
ORIGINAL RESEARCH REPORT

Planfulness: A Process-Focused Construct of Individual Differences in Goal Achievement

Rita M. Ludwig, Sanjay Srivastava and Elliot T. Berkman

Goal pursuit outcomes are partly caused by the way people think about goals. Specific patterns of thought can increase the likelihood of goal achievement, such as generating heuristics to automate goal-related decision making, orienting present-moment attention to the future to increase the salience of a distal goal, and contrasting the anticipated enjoyment of an achieved goal with the progress required to complete it. However, it is unknown whether there are stable individual differences in the tendency to deploy particular meta-cognitions during goal pursuit. A tool to assess such differences would help to identify and intervene on personal barriers to goal progress. Here, we define a new construct within the conscientiousness domain—planfulness—that captures a person’s proclivity to adopt efficient goal-related cognition in pursuit of their goals. We hypothesize that planfulness consists of three interrelated facets representing distinct mental processes, temporal orientation (TO), cognitive strategies (CS), and mental flexibility (MF), and that planfulness predicts goal achievement on an individual basis. We developed a 30-item Planfulness Scale with three subscales tested and refined across 5 studies and 10 samples (total unique $N = 4,318$) using iterative exploratory and confirmatory factor analysis on data collected from both student and on-line samples. The Planfulness Scale demonstrated both convergent and discriminant validity when compared to other measurements, and scale scores predicted goal progress in a longitudinal study. We find that planfulness is a useful new construct for self-regulation research, and the 30-item Planfulness Scale to be a valid and reliable measurement of real-world goal achievement.

Keywords: Goals; Self-regulation; Individual differences; Planning; Human performance

How do people turn their desires into reality? The process of transforming something wanted into something obtained involves goal setting and pursuit. A goal is a desired end state, but many people struggle to translate their initial goal formations to ultimate goal success. Previous experimental research indicates that if a goal is cognitively linked to a set of means to achieve it, that provides a framework for coordinating cognitive resources, decision-making, and behavior. In parallel, there has been some work on the trait construct of “planfulness,” but there is not yet a clear link between the experimental and trait-based approaches success in goal pursuit. The aim of the present work is to develop a measure of individual difference in cognitive processes that lead to successful goal pursuit and gather initial evidence about its validity. Such a measure would help identify specific ways that individuals can improve their likelihood of achievement success.

Background: Two Streams of Goal Pursuit Research

Our approach brings together insights from two parallel research programs. First, working primarily from an experimental approach, psychology research conducted within laboratory contexts has uncovered a number of specific patterns of thought that are likely to lead to effective goal pursuit. When individuals are experimentally induced to use strategies such as linking specific goals to particular patterns of behavior (Gollwitzer, 1993) and envisioning the distance between future goal achievement and the present reality of one’s progress (Oettingen, 2000), these strategies reliably increase rates of goal attainment in controlled lab experiments. This evidence suggests that the way one thinks about a desired goal contributes to its eventual achievement.

Second, researchers from a complementary approach focus on the influence of enduring traits or dispositions associated with successful goal pursuit. Work in this field often starts with phenotypic description, such as personality attributes represented in natural language or questionnaire items based on the descriptive language of applied practitioners (John, Naumann & Soto, 2008;

Saucier & Srivastava, 2015). Such work has established links between certain stable individual differences and achievement of 'real-world' outcomes. Among the Big Five personality traits, conscientiousness is predictive of successful outcomes in a variety of domains, including academics (Nofle & Robbins, 2007), health (Roberts, Walton, & Bogg, 2005), and work performance (Dudley, Orvis, Lebiecki, & Cortina, 2006). Grit, a trait describing one's degree of sustained effort toward long-term pursuits, is also associated with measures of academic success, including undergraduate GPA and National Spelling Bee ranking (Duckworth, Peterson, Matthews, & Kelly, 2007); grit is strongly associated with conscientiousness (Credé, Tynan, & Harms, 2016). These findings, amongst others (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007; Shiner, Masten, & Roberts, 2004), suggest that it is critical not to overlook the role of individual differences in considering the likelihood of successfully reaching a goal.

In the present work, we draw on insights from both experimental and individual-differences research. Our overarching goal is to weave these two threads together to account for naturally occurring variation in goal achievement across individuals as a function of meta-cognitive tendencies. To do this, we conducted a series of studies to build and validate a comprehensive measure of a construct within the conscientiousness domain, *planfulness*, which refers to individual differences in the tendency to use specific cognitive processes that experimental research has shown to promote goal achievement. As such, planfulness is hypothesized to be useful beyond broader measures for capturing individual differences in goal progress.

Planfulness

We use the term "planfulness" to describe an individual's tendency to engage in three particular patterns of thought with respect to goals. Previous work has described planfulness in terms of a general tendency to plan before acting (Frese, Stewart, & Hannover, 1987) but did not attempt to systematically link the construct to the cognitive habits that relate to successful self-regulation in the laboratory. We conceptualized of planfulness as relatively stable in individuals across time, circumstances, and goals. To the extent that individuals vary on a continuum of planfulness, those who are high in this trait are expected to tend to deploy more effective meta-cognition when they pursue goals and therefore be more likely to achieve them.

We propose that planfulness consists of three interrelated facets that map on to the broad cognitive strategies that self-regulation research shows to be reliably associated with improved goal outcomes in experimental settings: temporal orientation (TO) to the future implications of present behavior, mental flexibility (MF) in contextualizing one's actions in terms of one's goals, and cognitive strategies (CS) to anticipate and deal with potential obstacles. Importantly, this is not to suggest that these facets cover the entire scope of possible factors that contribute to goal achievement, but rather that the facets represent initial categories of practices that

are currently known to reliably improve goal progress. Additional facets would be hypothesized to the extent that additional cognitive strategies are discovered that also reliably improve goal progress and explain additional variance beyond these three.

Relation to Conscientiousness

The trait of planfulness is conceptually situated in the broader personality domain of conscientiousness. Conscientiousness, like the other Big Five personality characteristics, was initially identified via the lexical approach; that is, by using factor analysis on person-descriptive adjectives sampled from natural language to isolate distinguishable patterns of how individuals describe one another. This approach assumes that the coalescence of many words around a single concept indicates that that concept is socially important (Saucier & Srivastava, 2015). Although the exact number and type of conscientiousness facets is contested (Costa, McCrae, & Dye, 1991; Hill & Roberts 2011; MacCann, Duckworth, & Roberts, 2009) the resulting broad factor assesses individuals on continua of orderliness, self-control, punctuality, and dutifulness. High conscientiousness has been associated with a number of positive life outcomes, including academic success (Chamorro-Premuzic & Furnham, 2003) and longevity (Bogg & Roberts, 2004). There are also many other measures related to self-control and achievement that are not explicitly labeled as facets of conscientiousness (see Duckworth & Kern, 2011, for a meta-analysis) that are expected to overlap with planfulness.

Despite all that is known about conscientiousness, questions remain about the psychological mechanisms and processes that contribute to individual differences in conscientiousness and influence behavior and long-term outcomes (Roberts, Lejuez, Krueger, Richards, & Hill, 2012). A personality trait can be described as the central tendency of a distribution of states, each of which reflects momentary outputs of processing (Fleeson, 2001). In the case of conscientiousness, the constituent processes are only partially understood, in part because models of the lower-order structure of conscientiousness have largely been derived from factor-analytic studies of natural language or descriptive item sets (Roberts et al., 2012). Because of our focus on the specific cognitive processes identified as effective in laboratory experiments, the present research may help clarify some of the specific mechanisms by which highly conscientious individuals achieve goal successes, and in turn, may therefore reveal tactics that individuals low in conscientiousness could be taught in order to encourage goal achievement success.

Distinctiveness from Motivation and Impulsivity

Items that specifically tap individuals' level of motivation to pursue a goal or their level of impulse control were deliberately not included in the Planfulness Scale. Planfulness is distinct from motivation and impulsivity for three reasons. First, experimental research has found that effective goal-related cognition increases rates

of achievement regardless of an individual's level of motivation (Gollwitzer, 1999; Orbell & Sheeran, 2000) and capability for self-control of emotional impulses (Hofmann, Deutsch, Lancaster, & Banaji, 2010; Webb & Sheeran, 2003). Second, from a trait perspective, the relation between motivation and conscientiousness has long been unclear (Roberts et al., 2012). By avoiding items that refer to motivation, we were able to avoid this conceptual confusion and focus more precisely on cognitive processes that are conceptually distinct from motivation and impulse control. Third, although some types of impulsivity are conceptually highly similar to planfulness (e.g., lack of planning; Whiteside & Lynam, 2001) and are expected to relate negatively to it, impulsivity as it relates to in-the-moment behavior is conceptually orthogonal to planfulness (Frese et al., 1987), a divergence from constructs strictly related to self-control (Friese, & Hofmann, 2009). A person who is low in motivation and/or highly impulsive is not definitionally low in planfulness; one can be unmotivated but still plan, or plan carefully then act on impulse in the spur of the moment. Also, planfulness could compensate for low motivation or high impulsivity. For example, the effects of impulsive behavior might be mitigated by adhering to specific cognitive strategies, and low levels of motivation might be overcome by reflecting on how a current action brings goal achievement closer. These relationships can be borne out empirically but are part of the construct definition of planfulness.

The Present Research

Adopting a future time perspective, developing implementation intentions, and mentally contrasting the feeling of goal actualization with the current reality have been found to be especially beneficial to goal progress. The present research investigates whether there are stable differences across individuals in their tendency to use those processes. To the extent that those differences do exist, people who tend to engage in beneficial patterns of meta-cognition are likely to make progress in their goals. Therefore, the hypothesized three-facet structure of planfulness containing Temporal Orientation (TO), Mental Flexibility (MF), and Cognitive Strategies (CS) elements directed the development of the Planfulness Scale and the selection of items to represent of each of these three facets.

