity, and
from T2 physical activity to T3 automaticity. Opposite direction: paths from T1
automaticity to T2 physical activity, and from T2 automaticity to T3 physical activity. Auto-
regressions were also specified.

The examination of stationarity is recommended for cross-lagged models [15,16]. Stationarity means that there is an unchanging causal structure over time [16]. To find stationarity in this study, the structure of associations from T1 to T2 and T2 to T3 should be the same. This is done by comparing a model without constrains (Model 1) to a model where constrains are set to the auto-regressions (Model 2) and to a model where constraints are set to both auto-regressions and crossed paths (Model 3). If the latter model fits the data as well as the other models, stationarity is confirmed.

**Results**

**Descriptive statistics and correlations**

Correlations, means, and standard deviations are reported in Table 1. All correlations were statistically significant and had the expected (+/-) sign. Skewness and Kurtosis levels for all variables were within the levels recommended for a CFA with maximum-likelihood estimation [64]. Cross-sectional correlations between assumed independent variables were all below the levels considered problematic for multicollinearity ($r \geq .85$ or $r \geq .90$) [51–54]. All the factor loadings of the indicator variables of each construct were above 0.5 at every measurement point. Factor loadings are reported in Table 2.

---

Insert Table 1

---

Insert Table 2
Bearing in mind the dropout rate (only 53.8% of the original sample remained at T3), attrition analyses were performed, including age, sex, and all the study variables (intention, action plans, action control, physical activity, and automaticity). We found statistically significant differences for age only: those in the longitudinal sample were slightly younger ($M = 15.39$, $SD = 1.43$) than those who dropped out ($M = 15.84$, $SD = 1.21$) ($t(373.98) = 3.30, p < .01$). No statistically significant differences emerged for sex or other study variables.

Additionally, the study variables were compared by sex (men = 1, women = 2). Statistically significant differences favoring men were found at T1 intentions, T2 intentions, T1 action control, T2 automaticity, T1 physical activity, T2 physical activity, and T3 physical activity. For the rest of the variables, there were no statistically significant differences. A table summarizing statistical details on these findings is included as additional material (Appendix 1).

**Cross-lagged panel model**

A measurement model was estimated, which showed adequate fit to the data: $\chi^2 (713) = 1077.25, p < .001$; $\chi^2/df = 1.51$; CFI = .94; TLI = .93; AIC = 1457.25; RMSEA = .050 (CI 90% [.044; .056]).

The cross-lagged panel model was then constructed. Besides the auto-regressions, cross-lagged paths were initially specified in two directions: one based on the assumed sequence from previous research on the HAPA model [18], and another with the opposite directionality of associations (from planning to intentions, from action control to action plans, from physical activity to action control, and from automaticity to physical activity),
as depicted in Figure 1. A general model without constraints (Model 1) was stepwise compared to nested models to examine for stationarity assumptions. Namely, Model 1 was compared to a model where constraints were set to the auto-regressions (Model 2) and to a model with constraints in both auto-regressions and crossed paths (Model 3). Age and sex were included as covariates. We created a time proximity variable representing the time intervals between T1 and T2, and T2 and T3, and controlled time proximity for T2 and T3 variables. Table 3 summarizes the fit of these nested models.

Based on the AIC score, Model 3 showed the best fit with all the stationarity assumptions. However, the difference test ($\Delta \chi^2$) revealed that it did not significantly differ from the model with no constraints ($p > .05$), meaning that the same structure of longitudinal associations was supported in periods T1 to T2 and T2 to T3.

Figure 2 presents the results of Model 3’s path coefficients. This figure does not include cross-sectional within-time correlations. Correlations between variables were $r = .65$ or lower at T1; $r = .71$ or lower at T2, and $r = .75$ or lower at T3. These correlations were all below levels considered problematic for redundancy and collinearity ($r \geq .85$ or $r \geq .90$) [51–54].