We used the model of scale development suggested by Simms (2008) to balance theoretical and psychometric considerations with respect to inclusion of scale items. First, reflecting the substantive validity phase, we used our theory and related literature to define the planfulness construct and develop an initial item pool (Study 1 – see Supplemental Materials). Second, in the structural validity phase, we use an iterative process to refine scale items based on psychometric evaluation (Studies 2 through 4). Finally, we present Study 5 as the first step in the external validity phase, in which we test the Planfulness Scale on convergent, discriminant, and criterion validity. The latter is tested through the ability of our finalized scale to longitudinally predict goal achievement.

Studies 2 through 4: Iterative Scale Refinement

For the sake of brevity, we begin with reporting the formal psychometric testing of different versions of the Planfulness Scale. A full accounting of how the hypothesized structure of planfulness drove the initial item selection, and testing of a pilot pool of items in Study 1, can be found in the Supplemental Materials. The results of Study 1 supported the three-component structure of planfulness and provided 42 scale items for refinement in Studies 2 through 4.

The purpose of Studies 2 through 4 was to iteratively test and improve a selection of scale items in a hypothesized model of planfulness until adequate scale performance was reached, as assessed by psychometrics and measures of model fit. For clarity, each study here refers to a unique scale version, and each sample within a study refers to a unique group of people who completed that scale version. Our framework for this process was to begin with item and model testing, then make modifications of the scale based on the previous analyses, and to iterate through the cycle again by testing these modifications with either the same or a different version of the scale in a new, independent sample. For example, if certain items exhibited poor psychometrics in one sample, improvements in model fit with these items removed were confirmed in a new sample. All model tests were conducted using entirely independent samples, often resulting in the collection of multiple samples per study; in this way we ensured that the final scale and model versions would not be unduly influenced by the characteristics of any given sample.

As the methods for data collection and analysis across Studies 2 through 4 were the same, the following sections review these iterative studies collectively.

Method

Participants

A total of five independent samples were collected from Amazon's Mechanical Turk (mTurk) to test the three different scale versions used in Studies 2, 3, and 4. mTurk serves as an online repository for tasks (called "HITS") that participants (called "workers") complete in exchange for market rate-based payment. The workers on mTurk have been shown to be demographically diverse (Paolacci, Chandler, & Ipeirotis, 2010) and as attentive or more attentive during experiments than the typical university student sample (Hauser & Schwartz, 2014). Participants were individuals 18 years or older, native English speakers, and current U.S. residents. We paid for a completed HIT at the rate that we paid for lab-based experiments, between \$5–\$6 per half hour. Prior to being permitted access to the study, individuals were asked to provide their mTurk worker ID. These IDs were then prescreened to ensure that each sample consisted of unique participants. Workers who had previously participated in a planfulness study were not granted access to the survey, nor was any data recorded from those workers. In total, 813 participants were recruited for Study 2 (407 in A; 406 in B), 814 participants were recruited for Study 3 (416 in A; 398 in B), and 406 participants were recruited for Study 4.

Materials

Each study tested a different version of the Planfulness Scale. Each scale version was similar to the previous, with changes reflecting either the exclusion or slight word alteration of one or more items; these changes were based on the results of previous exploratory and confirmatory analyses as described below. Each version of the scale included items meant to assess the three hypothesized facets of planfulness, and included forward- and reverse-keyed statements. The additional seventeen items generated for Study 3 were primarily developed using the same tactic used to generate the initial pilot pool of items. Several original items were also created for this scale version in an attempt to balance the number of statements per each subscale; for example, the item, “When it comes to achieving my goals, I think of any misstep as a failure,” was developed for the MF subscale during this process. Participants in each sample responded to scale items using a five-point Likert scale (1 = *Strongly disagree*; 3 = *Neither disagree nor agree*; 5 = *Strongly agree*). Each scale version took between ten and twenty minutes to complete. See **Table 1** for a summary of scale version construction and results.

Procedure

All studies were approved by the University of Oregon Research Compliance Services. Each version of the Planfulness Scale was completed online. After electronically consenting to participate in the study, participants in each sample were asked to indicate their agreement with scale items. Data from individuals who failed attention checks during the study were not recorded. The collected data were cleaned of missing observations via listwise deletion prior to analysis. This procedure resulted in the following final sample sizes of 372 (Study 2A), 377 (Study 2B), 257 (Study 3A), 356 (Study 3B), and 373 (Study 4). The items on each version of the scale were examined for psychometric quality in SPSS23. All items, including those identified as having poor psychometrics, were tested in at least two independent samples to lessen the influence of any specific sample’s characteristics on item inclusion/exclusion decisions. Tests of the planfulness model fit were conducted using confirmatory factor analysis (CFA) with a maximum likelihood estimator in R

version 3.1.3 using lavaan, a structural equation modeling package (Rosseel, 2012). These tests used a two-step approach, testing first the measurement model (i.e., item loadings onto theoretical latent constructs) and then the structural model (i.e., including relationships amongst all latent constructs; Anderson & Gerbing, 1988). In the primary model, planfulness was established as a first-order (superordinate) latent variable that explains the three second-order (subordinate) latent variables of Temporal Orientation, Cognitive Strategies, and Mental Flexibility; these latent variables in turn explain each of their indicator scale items. The second-order variables were not allowed to covary directly with each other, so that all of the covariance among them is explained by the first-order variable. Based on the results from the principal components analysis of the pilot data which revealed a distinct factor for reverse-keyed items (see Supplemental Materials), an additional measurement factor (ACQ) was included in the model to capture variance potentially generated by participant acquiescence bias. In the primary model, each indicator item loads on this measurement factor, with loadings fixed to either +1 or -1 for forward- and reverse-keyed statements, respectively; additionally, this factor was set to have no correlation with the first-order planfulness factor, as is necessary for specification in models that include measurement factors (Billiet & McClendon, 2000). **Figure 1** is a visual depiction of this model. Alternative models, including a model without subscale factors, were also tested to determine the relative fit of the primary model (see **Figure 2**). Inspected indices of model fit included the comparative fit index (CFI), root mean square error of approximation (RMSEA), the χ^2 goodness-of-fit statistic, and the Aikake information criterion (AIC); these statistics were compared to determine relative model performance across samples and scale versions.

Results

Scale Psychometrics

As is shown in **Table 1**, Cronbach’s alpha was high for each scale version, ranging between .92 and .95 across all five independent samples. Inter-item correlations were examined closely to identify statements that poorly ($r < .30$) or negatively correlated with others either on the same subscale or the whole scale; if statements

Table 1: Summary of descriptives and psychometrics from studies 2, 3, and 4.

Study, Sample	N	# of Scale Items (reverse-coded)	Cronbach’s α	Mean scale score (standard deviation)	Inter-item r mean (variance)
2, A	372	43 (10)	.94	3.81 (.47)	.26 (.02)
2, B	377		.92	3.76 (.43)	.22 (.02)
3, A	257	60 (20)	.95	3.57 (.47)	.24 (.04)
3, B	356		.95	3.60 (.43)	.22 (.03)
4	373	55 (22)	.95	3.65 (.52)	.27 (.02)

Note: the number of scale items in this table reflects the total number of items per each *study*, but not necessarily the total number of items tested in each *sample*. Often, multiple different items were dropped during exploratory model testing before confirmatory tests were conducted in new, independent samples. Scale scores were calculated by averaging across all items in a given version of the scale, resulting in a possible range of 1–5.

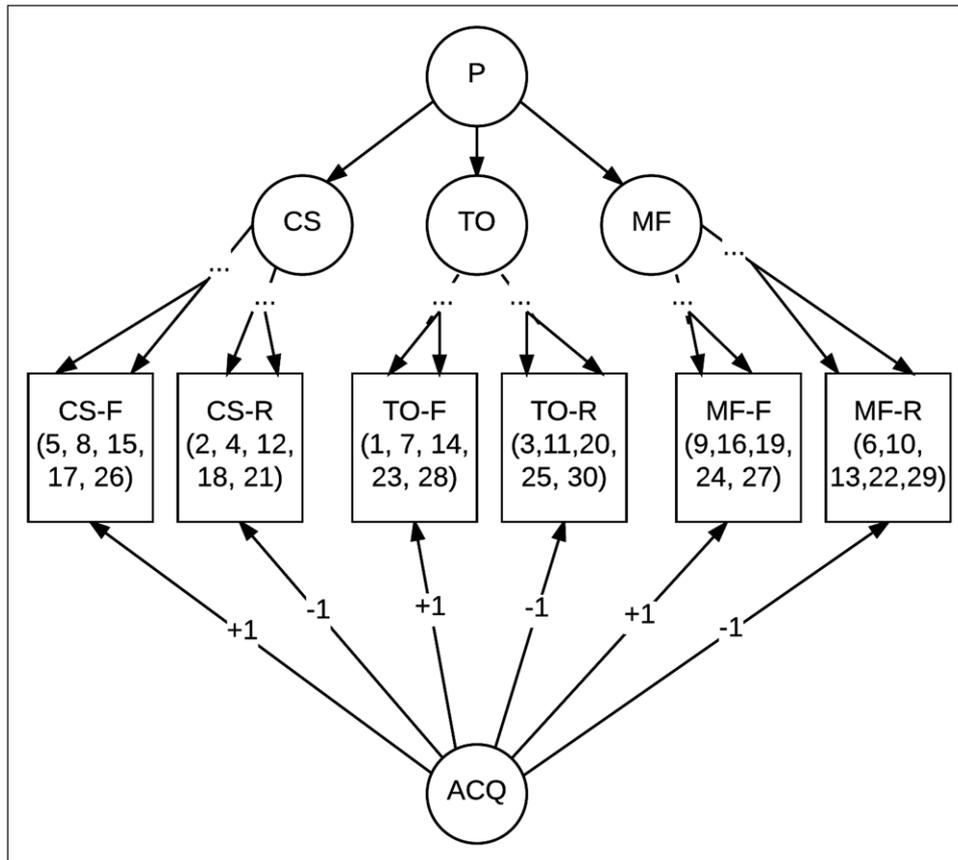


Figure 1: Structural equation model of the Planfulness Scale. P = Planfulness, ACQ = Acquiescence (measurement) factor, CS = Cognitive Strategies, TO = Temporal Orientation, MF = Mental Flexibility, -F = forward-keyed items, -R = reverse-keyed items. Circles indicate latent factors and squares represent the indicator items on the Planfulness Scale. For clarity, like-indicator items are grouped into singular squares here, although each of the 30 total items was run as a single indicator during analyses.