Results on the crossed paths were as follows:

Hypothesis 1: Contrary to our expectations, T1 intentions were not associated to T2 action plans, and T2 intentions were not associated to T3 action plans ($p > .05$). Thus, the hypothesized directional association was not confirmed. In fact, the opposite direction was
confirmed: Action plans had a non-expected association with subsequent intentions ($b = .38; \beta = .39$ from T1 to T2; $\beta = .42$ from T2 to T3, $p < .001$) (action plans $\rightarrow$ intention).

Hypothesis 2: Action plans presented a longitudinal association (“effect”) with subsequent action control ($b = .36; \beta = .40$ from T1 to T2; $\beta = .40$ from T2 to T3, $p < .001$), in accordance to our assumptions. Also as expected, action control was not longitudinally associated to subsequent action plans ($p > .05$).

Hypothesis 3: The directional expected longitudinal association was confirmed: action control was associated to subsequent physical activity ($b = .24; \beta = .14$ from T1 to T2; $\beta = .12$ from T2 to T3, $p < .01$). Also as hypothesized, physical activity behavior was not associated to subsequent action control ($p > .05$).

Hypothesis 4: Physical activity was associated to subsequent automaticity ($b = .18; \beta = .22$ from T1 to T2; $\beta = .18$ from T2 to T3, $p < .001$). However, automaticity was not found to be longitudinally associated to physical activity ($b = .07; \beta = .06$ from T1 to T2; $\beta = .07$ from T2 to T3, $p = .15$). Thus, instead of the hypothesized reciprocal association between physical activity behavior and automaticity, the relationship was directional (physical activity $\rightarrow$ automaticity).

Based on the results of this specific sample, frequent physical activity seems to be longitudinally derived from a set of self-regulatory variables, but not from automaticity. We comment on this in the discussion section.

Results for the auto-regressions specified in the model (for T1 to T2 and T2 to T3) were as follows: intentions, $b = .24$; action plans, $b = .82$; action control, $b = .23$; frequency of physical activity, $b = .44$ and automaticity, $b = .29$. Auto-regression associations were all
When considering standardized coefficients (Figure 2), it becomes apparent that action plans were the most stable variable (T1 to T2, $\beta = .75$; T2 to T3, $\beta = .84$). In contrast, the least stable ones were intentions (T1 to T2, $\beta = .26$ and T2 to T3, $\beta = .24$) and action control (T1 to T2, $\beta = .26$ and T2 to T3, $\beta = .22$). Although automaticity auto-regressions were not the lowest, automaticity was the third least stable variable ($b = .29$, T1 to T2: $\beta = .24$ and T2 to T3: $\beta = .28$). We will take up this in the discussion section.

Concerning the variables introduced as covariates, their associations with T1 variables were as follows: Sex was negatively related to physical activity ($r = -.30$, $p < .001$), and negatively and marginally related to action control ($r = -.15$, $p = .05$), meaning that being a male was related to higher levels of physical activity and self-monitoring. Age was negatively related to T1 physical activity ($r = -.18$, $p < .05$), T1 automaticity ($r = -.19$, $p < .001$), T1 action control ($r = -.34$, $p < .001$), and T1 action plans ($r = -.19$, $p < .001$), suggesting that self-regulation and activity levels were lower in older participants. Time proximity only correlated with T1 intentions ($r = .17$, $p < .01$). Regarding T2 variables, sex only had a negative association with automaticity ($\beta = -.16$, $p < .05$); age was negatively related to physical activity ($\beta = -.12$, $p < .05$), and time proximity was positively related to automaticity ($\beta = .17$, $p < .05$). The relationship of the covariates with T3 variables was the following: Sex only had a negative association with physical activity ($\beta = -.13$, $p < .05$); age was marginally and negatively related to intentions ($\beta = -.11$, $p = .05$) and action plans ($\beta = -.15$, $p < .01$), and time proximity was associated to intentions ($\beta = .23$, $p < .001$), action
plans ($\beta = .21, p < .001$), action control ($\beta = .15, p < .01$), physical activity ($\beta = .15, p < .01$), and automaticity ($\beta = .21, p < .01$). Generally, when associations were found for sex, males had higher levels of physical activity, self-regulation, behavior, and habit. When associations were found for age, younger participants had higher self-regulation, more frequent behavior, and stronger habits. We are aware that the cross-lagged model presented in this section may be considered complex, and that complexity, along with sample size, may raise some concerns. Therefore, we conducted additional analyses in simpler cross-lagged models (consisting of three variables and covariates each one). Results basically replicated the findings from the more complex cross-lagged model (see Appendix 2 for details).