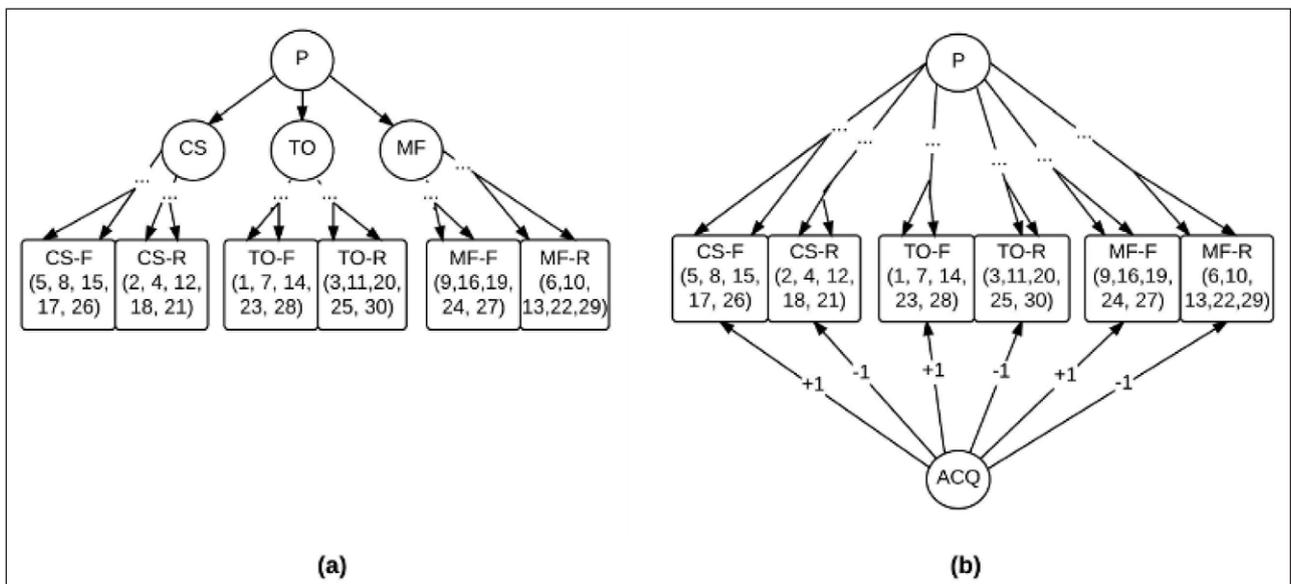


Figure 2: Alternative Planfulness Scale models. Two alternative models that did not significantly fit the data better than the hypothesized primary model. Model a does not include an acquiescence measurement factor; Model b does not include latent subfactors.

identified during this process were confirmed to have poor psychometric quality in a subsequent sample, those items were removed in the following scale version. An

example item identified during this process is the CF statement, “I do not get caught up in thinking about the barriers to my goals”. This statement’s corrected item-total

correlation, a measure of an individual item's correlation with the whole scale, was low in both Study 3 samples ($r = .10$ in A, $r = -.05$ in B), and it negatively correlated with multiple items on the CF subscale in both samples. This process resulted in three items being removed from the scale version in Study 2 to Study 3, and four items removed from the version in Study 3 to the one in Study 4. Finally, scale scores were calculated by averaging all item responses on each version; distributions of these scores were slightly negatively skewed with means close to 4, also shown in **Table 1**.

Confirmatory Factor Analysis

Results of CFAs conducted on each sample's data supported the primary model as comparatively best fitting and revealed incremental improvement in model fit across scale versions. The CFI and RMSEA of the full planfulness model tested in Study 2A were .71 and .081, respectively, and both of these values improved to a CFI of .74 and a RMSEA of .068 for the model tested in Study 4. As expected, in each sample the primary model was determined to be of better fit than the alternative models, as indicated by a reduction in AIC and by a statistically significant change in chi-square test. In all alternative

model tests across all samples the reduction in the chi-square statistic in the primary model was significant at $p < .001$. See **Table 2** for a summary of model test results.

Standardized factor loadings and variances were inspected for the individual scale item indicators as well as the second-order subscale factors. In each sample, the majority of item indicators loaded on their respective subscale latent factors above .30, indicating unidimensionality of the subscales. For example, in Study 3B all but 5 indicators (out of 60 total) loaded at .30 or greater, with an average loading of .50. Of the five with low loadings, two were previously flagged during psychometric evaluation, including the aforementioned item, "I do not get caught up in thinking about the barriers to my goals," which loaded at .013 on the CF subscale. All item variances were significant at $p < .001$ in every sample. Review of the subscale standardized factor loadings and variances indicated a high degree of multicollinearity. In every sample, each subscale loaded at .80 or greater onto the first-order planfulness factor in the primary model. Additionally, while the variances of each of the subscales did reach significance at the .05 level in every sample, the reported p values were surprisingly high (i.e., $\geq .001$) given the large sample sizes. For example,

Table 2: Summary of CFA results of primary and alternative models in studies 2, 3, and 4.

Study, Sample (N)	Model	χ^2 (df)	AIC	CFI	RMSEA [90% CI]
2, A (372)	Primary	2941.9 (856)	36528.11	0.71	.081 [.078, .084]
	Without Acquiescence Factor	3511.1 (857)	37095.23	.63	.091 [.088, .094]
	§Without Planfulness Factor	2941.9 (856)	36528.11	0.71	.081 [.078, .084]
	Without Subscale Factors	3131.1 (859)	36711.24	.68	.084 [.081, .088]
2, B (377)	Primary	2691.1 (856)	37876.05	0.70	.075 [.072, .079]
	Without Acquiescence Factor	3324.1 (857)	38507.02	.60	.087 [.084, .091]
	§Without Planfulness Factor	2689.6 (854)	37878.54	.70	.076 [.072, .079]
	Without Subscale Factors	2872.3 (859)	38051.23	.67	.079 [.076, .082]
3, A (257)	Primary	3916.5 (1706)	35998.36	.71	.071 [.068, .074]
	Without Acquiescence Factor	4514.2 (1707)	36594.01	.64	.080 [.077, .083]
	§Without Planfulness Factor	3916.5 (1706)	35998.36	.71	.071 [.068, .074]
	Without Subscale Factors	4250.5 (1709)	36326.33	.67	.076 [.073, .079]
3, B (356)	Primary	4229.4 (1706)	48448.12	0.72	.064 [.062, .067]
	Without Acquiescence Factor	4820.5 (1707)	49037.15	.66	.072 [.069, .074]
	§Without Planfulness Factor	4229.4 (1706)	48448.12	0.72	.064 [.062, .067]
	Without Subscale Factors	4612 (1709)	48824.72	.68	.069 [.067, .071]
4 (373)	Primary	3719.6 (1373)	45703.68	.74	.068 [.065, .070]
	Without Acquiescence Factor	4083.6 (1374)	46065.66	.70	.073 [.070, .075]
	§Without Planfulness Factor	3719.6 (1373)	45703.68	.74	.068 [.065, .070]
	Without Subscale Factors	4057.1 (1376)	46035.19	.70	.072 [.070, .075]

Note: Across all comparisons the primary model had significantly improved fit over each alternative model, as determined by a change in χ^2 test ($p < .001$) and comparison of the Aikake information criterion (AIC). A § denotes the equivalent model with the same fit indices as the primary model. For a visual depiction of each alternative model, see Figure 2.

in Study Four the reported variance of the TO subscale was 0.14, $p = .010$, and the reported variance of the MF subscale was 0.22, $p = .001$. Finally, the variance of the acquiescence measurement factor was significant at $p < .001$ in every sample.

Discussion

The iterative process of scale development resulted in improved model fit of the Planfulness Scale across five independent samples. Troublesome items were identified during multiple steps of data inspection and were confirmed to be poorly performing in new independent data sets before being removed from the scale. The addition of novel items in Study 3 helped to balance the number of items representing each planfulness facet. Together, these procedures led to improved unidimensionality of each of the subscales and consequently increased the fit of the primary model from Study 2A to Study 4.

The mean scale score distribution was slightly negatively skewed. A possible reason for this is that items that were designed to be more socially desirable but indicate low planfulness were disproportionately cleaned out for poor performance across versions. For example, of the four items removed between Studies 3 and 4, three were intended to sound socially desirable, including the aforementioned CS item. We conclude that acquiescence bias was present in responses to the tested scale versions, leading us to generate new socially desirable items for Study 5 in an attempt to reduce the observed skewness.

The data produced unclear results as to the unidimensionality of the overall scale. The low ($r < .30$) mean inter-item correlation across all scale versions suggests that the scale has multidimensional qualities; this is also supported by the previously reviewed PCA conducted on the pilot scale data in Study 1 (see Supplementary Materials). The CFAs conducted on the primary models in each sample reveal the loadings of all of the second-order facet factors on the first-order planfulness factor to be very high ($>.80$), as were the subscale covariances in the equivalent models. As the observed p -value of the subscale variances in the primary model was also surprisingly high in some samples, these results are possibly indicative of empirical under-identification (Kenny, 1979). Empirical under-identification of the primary model would suggest that the subscales do not contribute enough unique variance to be reliably estimated as distinct parameters, and thus the scale may be unidimensional. In contrast, the results of the series of alternative model tests do not seem to support this conclusion, as the primary model fit the data significantly better than the model that collapsed all subscale factors into one singular planfulness factor. Taking into account the theoretical basis for the hypothesized structure of planfulness, the conclusion to be drawn from these mixed results may be that while there are distinct categories of planfulness behaviors, scores on the Planfulness Scale should be calculated in aggregate rather than by individual facet scores. This conclusion is further tested in Study 5, which includes a new balanced scale version as well as novel model tests.