Discussion

Although conscious, goal-directed health behavior processes can be conceptualized as distinct and opposed to unconscious habits [65], they can also be conceived as sequentially related: a conscious goal-directed action, deliberately controlled at the beginning, may precede and turn into a habit over time. Traditionally, social-cognitive health behavior models, such as the HAPA [27], have not incorporated the habit construct. Recent research has emphasized the need to examine, rather than just assume theoretically, the directionality of associations between variables related to health behaviors [14]. This study aimed to address such need. Generally, the results from the cross-lagged panel model support the assumed directional longitudinal associations that were proposed for a sample of high school students. The study also found some unexpected but thought-provoking results, which are discussed later in this section.
We found that action plans were longitudinally related to subsequent action control, and action control was longitudinally related to subsequent behavior. This was in line with hypotheses 2 and 3, and with the sequence and directionality assumed in previous studies [18,36]. To our knowledge, and per recent literature [14], this was the first time that a HAPA set of longitudinal associations was tested, and not only assumed, which constitutes one of the main contributions of this study. Additionally, and beyond the traditional HAPA proposal, frequent physical activity was longitudinally associated with automaticity, suggesting that self-regulation may precede habits and play a role in habit-formation processes.

Other noteworthy results relate to the information we collected on the stability of associations over time. In its usual formulation, HAPA makes no specific hypothesis concerning the association stability [27]. Our study found that the structure of longitudinal associations between variables remained the same in both periods of study (T1 to T2 and T2 to T3). Moreover, based on the auto-regression coefficients, action plans were the most stable variable, and intentions were the least stable. Although the stability of intentions has been studied before [34], the stability of action plans has not received, to our knowledge, much attention in previous research.

Not every finding was as expected. One surprising result that contradicted Hypothesis 1 was that intention had no longitudinal association to (or no effect on) subsequent action plans. More hypotheses are needed in this regard. One possibility is that intentions may be necessary for the formulation of new plans, but less so for plans that might be already formulated [66]. Our participants reported action plans at levels slightly above the mid-point of the resting scale ($M = 3.46$ at T1) [3].
Another unexpected result was that action plans at one point in time were associated with intentions at a later point. Hence, it seems that plans played both a self-regulatory (volitional) and a motivational role. Although we only hypothesized on the self-regulatory role of plans, the motivational role has been reported before [67]. By recalling action plans, participants may also recall experiences of success and, therefore, motivational scenarios.

Another possibility for the unexpected results in the intention-action plan associations is redundancy because it could render a vital predictor (e.g., intentions) less relevant. However, as mentioned in the results for the cross-lagged model, the cross-sectional correlations between study variables at T1 and T2 were all below the levels suggested in the literature as indicative of redundancy problems [51–54].

The lack of a longitudinal association from automaticity to subsequent behavior was also surprising and contradicted our Hypothesis 4. It may be that, in this sample, the habit of physical activity was not yet formed. Behavior is mostly under the control of conscious, self-regulatory mechanisms, and not yet under the control of an automatic impulse to act. Automaticity means ranged from 3.67 at T1 to 3.47 at T3 (in a 1–7 scale). We also found automaticity stability over time to be lower than expected. We must recognize that this observational study did not measure if conditions for habit maintenance already existed in the context. For instance, we did not measure context stability, which is considered a critical factor in habit formation and prevention of habit disruption processes [28,68]. As reported afterward by the high school authorities, students should have 40 to 80 minutes of physical education per week. However, the program may vary through the academic year, and physical education is often not prioritized when there are other academic obligations.