Study 5: Reliability and Validity of a Balanced Planfulness Scale

The previous studies supported the hypothesized model of planfulness and narrowed down a collection of items that together improved the fit of this model. In Study 5 we sought to generate a balanced scale with equivalent numbers of forward- and reverse-keyed items per each subscale. Balancing scales across forward- and reverse-keyed items have been found to be beneficial for reducing acquiescence bias in participant responses, but such scales might also have effects on the structure of the data by causing the reverse-keyed items to load on a different factor from the forward-keyed ones (Kulas, Klahr, & Knights, 2018; Vautier & Pohl, 2009). We found this pattern with earlier versions of the planfulness scale, which is why we continue to include a measurement factor accounting for the keying of the items.

The balanced scale version was again inspected for psychometric quality, and tested using CFA against alternative models; additionally, given previous concerns about the significance of second-order subscale factor variances, a parceled model was also included to better examine the fit of the hypothesized structural model. We hypothesized that the primary model would continue to fit the data better than alternative models, as shown in Studies 2–4. We further hypothesized that tests would support the ability of each subscale to capture distinguishable patterns of thought about goals, as indicated by significant variance of the subscale factors. Three independent samples, A, B, and C, were collected for model testing using this balanced Planfulness Scale.

The balanced scale also underwent additional validity tests. In Samples A and B, the Planfulness Scale was compared with other related but theoretically distinct scales to test for convergent and discriminant validity. Planfulness was expected to correlate strongly ($r > 0.50$) with measures of conscientiousness or related constructs, such as grit, and to have weaker correlations ($r < .30$) with unrelated measures, such as extraversion. We additionally wanted to test whether planfulness is separable from the expected related constructs. Prior work has suggested that doing so using multiple regression of observed variables can lead to inflated Type I errors (Westfall & Yarkoni, 2016). Therefore, we designed a structural equation model wherein each construct was included as a latent variable explaining a latent factor with questions asking about specific goal-related outcomes as indicators. As this outcome factor was specific to goal achievement, we expected the Planfulness Scale to explain its latent variance beyond that of the other measures. Six comparison scales were included in Sample 5A, and four of these same scales were tested again in 5B.

Finally, longitudinal data were collected in Sample 5C to test the predictive validity of the balanced scale and estimate its test-retest reliability. In this sample, individuals completed the scale and also reported about three specific goals that they wished to achieve in the coming months; measures of planfulness and goal progress were both collected at two timepoints separated by three

months. We hypothesized that the test-retest reliability of planfulness scores would be high, $r > .50$. We additionally hypothesized a strong positive relationship between planfulness at the first timepoint and goal progress at the second timepoint (i.e., change in goal completion in the interim). If supported, the results would reveal the ability of the Planfulness Scale to predict variance in goal achievement as a function of specific patterns of goal-related thought.

Method

Participants

All participants were native English speakers, and data from individuals who failed survey attention checks were not recorded. Participants for Samples A and B were recruited from mTurk using the same methods as described previously, including the additional measure to ensure that each sample consisted of individuals who had not previously taken any version of the Planfulness Scale, and were again compensated at approximately \$5–\$6 dollars per half hour. Sample A included 410 individuals, 51% of whom were male, primarily identified as non-Hispanic Caucasian (79%; the two next largest groups were 7% Asian or Pacific Islander and 6% non-Hispanic Black), and who ranged in age from 18–78 ($M = 36.41$, $SD = 11.22$). There were 411 participants in Sample B; this sample was primarily male (58%) and non-Hispanic Caucasian (81%; 7% non-Hispanic Black, 6% Asian or Pacific Islander), and included individuals ranged in age from 19–84 ($M = 35.46$, $SD = 11.25$). Sample C consisted of 1,192 participants recruited through a nationally representative Qualtrics panel of adults. In this sample, 62% of participants were women, 80% identified as Caucasian non-Hispanic (8% Black non-Hispanic, 6% Hispanic), and participants ranged in age from 19 to 89 years ($M = 47.73$, $SD = 16.64$). All 50 U.S. states were represented by participants in this sample; a choropleth map showing response rate by state is available in Appendix B.

Materials

The balanced Planfulness Scale was tested in all three samples. To decrease potential error induced by respondent social desirability bias, we included items that assessed low planfulness while maintaining high face-valid social desirability, such as, “I prefer my days to be spontaneous rather than scheduled.” We also structured this version to balance forward- and reverse-coded items across each subscale, and to represent each subscale equivalently across total items. The balanced scale therefore had a total of 30 items; 10 total from each subscale, with 5 of each of these items reverse-coded. As with prior versions of the scale, all subscale items were presented in a fixed, interleaved pattern. This version of the scale took approximately seven minutes to complete. The full text of the scale can be found in Appendix A.

Participants in Samples A and B additionally completed the Big Five Inventory (BFI, 44 items; John & Srivastava, 1999), the Brief Self-Control Scale (BSCS, 13 items; Tangney, Baumeister, & Boone, 2004), the Barratt Impulsiveness

Scale (BIS-11, 30 items; Patton, Stanford & Barratt, 1995), and the 12-item Grit scale (Duckworth, Peterson, Matthews, & Kelly, 2007). Sample A participants also completed the Mindful Attention Awareness Scale (MAAS, 15 items; Brown & Ryan, 2003), and the Need for Cognition scale (18 items; Cacioppo, Petty, & Kao, 1984). Finally, Sample B participants completed an ‘Outcomes’ questionnaire. Outcome items were designed to be face-valid questions about specific, tangible goal achievement experiences; for example, one item read, “I meet my deadlines on time”. Participants responded to questionnaire items using a five-point Likert scale (1 = *Strongly disagree*; 3 = *Neither disagree nor agree*; 5 = *Strongly agree*). There were nine outcome items, three of which were reverse-coded; see Appendix D. Each scale took between ten to twenty minutes to complete.

Finally, participants in Sample C answered questions about goals and goal progress at each of two timepoints (adapted from Palfai & Weafer, 2006, based on the approach described in McGregor & Little, 1998). At the first timepoint, they were prompted to, “[p]lease take a few minutes and write down the top three personal projects that you are currently engaging in or considering for the next three months” (see Appendix E for full instructions). Inspection of the projects that participants listed revealed them as varying in size (“plan for retirement,” “plant & grow starter plants”) and specificity (“saving,” “send out 10 job applications”). After recording these projects using three free-response text boxes, they then answered three questions in a fixed order per each goal; these three questions were also asked at timepoint two, after participants viewed and confirmed their verbatim projects, which were ported from timepoint one. The first question, “How committed to this project are you?” and the second question, “How important is this project to you?” were both answered using a five-point Likert scale (1 = *Not at all committed/important*; 3 = *Moderately committed/important*; 5 = *Extremely committed/important*). For the third question, “What percentage of this project is completed right now?” participants responded with percentages ranging from 0–100. The third question allowed us to calculate two dependent measures of goal progress, and the other two questions were included as potential moderators, although we did not include these in the present analyses. If participants did not recognize the three provided projects at the second timepoint as their own, they were asked to re-enter their projects to the best of their ability. They then were subsequently asked to rate how confident they felt about their ability to recall each of those projects from memory on a scale of 1 = *Not at All Confident* to 3 = *Very Confident*.

Two dependent measures of goal progress were calculated. The first dependent measure represented the raw change in percentage complete of a goal from timepoint one at timepoint two, T2%–T1%. While face valid, this index is an inadequate measurement of progress made towards goal completion; for example, a participant who moved from 10% to 20% complete would obtain the same score as another participant who moved from 90% to 100%. Thus, the *a priori* decision was made to calculate a second variable to measure the amount of progress

made as a function of the amount of progress remaining, given by the formula $(T2\% - T1\%) / (100 - T1\%)$. For example, on this measure moving from 10% to 20% yields 11%, whereas moving from 90% to 100% yields 100% of progress towards completion made. These indices are calculated per each goal and then averaged across all goals to form composites of average raw percentage change.

Procedures

The research conducted in Study 5 was approved by the University of Oregon Research Compliance Services prior to commencement. Participants in all samples completed the study online via a Qualtrics survey and provided electronic consent before proceeding to the survey questions. Participants in Samples A and B were presented with all scale measures in a randomized order, except for the Outcomes Questionnaire, which was presented to Sample B participants immediately following consent in order to avoid any carry-over effects from the other measures. If participants did not provide an answer to a survey item, they were prompted with a request to do so; this process resulted in complete data from every individual in both samples. Participants in Sample C were invited to complete the study at two timepoints separated by a three-month period. At the first timepoint, participants were given a fixed order of the Planfulness Scale followed by input of their three goals, and the three questions about those goals. At the start of the second timepoint participants were shown their goals from the first timepoint verbatim and confirmed that they were accurate. If not, participants entered in their three original goals to the best of their recollection. All participants then completed the Planfulness Scale and goal-related questions in randomized order. Of the 1,192 participants who completed the study at the first timepoint, 501 (42%) participants returned to complete the second session. Lists of goals were matched across timepoints, and cases that did not include matching goals at each timepoint were excluded from regression analysis; individuals who had indicated that their goals were 100% complete at timepoint one were also excluded. Finally, cases with missing values were removed with listwise deletion, resulting in a total $N = 457$ (91%) of observations included in tests of predictive validity and test-retest reliability. Model testing and scale psychometrics were performed on the dataset from the first timepoint, cleaned of missing observations using listwise deletion, for a final N of 1,188.