Variations in cue availability provided by physical education lessons may have, therefore,
influenced automaticity levels [50]. Future research should try to control for cue availability over time.

Some limitations of this study must be mentioned. Although the study was designed to test for the directionality of longitudinal associations over time and, overall, supported the sequence of variables previously hypothesized, the design was still correlational and does not provide the most solid foundation for causality inferences. Moreover, generalization to other samples, habits, behaviors, and situations is not advised. We suggest instead that replications of this study should be carried out with different samples and target behaviors as they may entail degrees of conscious self-regulation and automaticity. Additionally, the behaviors in this study were self-reported, which can introduce some bias to the information due to recall imprecisions and social desirability. Future studies should add some objective measures of behavior.

Additionally, although ‘last week’ is frequently used in the Spanish wording of the question selected from IPAQ [46], and it has also been used in other languages [41,46,47], it may raise some concerns for lack of clarity [46]. Some respondents may estimate their answers based on the last seven days, while others may refer to the last Sunday-to-Saturday period. Future research may include a small calendar as a visual aid for participants to interpret the written question more precisely [46]. The complexity of the tested model and the size of the sample may also raise some concerns. Appendix 2 shows results based on simpler models. Results are consistent with those of the complex model, suggesting that our conclusions are reliable.

This work has important implications for intervention. Action plans emerged as a very relevant variable because of their association with motivational (intention) and volitional (action control) processes. This is in line with previous reports on the crucial role
of plans for health-related behaviors and health promotion interventions [69]. The findings also provided further evidence of the contribution of self-monitoring (i.e., action control) to account for behavior [18,66]. Self-monitoring plays a vital role in bridging the intention-behavior gap, which is particularly relevant in physical activity promotion interventions [70]. Most prior studies have been conducted in adults only [14], so it is important to highlight that this study was conducted on a sample of adolescents, and its results suggest that a sequence of conscious self-regulation variables is relevant to that population as well.

In conclusion, our study provided evidence to suggest that conscious and deliberate processes of change in health behaviors are not unrelated to automaticity but can instead be integrated. According to our findings, plans were longitudinally and directionally associated with subsequent action control and intentions; action control was longitudinally associated with subsequent frequency of physical activity, and frequent physical activity was longitudinally associated with subsequent automaticity (physical activity habit). Understanding how these processes unfold is vital to further our knowledge of health behavior change and how it may be maintained over time, and it contributes to theory refinement and to inform the design of effective behavioral interventions targeted to protecting and promoting health.

**Statements regarding informed consent and ethical approval**

*Per Costa Rican legislation, parental informed consent and adolescent informed assent was obtained for all participants included in the study. All procedures performed in studies involving human subjects were in accordance with the standards of the Ethical Review Committee and the 1964 Declaration of and its amendments.*
References


Running head: SELF-REGULATION AND HABIT


Landis RS, Beal DJ, Tesluk PE. A comparison of approaches to forming composite


### Table 1. Bivariate correlations, means, and standard deviations at three time points.

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
<th>T4</th>
<th>T5</th>
<th>T6</th>
<th>T7</th>
<th>T8</th>
<th>T9</th>
<th>T10</th>
<th>T11</th>
<th>T12</th>
<th>T13</th>
<th>T14</th>
<th>T15</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. T1 Intention</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. T1 Action Planning</td>
<td>.73***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. T1 Action Control</td>
<td>.72***</td>
<td>.71***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. T1 Physical Activity</td>
<td>.50***</td>
<td>.44***</td>
<td>.56***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. T1 Automaticity</td>
<td>.43***</td>
<td>.41***</td>
<td>.41***</td>
<td>.32***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. T2 Intention</td>
<td>.60***</td>
<td>.53***</td>
<td>.62***</td>
<td>.44***</td>
<td>.34***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. T2 Action Planning</td>
<td>.41***</td>
<td>.37***</td>
<td>.38***</td>
<td>.16***</td>
<td>.13*</td>
<td>.48***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. T2 Action Control</td>
<td>.24***</td>
<td>.21***</td>
<td>.25***</td>
<td>.20***</td>
<td>.28***</td>
<td>.23***</td>
<td>.18***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. T2 Physical Activity</td>
<td>.46***</td>
<td>.41***</td>
<td>.53***</td>
<td>.63***</td>
<td>.35***</td>
<td>.45***</td>
<td>.30***</td>
<td>.20***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. T2 Automaticity</td>
<td>.36***</td>
<td>.31***</td>
<td>.36***</td>
<td>.32***</td>
<td>.34***</td>
<td>.34***</td>
<td>.21**</td>
<td>.17*</td>
<td>.39***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. T3 Intention</td>
<td>.55***</td>
<td>.53***</td>
<td>.58***</td>
<td>.43***</td>
<td>.34***</td>
<td>.57***</td>
<td>.42***</td>
<td>.22***</td>
<td>.42***</td>
<td>.28***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. T3 Action Planning</td>
<td>.51***</td>
<td>.55***</td>
<td>.54***</td>
<td>.33***</td>
<td>.36***</td>
<td>.50***</td>
<td>.48***</td>
<td>.22***</td>
<td>.44***</td>
<td>.37***</td>
<td>.76***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. T3 Action Control</td>
<td>.24***</td>
<td>.18***</td>
<td>.25***</td>
<td>.13***</td>
<td>.18*</td>
<td>.19***</td>
<td>.21***</td>
<td>.52***</td>
<td>.22***</td>
<td>.22***</td>
<td>.11 n.s.</td>
<td>.27***</td>
<td>.34***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. T3 Physical Activity</td>
<td>.38***</td>
<td>.35***</td>
<td>.45***</td>
<td>.57***</td>
<td>.31***</td>
<td>.35***</td>
<td>.27***</td>
<td>.14**</td>
<td>.60***</td>
<td>.32***</td>
<td>.47***</td>
<td>.45***</td>
<td>.20***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. T3 Automaticity</td>
<td>.33***</td>
<td>.33***</td>
<td>.36***</td>
<td>.34***</td>
<td>.41***</td>
<td>.40***</td>
<td>.21**</td>
<td>.23***</td>
<td>.39***</td>
<td>.39***</td>
<td>.39***</td>
<td>.33***</td>
<td>.23***</td>
<td>.44***</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>- .22**</td>
<td>- .26**</td>
<td>- .31**</td>
<td>- .22**</td>
<td>- .20**</td>
<td>- .13</td>
<td>- .22*</td>
<td>- .05</td>
<td>- .27**</td>
<td>- .20**</td>
<td>- .33**</td>
<td>- .39**</td>
<td>- .18*</td>
<td>- .19**</td>
<td>- .15*</td>
</tr>
<tr>
<td>M</td>
<td>3.89</td>
<td>3.46</td>
<td>3.45</td>
<td>2.29</td>
<td>3.67</td>
<td>3.76</td>
<td>3.27</td>
<td>2.59</td>
<td>1.99</td>
<td>3.60</td>
<td>3.77</td>
<td>3.27</td>
<td>2.60</td>
<td>1.94</td>
<td>3.47</td>
</tr>
<tr>
<td>SD</td>
<td>1.22</td>
<td>1.46</td>
<td>1.14</td>
<td>2.14</td>
<td>1.63</td>
<td>1.23</td>
<td>1.89</td>
<td>1.23</td>
<td>2.02</td>
<td>1.92</td>
<td>1.27</td>
<td>1.45</td>
<td>1.32</td>
<td>2.01</td>
<td>1.96</td>
</tr>
</tbody>
</table>

Notes: *** p < .001; ** p < .01; * p < .05. Factor correlations are above the diagonal. Correlations between composite scores of each variable (averaged items) are below the diagonal.
Table 2. Factor loadings for variables at each time point