Scale psychometrics were inspected in the same manner as the previous studies using SPSS23. Bivariate correlations of scale scores in Samples A and B were also conducted using SPSS. Model testing was again completed in R version 3.1.3 using lavaan, and included the previously described primary and alternate models. A new, parceled model was also included to address concerns about empirical underidentification uncovered during Studies 2–4. Parceled models do not allow inferences about individual items, but the benefits of using parcels include a reduction of sampling error and decreased likelihood of correlated residuals, both of which can clarify the

structural relationship among latent factors in the model (Little, Cunningham, & Shahar, 2002). We therefore built three parcels for each subscale using a blend of two techniques discussed in Little, Cunningham, and Shahar (2002) in order to increase parcel unidimensionality and test the hypothesized structure of the scale. The TO parcels were built using a domain-representative approach, by splitting all items into categories of ‘past’, ‘present’, and ‘future’ focus, and then dividing these items evenly across parcels. The CS and MF subscales lacked obvious categories in their respective items, and thus parcels for these subscales were constructed using an item-to-construct balance approach. In this approach, a principal components analysis is run on each subscale and items are equivalently distributed across parcels based on their factor loadings, such that each parcel has approximately the same average loading of items on the constructs; three parcels are recommended as the minimum requirement for model identification. The PCA was first run on the largest dataset, Sample C, to determine parcel components, and then these same components were used in Samples A and B’s models (see **Table 4** for parcel items). The model to test whether planfulness could be considered separable from related constructs was similarly constructed. Three parcels each were generated using an item-to-construct balance approach for grit, conscientiousness, impulsivity, self-control, and the Outcomes Questionnaire. These parcels were explained by their respective latent construct variables, and the latent variables of planfulness, grit, impulsivity, self-control, and conscientiousness in turn explained the latent outcomes variable (see **Figure 3**). Finally tests of Sample C’s planfulness test-retest reliability and predictive validity were conducted in R.

Results

Scale Psychometrics

Despite reducing the number of items in the scale, Cronbach’s alpha remained high in all samples, $\alpha \geq .89$. Compared to previous studies, the average inter-item correlation increased while the variance decreased across all samples. Inspection of the corrected item-total correlations in each sample revealed no poorly ($r < .30$) correlating items in Sample A, but revealed the same three items (3, 18, and 28) in Samples B and C as correlating between a range of .19 and .27. Further investigation of these items’ performance on their respective subscales indicated that item 18 correlated well ($r \geq .35$) with the rest of the CS items in both samples; however, items 3 and 28 continued to have low correlation ($r \leq .27$) with the whole TO subscale in both samples. Finally, planfulness scores were calculated by averaging across all item responses, resulting in a possible score range of 1–5. The mean scale scores did decrease on average as compared to the previous studies (e.g., $M = 3.69$ in Sample A here vs. $M = 3.81$ in Study 2A), and mean scores continued to be slightly negatively skewed. A truncated range of scale scores was also observed in each sample, as scores began at approximately 2. See **Table 3** for a summary of all scale psychometrics in each sample.

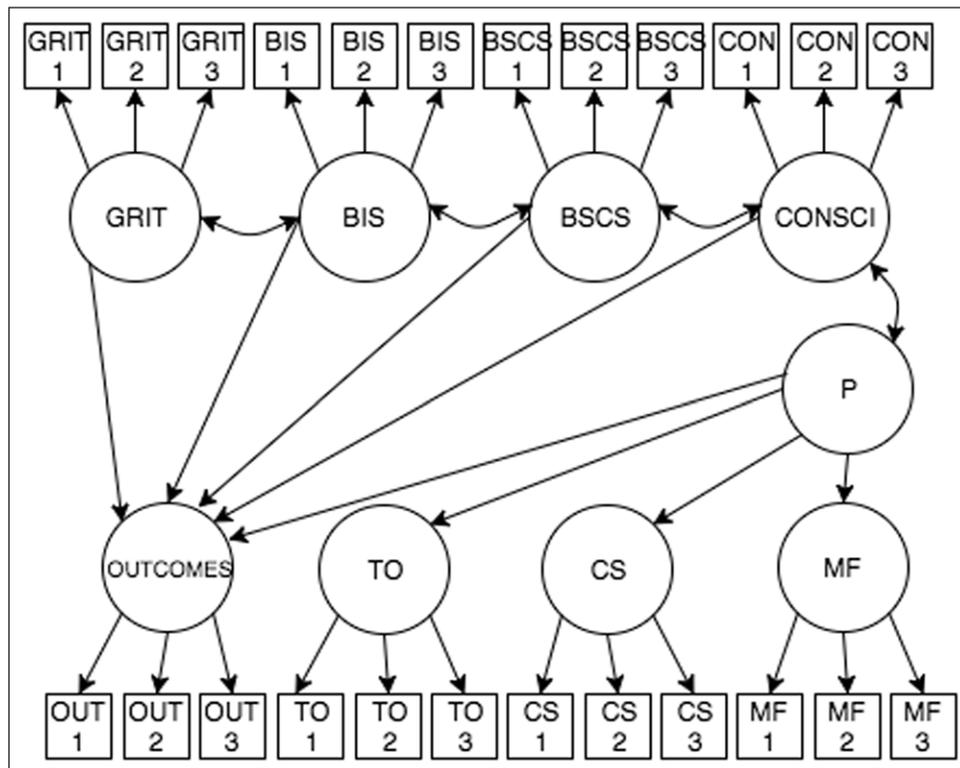


Figure 3: Structural equation model testing separability of constructs. BIS = Barratt Impulsivity Scale, BSCS = Brief Self-control Scale, CONSCI = Conscientiousness facet of the Big Five Inventory, P = Planfulness, TO, CS, MF = Planfulness facets. In this model, all latent construct factors are explaining the variable representing the Outcomes Questionnaire given to Sample 5B. All indicators are parcels constructed using an item-to-construct balance approach, except for the TO parcels which were constructed with a domain-representative method. All construct latent factors are allowed to covary with each other, although only neighbor covariances are depicted here for clarity of the structure.

Confirmatory Factor Analysis

Results of model tests confirmed two of our hypotheses. First, the balanced Planfulness Scale improved in fit over previous scale versions according to various indices (ex: Sample C CFI = .82, RMSEA = .066, see **Table 4** for full model test results). Standardized loadings of individual items on their respective subscales in the primary model remained high in all samples, $M(A) = .57$, $M(B) = .56$, $M(C) = .49$. Second, the primary model fit significantly better than each tested alternative model, across all samples, according to change in chi-square tests (all tests significant at $p < .001$) and smaller values of the AIC. Concerns from the previous studies were confirmed as across samples the variances of at least one of the second-order factors did not reach significance. Which subscale factor variance failed to reach significance differed by sample: Sample A, $\text{var}(CS) p = 0.17$, Sample B, $\text{var}(TO) p = 0.49$, $\text{var}(MF) p = .25$; Sample C, $\text{var}(TO) p = .80$. Parceled models were run to address our concerns about empirical under-identification. The subscale variances reached statistical significance in the parceled models in all three samples, supporting multidimensionality of the subscales; the difference between the parceled and non-parceled models suggests item variance obscures multidimensionality. In Samples A and B the variance of the CS subscale was significant at $p = .011$, while all variances were significant in Sample C at $p < .001$. Correlations amongst the subscales were very high in all samples, ranging from .92 to .96.

Convergent and Discriminant Validity

All scores for the comparison scales were calculated according to their original instructions, resulting in a final range of scores from 1–4 for the BIS-11, 1–6 for the MAAS, and 1–5 for all other included scales. As hypothesized, the correlation of Planfulness scores with other scale scores tracked closely to those of the conscientiousness scores, correlating strongly with related measures and exhibiting weak correlations with unrelated metrics in both Samples A and B (see **Table 5**). Specifically, scores on the Planfulness Scale correlated the most highly with Grit ($r = .76$ in both samples) and the BIS-11 ($r = -.74$ in A, $r = -.75$ in B), and the lowest with extraversion ($r = .22$ in A, $r = .26$ in B) and openness ($r = .27$ in A, $r = .24$ in B). Planfulness scores correlated highly with conscientiousness itself, as expected, $r = .71$ in A, $r = .75$ in B.

Separability of Constructs

Responses to the Outcomes Questionnaire were gathered in Sample B to test the incremental validity of the Planfulness scale in predicting goal outcomes above and beyond related constructs. Scores on the Outcomes Questionnaire were averaged across all nine items (reverse-coding three), resulting in a range of 1–5, where five indicates a strong self-reported ability to achieve set goals. In Sample B, the mean Outcomes score was 3.76 ($SD = 0.62$) and the Cronbach’s alpha for the nine items was .87. The items were entered into a formative model wherein all items loaded on one general Outcomes construct; this was then

Table 3: Summary of descriptives and psychometrics from Study 5 samples.

Sample (N)	Scale	Cronbach's α (# of items)	Mean scale score (standard deviation)	Inter-item r mean (variance)
A (410)	Planfulness	.93 (30)	3.69 (.52)	.31 (.01)
	Grit	.90 (12)	3.55 (.73)	.44 (.02)
	BIS-11	.88 (29)	1.88 (.37)	.21 (.02)
	Extraversion	.91 (8)	2.82 (1.0)	.55 (.02)
	Agreeableness	.87 (9)	3.78 (.79)	.42 (.01)
	Conscientiousness	.90 (9)	3.95 (.76)	.51 (.01)
	Neuroticism	.92 (8)	2.67 (1.0)	.58 (.01)
	Openness	.85 (10)	3.73 (.69)	.38 (.02)
	BSCS	.91 (13)	3.49 (.80)	.44 (.01)
	NforCognition	.95 (17)	3.31 (.83)	.55 (.01)
	MAAS	.93 (15)	4.26 (.94)	.46 (.01)
B (411)	Planfulness	.92 (30)	3.62 (.54)	.29 (.02)
	Grit	.90 (12)	3.51 (.71)	.43 (.02)
	BIS-11	.88 (29)	1.89 (.38)	.21 (.03)
	Extraversion	.91 (8)	2.83 (.98)	.56 (.01)
	Agreeableness	.85 (9)	3.77 (.74)	.39 (.01)
	Conscientiousness	.89 (9)	3.86 (.78)	.50 (.01)
	Neuroticism	.92 (8)	2.68 (1.0)	.60 (.01)
	Openness	.85 (10)	3.71 (.69)	.37 (.02)
	BSCS	.90 (13)	3.44 (.78)	.42 (.01)
C (1,188)	Planfulness	.89 (30)	3.61 (.44)	.23 (.02)

Note: Sample C data is from the first timepoint. Scale scores ranged from 1–4 on the BIS-11, 1–6 on the MAAS, and 1–5 for all other scales.

Table 4: Summary of CFA results of primary and alternative models in Study 5 samples.

Sample (N)	Model	χ^2 (df)	AIC	CFI	RMSEA [90% CI]
A (410)	Primary	1388.73 (401)	28207.53	.80	.078 [.073, .082]
	Without Acquiescence Factor	1640.50 (402)	28457.3	.76	.087 [.082, .091]
	Without Subscale Factors	1452.68 (404)	28265.48	.79	.080 [.075, .084]
	Parceled	100.45 (24)	5102.23	.97	.088 [.071, .106]
B (411)	Primary	1272.05 (401)	29285.71	.83	.073 [.068, .077]
	Without Acquiescence Factor	1747.96 (402)	29759.62	.74	.090 [.086, .095]
	Without Subscale Factors	1315.74 (404)	29323.4	.82	.074 [.070, .079]
	Parceled	111.91 (24)	5290.46	.96	.094 [.077, .112]
C (1,188)	Primary	2453.16 (401)	82645.69	.82	.066 [.063, .068]
	Without Acquiescence Factor	3980.66 (402)	84171.19	.68	.087 [.084, .089]
	Without Subscale Factors	2547.22 (404)	82733.75	.81	.067 [.064, .069]
	Parceled	200.45 (24)	14567.54	.96	.079 [.069, .089]

Note: Across all comparisons the primary model had significantly improved fit over each of the two alternative models, as determined by a change in χ^2 test ($p < .001$) and comparison of the Aikake information criterion (AIC). All parceled models consisted of the same component items per each subscale, TO (1, 7, 25; 3, 23, 30; 11, 14, 28, 20), CS (8, 2, 15; 5, 18, 21; 17, 4, 12, 26), and MF (13, 9, 22; 24, 29, 10; 6, 16, 19, 27).

Table 5: Correlation matrix of measurements collected in Study Five, Samples A and B.

	1	2	3	4	5	6	7	8	9	10
	Planfulness	Grit	BIS-11	BFI – E	BFI – A	BFI – C	BFI – N	BFI – O	BSCS	Need for Cognition
1 Planfulness		.76	-.75	.26	.37	.75	-.51	.24	.69	
2 Grit		.76	-.73	.32	.39	.82	-.58	.16	.77	
3 BIS-11		-.74	-.75	-.18	-.36	-.74	.52	-.21	-.77	
4 BFI – E		.22	.35	-.16	.41	.33	-.46	.27	.29	
5 BFI – A		.38	.48	.42	.33	.47	-.51	.23	.44	
6 BFI – C		.71	.77	-.73	.31	.48	-.61	.21	.79	
7 BFI – N		-.44	-.62	.47	-.48	-.54	-.57	-.17	-.60	
8 BFI – O		.27	.29	-.27	.3	.24	-.24		.18	
9 BSCS		.63	.73	-.72	.26	.49	-.56	.17		
10 NforCognition		.43	.46	-.47	.28	.21	-.38	.61	.31	
11 MAAS		.53	.69	-.66	.23	.48	-.57	.28	.63	.38

Note: All correlations are significant at the 0.01 level (2-tailed; S5A $N = 410$, S5B $N = 411$). Numbers above the diagonal represent S5B, below represent S5A data. BIS-11 = Barratt Impulsivity Scale, BSCS = the Brief Self-Control Scale, MAAS = Mindful Attention Awareness Scale, and BFI = the Big Five Inventory and traits of extraversion, agreeableness, conscientiousness, neuroticism, and openness. Higher scores on the BIS-11 are associated with greater impulsivity.

Table 6: Parameter estimates of factor loadings and residuals for Outcomes Questionnaire items on a shared latent factor.

Variables	Factor Loadings			Measurement Errors		
	Unstd.	SE	Std.	Unstd.	SE	Std.
Outcomes						
Q1	1	–	.82	.27	.03	.33
Q2 (REV)	.95	.07	.67	.61	.05	.55
Q3 (REV)	.85	.08	.54	.98	.07	.55
Q4	.68	.05	.63	.39	.03	.60
Q5	.48	.05	.49	.40	.03	.76
Q6 (REV)	.45	.06	.37	.71	.05	.86
Q7	.74	.05	.68	.36	.03	.54
Q8	.78	.05	.66	.40	.03	.56
Q9	.93	.05	.85	.18	.02	.28

Note: Unstd. = Unstandardized; std. = standardized, REV = reverse-keyed item. Model fit statistics are as follows: χ^2 (df) = 445.12 (27), CFI = 0.76, RMSEA = .19 [0.18, 0.21]. All loadings and variances of variables are significant at $p < 0.00$.

tested using confirmatory factor analysis with a maximum likelihood estimator. The items loaded moderately (standardized coefficient $>.35$) to highly ($>.80$) on the latent Outcome factor, and all loadings and variances were significant. For a full report of item loadings, see **Table 6**.

The scores for the Planfulness scale and the comparison measures, Grit, the BIS-11, the BSCS, and the conscientiousness facet of the BFI, were calculated as previously noted and their descriptives are shown in **Table 3**.

Results of a principal components analysis conducted for each scale measurement resulted in the construction of three parcels balanced across factor loadings (see **Table 7a** for parcel details). These parcels were then included in a structural equation model with their respective latent constructs, and all latent constructs were modeled as predictors of the Outcomes variable (**Figure 3**). Testing this model supported our hypothesis. Latent Planfulness significantly predicted variation in Outcomes while controlling for latent Grit, Impulsivity, Self-control, and Conscientiousness, with the standardized coefficient for latent Planfulness = .51, $p < .001$. Furthermore, though conscientiousness was the only other construct that was significantly related to Outcomes, this relationship was weaker than that of planfulness (Beta = .37, $p = .01$), although comparison of this model with one wherein the paths from Planfulness and Conscientiousness to Outcomes were set to be equivalent suggests that this difference did not reach statistical significance, change in chi-square $p = .13$. Overall, the variance of each Outcome parcel was $\sim.30$, meaning that roughly 70% of the variance in each Outcome parcel was accounted for by the other latent variables in the model. Finally, a high degree of multicollinearity ($>.7$) was again observed amongst the related constructs; see **Tables 7a** and **7b**.

Test-retest Reliability and Predictive Validity

Scale scores computed from each timepoint in Sample C were compared with a Pearson correlation. The balanced scale was found to have high test-retest reliability, $r(457) = .81$, 95% CI = (.77, .84). To test for predictive validity, two interrelated indices of goal progress were constructed *a priori* from the third goal-related question about goal percent completion:

Table 7a: Parameter estimates of factor loadings and residuals for a model testing separability of constructs in Sample 5B.

Variables	Factor Loadings			<i>p</i>	Measurement Errors		
	Unstd.	SE	Std.		Unstd.	SE	Std.
Planfulness							
Outcomes	0.50	0.10	0.51	0.00			
MF	1.00	–	1.00	–	0.00	0.01	0.00
TO	0.67	0.04	0.86	0.00	0.06	0.01	0.19
CS	0.85	0.04	0.93	0.00	0.04	0.01	0.13
Grit							
Outcomes	0.01	0.14	0.01	0.92			
Parcel 1	1.00	–	0.88	–	0.12	0.01	0.23
Parcel 2	1.06	0.04	0.89	0.00	0.12	0.01	0.21
Parcel 3	1.04	0.05	0.81	0.00	0.24	0.02	0.34
Conscientiousness							
Outcomes	0.27	0.11	0.37	0.01			
Parcel 1	1.00	–	0.86	–	0.24	0.02	0.26
Parcel 2	0.66	0.03	0.78	0.00	0.19	0.02	0.39
Parcel 3	1.00	0.04	0.90	0.00	0.17	0.02	0.20
Impulsivity (BIS-11)							
Outcomes	0.06	0.15	0.04	0.67			
Parcel 1	1.00	–	0.81	–	0.06	0.01	0.35
Parcel 2	1.15	0.05	0.93	0.00	0.03	0.00	0.14
Parcel 3	1.10	0.06	0.85	0.00	0.06	0.01	0.29
Self-control (BSCS)							
Outcomes	0.07	0.09	0.08	0.43			
Parcel 1	1.00	–	0.87	–	0.16	0.02	0.25
Parcel 2	1.09	0.05	0.88	0.00	0.16	0.02	0.22
Parcel 3	1.13	0.05	0.88	0.00	0.18	0.02	0.22
Outcomes							
Parcel 1	1.00	–	0.82	–	0.18	0.02	0.33
Parcel 2	0.97	0.05	0.88	0.00	0.10	0.01	0.23
Parcel 3	0.95	0.05	0.86	0.00	0.11	0.01	0.26