<table>
<thead>
<tr>
<th></th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intention</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>.87</td>
<td>.86</td>
<td>.87</td>
</tr>
<tr>
<td>Item 2</td>
<td>.85</td>
<td>.90</td>
<td>.87</td>
</tr>
<tr>
<td>Item 3</td>
<td>.90</td>
<td>.91</td>
<td>.89</td>
</tr>
<tr>
<td><strong>Action Plans</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>.77</td>
<td>.81</td>
<td>.82</td>
</tr>
<tr>
<td>Item 2</td>
<td>.87</td>
<td>.82</td>
<td>.87</td>
</tr>
<tr>
<td>Item 3</td>
<td>.80</td>
<td>.82</td>
<td>.88</td>
</tr>
<tr>
<td><strong>Action Control</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parcel 1</td>
<td>.81</td>
<td>.78</td>
<td>.90</td>
</tr>
<tr>
<td>Parcel 2</td>
<td>.78</td>
<td>.83</td>
<td>.91</td>
</tr>
<tr>
<td>Parcel 3</td>
<td>.80</td>
<td>.86</td>
<td>.84</td>
</tr>
<tr>
<td><strong>Automaticity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Item 1</td>
<td>.74</td>
<td>.84</td>
<td>.90</td>
</tr>
<tr>
<td>Item 2</td>
<td>.75</td>
<td>.90</td>
<td>.89</td>
</tr>
<tr>
<td>Item 3</td>
<td>.75</td>
<td>.77</td>
<td>.84</td>
</tr>
<tr>
<td>Item 4</td>
<td>.60</td>
<td>.61</td>
<td>.79</td>
</tr>
</tbody>
</table>
Table 3. Fit of the cross-lagged panel nested models

<table>
<thead>
<tr>
<th>Model</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\chi^2$/df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA CI 90%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[Lo; Hi]</td>
</tr>
<tr>
<td>Model 1</td>
<td>1651.88</td>
<td>878</td>
<td>1.88</td>
<td>.89</td>
<td>.87</td>
<td>.066 [.061; .071]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2057.88</td>
</tr>
<tr>
<td>Model 2</td>
<td>1652.39</td>
<td>883</td>
<td>1.87</td>
<td>.89</td>
<td>.87</td>
<td>.066 [.061; .071]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2048.39 .50</td>
</tr>
<tr>
<td>Model 3</td>
<td>1656.54</td>
<td>891</td>
<td>1.86</td>
<td>.89</td>
<td>.88</td>
<td>.065 [.060; .070]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2036.54 4.66</td>
</tr>
</tbody>
</table>
Figure 1. The specified cross-lagged model. Correlations between exogenous variables and disturbances in cross-sectional data are specified but not depicted in this figure. Paths from time proximity, sex and age at T1, T2 and T3 variables are also specified.
Figure 2. Results of the cross-lagged model. Notes: Standardized (left) and unstandardized (right) coefficients are reported for every path. Non-significant paths ($p > .05$) are depicted with dashed lines. Paths from the sex, age, and time proximity covariates, and within-time correlations are specified but not depicted in this figure. ***$p < .001$, **$p < .01$. Within-time correlations among variables (or disturbances) were between $r = .00$ and $.74$. 

Intention T1 → Intention T2: 0.26/24***
Intention T2 → Intention T3: 0.24/24***
Action Plans T1 → Action Plans T2: 0.39/0.38***
Action Plans T2 → Action Plans T3: 0.42/0.38***
Action Control T1 → Action Control T2: 0.26/0.23***
Action Control T2 → Action Control T3: 0.22/0.23***
Physical Activity T1 → Physical Activity T2: 0.46/0.44***
Physical Activity T2 → Physical Activity T3: 0.44/0.44***
Automaticity T1 → Automaticity T2: 0.24/0.29***
Automaticity T2 → Automaticity T3: 0.28/0.29***