Note: Unstd. = Unstandardized; std. = standardized. Model fit statistics are as follows: χ^2 (df) = 871.17(234), CFI = 0.93, RMSEA = .081 [0.076, 0.087]. All variances of variables are significant at $p < 0.00$, except for an observed moderate Heywood case for the MF latent variable, $p = .89$. Parcels contain the following items: Grit P1 = 6,8,11,12, P2 = 1,2,9,10, P3 = 3,4,5,7; BIS-11 P1 = 4,5,7,13,17,19,21,25,27,30, P2 = 1,3,8,9,10,14,15,16, 22, 28, P3 = 2,6,11,12,18,20,23,24,26,29; BSCS P1 = 1,4,6,10,12, P2 = 2,5,8,1, P3 = 3,7,9,13; BFI-C P1 = 8,18,28, P2 = 3,13,38, P3 = 23,33,43; Outcomes P1 = 3,8,9, P2 = 2,5,7, P3 = 1,4,6.

average raw change in goal percent completion, which ranged from –64% to 96.67% ($M = 17.02$, $SD = 24.62$), and average percent progress from timepoint one, which ranged from –332% to 100% ($M = 16.12$, $SD = 57.80$). A Pearson correlation was also used to examine the correlation between each goal's percent completion at baseline and endpoint, which was on average $r = .41$ (Goal 1 $r = .41$, Goal 2 $r = .32$, Goal 3 $r = .48$). The composites were then entered in a regression with timepoint one planfulness score as

the predictor. Results showed no significant association between planfulness score and average raw change in goal percent completion, $b = 4.55$, $t(453) = 1.70$, $p = .09$. However, as hypothesized, Planfulness score significantly predicted progress toward average goal completion, $b = 12.30$, $t(453) = 2.07$, $p = .04$. In other words, each one-point increase on the Planfulness Scale was associated with 12% more goal progress to completion, averaging across all goals reported at the first timepoint.

Table 7b: Parameter estimates of factor variances and covariances for a model testing separability of constructs in Sample 5B.

Latent Variables	Variances & Covariances		
	Unstandardized	SE	Standardized
Planfulness	0.38	0.04	1.00
Grit	0.35	0.03	0.89
Conscientiousness	0.44	0.04	0.85
Impulsivity	-0.17	0.02	-0.82
Self-control	0.335	0.03	0.78
Grit	0.41	0.04	1.00
Conscientiousness	0.5	0.04	0.93
Impulsivity	-0.18	0.02	-0.8
Self-control	0.38	0.03	0.85
Conscientiousness	0.69	0.06	1.00
Impulsivity	-0.24	0.02	-0.82
Self-control	0.52	0.05	0.89
Impulsivity	0.12	0.01	1.00
Self-control	-0.21	0.02	-0.86
Outcomes	0.07	0.01	0.19

Note: Estimates are from the full model of separability of constructs depicted in Figure 3. Model fit statistics are as follows: χ^2 (df) = 871.17(234), CFI = 0.93, RMSEA = .081 [0.076, 0.087].

Discussion

The balanced Planfulness Scale was the strongest version of the scale tested. Psychometric analyses indicated that even with a reduced number of items, internal consistency remained high, as did other metrics of unidimensionality. The distribution of scale scores appeared more normal for the balanced scale, although negative skew was still evident in the high observed mean and truncated range. However, negative skew was also evident in the distributions of several other measurement scores collected in Samples A and B, which may suggest that the results were reflective of the characteristics of the samples rather than of a weakness of the balanced scale. Analyses also identified two items on the TO facet (3 and 8) that correlated weakly with the full scale as well as other TO items. These results may be a function of the reduced subscale item size enhancing a multifaceted nature of TO. It has been previously shown that time perspective is a multidimensional construct (Zimbardo & Boyd, 1999), and a clear delineation of past, present, and future-focused items was evident when constructing TO parcels for model testing. Therefore, we believe that the low inter-item correlations among questions addressing different timeframes in the TO subscale are to be theoretically expected given the distinct temporal reference points.

The balanced scale was also found to be the best fitting model yet according to various fit indices. As with previous versions, the primary planfulness model was shown to be

of significantly better fit than alternate models. However, at least one subscale variance failed to reach significance across all three samples, making it unclear whether each subscale was contributing unique information to the model. Parceled models were constructed to test the dimensionality of the scale by focusing on the structural relationship amongst latent factors. Across all samples the variances of the subscales reached statistical significance when parcelation was used, although correlations amongst them remained high. These results further support the previous conclusion; namely, that while there are three distinct facets to planfulness, the Planfulness Scale should be considered in its entirety to represent a single construct, and not scored only according to its subscales.

The balanced Planfulness Scale was compared against other validated measurements and tested for its association with real-world achievement outcomes. Planfulness scores and BFI conscientiousness scores were found to share patterns of correlations with other measurements, including an exhibited inverse relationship with the BIS-11. While it may not be surprising that highly conscientious and planful individuals also self-report as being low in impulsivity, this association does not necessarily mean that the two constructs are simply inversely related. As discussed in the introduction, it may be that being highly planful can moderate or mask the effect of certain types of impulsivity in the context of goal pursuit. For example, part of being planful could be to anticipate contexts in which one is likely to make potentially maladaptive spur-of-the-moment decisions and to take steps to avoid those situations to mitigate the effect of impulsiveness.

The results from Samples A and B are evidence of the convergent and discriminant validity of the balanced Planfulness Scale and support the hypothesis that planfulness falls within the conscientiousness domain. A test of incremental predictive validity further supported the ability of the Planfulness Scale to explain unique variance in self-reported general ability to achieve goals, above and beyond other measures within this domain. This incremental predictive effect was observed even in the presence of a high degree of collinearity with related constructs such as grit and conscientiousness. Given that this test was only conducted in one sample and relied on comparing self-report measures, future tests of construct validity will need to focus on using non-self-report measures to further untangle the relationship between planfulness and other constructs.

The longitudinal component of this study additionally supports the Planfulness Scale as being reliable over time, as scale scores separated by a period of three months were found to be highly correlated. Critically, scale scores from the first timepoint were prospectively predictive of idiographic real-world goal progress. Given that it is not uncommon for people to make progress towards a goal and then either get stuck or abandon it altogether, we calculated *a priori* a variable that weighted achievement of a goal more heavily than raw change in reported progress toward that goal. Indeed, we found that the Planfulness Scale was not predictive of raw goal completion change, but

instead predicted goal progress *as a function of the amount left to make until completion*, as hypothesized. Importantly, Planfulness scores represent the degree to which individuals tend to engage in specific mental processes that have been previously supported as effective for goal achievement in laboratory research. Our longitudinal results therefore show that individuals who report engaging in these patterns of thought are more likely to achieve their goals— not just make incremental progress towards them— in the real world.

General Discussion

We proposed the construct of planfulness to build upon insights from parallel lines of inquiry on traits related to achievement, from the personality field, and mental processes related to goal success, from the self-regulation field. We defined planfulness as the tendency to engage in specific mental processes with respect to goals, which we hypothesized would explain between-person variability in effective goal-related cognition. We articulated the theoretical rationale behind the construct, and developed a scale to measure it in the general population using an iterative process of item testing and guided adjustment across five studies and ten independent samples. This process improved each subsequent scale version as indicated by psychometrics and fit indices, culminating in the balanced scale version tested in Study 5. Further tests of reliability and validity supported the balanced Planfulness Scale as reliable and valid metric for assessing individuals' thoughts about their goals and predicting real-world outcomes.

Planfulness offers a unique and specific perspective on goals as the construct provides a bridge between two broad areas of research. The study of goal setting and achievement in personality psychology is trait-oriented, describing broad phenotypic characteristics of individuals that are stable across time and circumstances (e.g., conscientiousness). In contrast, the laboratory research on goal achievement is process-oriented, focusing on the specific mental strategies that are likely to encourage goal achievement in the general population. These two research traditions investigate the same general topic with divergent methods and assumptions, and as such, tend not to directly interact or inform one another. Conceptualizing planfulness as a stable individual difference in the tendency to deploy specific mental processes joins these two theoretical worlds, and developing a valid and reliable measure of the construct provides a research instrument that can be beneficial to both. More generally, our approach to studying planfulness might serve as a model for conceptualizing how specific cognitions or behaviors can accumulate across time and manifest as an observable "trait", thereby allowing for longitudinal predictions at an individual level based on brief mental processes.

One seemingly paradoxical implication of this hybrid perspective of goal achievement is that we argue for the existence of stable individual differences in goal-related cognition, and also that the processes can be altered within a person. The present study provides evidence for stability over time in the tendency to use these strategies, and at the same time there is ample evidence that these

same patterns of thought about goals can be manipulated in a laboratory setting. Together, these observations suggest that people tend to deploy similar patterns of thought about goals, but that those patterns of thought are not so entrenched that they cannot be influenced by brief manipulations. It might be that people are simply unaware of alternative ways of thinking about goals than the ones that they have used previously.

The Planfulness Scale was constructed to capture individual differences in three specific processes that are known to reliably improve the likelihood of goal achievement. The scale items were selected to tap these processes specifically, so the scope of the scale is deliberately circumscribed around temporal orientation, mental flexibility, and cognitive strategies. As noted in the introduction, we theorize that additional facets could be added to the extent that additional cognitive processes are identified that reliably increase goal progress beyond these initial three facets. However, the ability of the Planfulness Scale to predict goal progress with only the three facets included suggests that the extant scale, on its own, explains meaningful variance in goal-related behaviors.

Limitations and Future Directions

There is an inherent limitation to studying planfulness that may have impacted the results reported here. Being high in planfulness involves a degree of goal meta-cognition (the ability to think about how one thinks about goals), as indicated by items such as, "I can easily identify why I have not achieved goals in the past." It may be that those low in planfulness do not regularly reflect on how they approach goals, or are less accurate in recognizing the specific steps they take that lead to achievement success or failure. In turn, these individuals may then overestimate the degree to which they engage in planful thoughts or behaviors, exhibiting the so-called Dunning-Kruger effect (Kruger & Dunning, 1999). This might introduce heterogeneity to how a group of individuals use the scale in addition to other sources of noise typical of self-report measures. If this effect indeed is occurring in low planfulness individuals, it could help to explain why the distribution of scores from Study 5 is approaching a normal shape, but with a truncated range (histogram in Appendix C).

The fact that the Planfulness Scale is a significant predictor of real-world goal progress somewhat mitigates the aforementioned concern, however goal progress itself was measured through self-report items in this research. This is a second limitation of the work presented here—we did not collect an objective, behavioral measure of goal progress. It may be then that our results represent a strong correlation between thinking about goals and thinking that progress has been made, regardless of the objective truth. Future research should focus on validation of the scale using real-world goals that can be objectively measured in order to explore this possibility.

Another limitation of this research is the lack of a direct comparison between the Planfulness Scale and measures of conscientiousness in their ability to predict objective measures of goal progress. We introduce planfulness here as a way of shifting the focus of trait-level measurements

away from the structure of personality and toward processes that result in habitual thoughts and behaviors. We provide initial evidence that this conceptualization exhibits reliability and validity in concordance with traditional personality measurements, as well as evidence based on self-report measurements that planfulness explains distinct variance in goal outcomes from related measures within the conscientiousness domain. Planfulness and conscientiousness are not theoretical “competitors” because planfulness is theorized as a set of processes in the conscientiousness domain. However, researchers and practitioners may nevertheless want to know as a practical matter which scale provides better or stronger predictive validity or whether the Planfulness Scale has incremental validity over existing Big Five scales. Having moved through the first of two phases in the scale construction model suggested by Simms (2008), an important next step to take following this work in the external validity phase is to thoroughly inspect whether the Planfulness Scale is able to capture information in a unique way from other measures of related traits using objective measurements of behavior. Future work will determine the relative performance of the Planfulness Scale, but we believe we have provided evidence here of its stand-alone value.

Conclusion

Having undergone thorough testing in multiple unique samples, the Planfulness Scale contains items representing three categories of goal cognition that contribute overall to general propensity for goal achievement and success. An initial test of predictive validity suggests that planfulness is positively correlated with progress on real-world goals across time. This initial scale construction and validation process should be followed up with tests of the Planfulness Scale’s ability to predict goal outcomes that are not self-report.

Data Accessibility Statement

De-identified data and analysis scripts are available for download on the Planfulness Scale Open Science Framework page: <https://osf.io/56ja2/>.

Additional Files

The additional files for this article can be found as follows:

- **Appendix A.** Balanced Planfulness Scale from Study Five. DOI: <https://doi.org/10.1525/collabra.136.s1>
- **Appendix B.** Study Five State Response Frequencies. DOI: <https://doi.org/10.1525/collabra.136.s1>
- **Appendix C.** Histogram of Planfulness scores from Study Five. DOI: <https://doi.org/10.1525/collabra.136.s1>
- **Appendix D.** Outcome Questionnaire used in Study 5B. DOI: <https://doi.org/10.1525/collabra.136.s1>
- **Appendix E.** Instructions for listing goals used in Study 5C. DOI: <https://doi.org/10.1525/collabra.136.s1>
- **Study 1.** Initial Planfulness Scale Development. DOI: <https://doi.org/10.1525/collabra.136.s2>

Funding Information

This work was supported by NIH Grants CA175241 and CA211224 to Elliot T. Berkman.

Competing Interests

The authors have no competing interests to declare.

Author Contributions

Rita M. Ludwig contributed to theory development, collected and analyzed the data, and contributed to writing the manuscript. Sanjay Srivastava and Elliot T. Berkman contributed to theory development and writing the manuscript.

References

- Anderson, J. C., & Gerbing, D. W.** (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411–423. DOI: <https://doi.org/10.1037/0033-2909.103.3.411>
- Billiet, J. B., & McClendon, M. J.** (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling, 7*(4), 608–628. DOI: https://doi.org/10.1207/S15328007SEM0704_5
- Bogg, T., & Roberts, B. W.** (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin, 130*(6), 887–919. DOI: <https://doi.org/10.1037/0033-2909.130.6.887>
- Brown, K. W., & Ryan, R. M.** (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*, 822–848. DOI: <https://doi.org/10.1037/0022-3514.84.4.822>
- Cacioppo, J. T., Petty, R. E., & Kao, C. F.** (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306–307. DOI: https://doi.org/10.1207/s15327752jpa4803_13
- Chamorro-Premuzic, T., & Furnham, A.** (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality, 37*(4), 319–338. DOI: [https://doi.org/10.1016/S0092-6566\(02\)00578-0](https://doi.org/10.1016/S0092-6566(02)00578-0)
- Costa, P. T., McCrae, R. R., & Dye, D. A.** (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO Personality Inventory. *Personality and Individual Differences, 12*(9), 887–898. DOI: [https://doi.org/10.1016/0191-8869\(91\)90177-D](https://doi.org/10.1016/0191-8869(91)90177-D)
- Credé, M., Tynan, M. C., & Harms, P. D.** (2016). Much Ado About Grit: A Meta-Analytic Synthesis of the Grit Literature. *Journal of Personality and Social Psychology*. Advance online publication. DOI: <https://doi.org/10.1037/pspp0000102>
- Duckworth, A. L., & Kern, M. L.** (2011). Journal of Research in Personality. *Journal of Research in Personality, 1*–10. DOI: <https://doi.org/10.1016/j.jrp.2011.02.004>

- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R.** (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. DOI: <https://doi.org/10.1037/0022-3514.92.6.1087>
- Dudley, N. M., Orvis, K. A., Lebiecki, J. E., & Cortina, J. M.** (2006). A meta-analytic investigation of conscientiousness in the prediction of job performance: Examining the intercorrelations and the incremental validity of narrow traits. *Journal of Applied Psychology, 91*(1), 40–57. DOI: <https://doi.org/10.1037/0021-9010.91.1.40>
- Fleeson, W.** (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*(6), 1011–1027. DOI: <https://doi.org/10.1037/0022-3514.80.6.1011>
- Frese, M., Stewart, J., & Hannover, B.** (1987). Goal orientation and planfulness: Action styles as personality concepts. *Journal of Personality and Social Psychology, 52*(6), 1182. DOI: <https://doi.org/10.1037/0022-3514.52.6.1182>
- Friese, M., & Hofmann, W.** (2009). Control me or I will control you: Impulses, trait self-control, and the guidance of behavior. *Journal of Research in Personality, 43*(5), 795–805. DOI: <https://doi.org/10.1016/j.jrp.2009.07.004>
- Gollwitzer, P. M.** (1993). Goal achievement: The role of intentions. *European Review of Social Psychology, 4*(1), 141–185. DOI: <https://doi.org/10.1080/14792779343000059>
- Gollwitzer, P. M.** (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*(7), 493–503. DOI: <https://doi.org/10.1037/0003-066X.54.7.493>
- Hill, P. L., & Roberts, B. W.** (2011). The role of adherence in the relationship between conscientiousness and perceived health. *Health Psychology, 30*(6), 797–804. DOI: <https://doi.org/10.1037/a0023860>
- Hofmann, W., Deutsch, R., Lancaster, K., & Banaji, M. R.** (2010). Cooling the heat of temptation: Mental self-control and the automatic evaluation of tempting stimuli. *European Journal of Social Psychology, 40*(1), 17–25. DOI: <https://doi.org/10.1002/ejsp.708>
- John, O. P., Naumann, L. P., & Soto, C. J.** (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In: John, O. P., Robins, R. W., & Pervin, L. A. (Eds.), *Handbook of Personality: Theory and Research*, 114–158. New York, NY: Guilford Press.
- John, O. P., & Srivastava, S.** (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research, 2*, 102–138.
- Kenny, D. A.** (1979). *Correlation and Causality*. New York: Wiley-Interscience.
- Kruger, J., & Dunning, D.** (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134. DOI: <https://doi.org/10.1037/0022-3514.77.6.1121>
- Kulas, J., Klahr, R., & Knights, L.** (2018). Confound it! Social Desirability and the “Reverse-Scoring” Method Effect. *European Journal of Psychological Assessment*. DOI: <https://doi.org/10.1027/1015-5759/a000459>
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F.** (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*(2), 151–173. DOI: https://doi.org/10.1207/S15328007SEM0902_1
- MacCann, C., Duckworth, A. L., & Roberts, R. D.** (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences, 19*(4), 451–458. DOI: <https://doi.org/10.1016/j.lindif.2009.03.007>
- McGregor, I., & Little, B. R.** (1998). Personal projects, happiness, and meaning: On doing well and being yourself. *Journal of Personality and Social Psychology, 74*(2), 494–512. DOI: <https://doi.org/10.1037/0022-3514.74.2.494>
- Noftle, E. E., & Robins, R. W.** (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology, 93*(1), 116–130. DOI: <https://doi.org/10.1037/0022-3514.93.1.116>
- Oettingen, G.** (2000). Expectancy effects on behavior depend on self-regulatory thought. *Social Cognition, 18*(2), 101–129. DOI: <https://doi.org/10.1521/soco.2000.18.2.101>
- Orbell, S., & Sheeran, P.** (2000). Motivational and Volitional Processes in Action Initiation: A Field Study of the Role of Implementation Intentions. *Journal of Applied Social Psychology, 30*(4), 780–797. DOI: <https://doi.org/10.1111/j.1559-1816.2000.tb02823.x>
- Palfai, T. P., & Weafer, J.** (2006). College student drinking and meaning in the pursuit of life goals. *Psychology of Addictive Behaviors, 20*(2), 131–134. DOI: <https://doi.org/10.1037/0893-164X.20.2.131>
- Paolacci, G., Chandler, J., & Ipeirotis, P. G.** (2010). Running experiments on amazon mechanical turk. *Judgment and Decision making, 5*(5), 411–419.
- Patton, J. H., & Stanford, M. S.** (1995). Factor structure of the Barratt impulsiveness scale. *Journal of Clinical Psychology, 51*(6), 768–774. DOI: [https://doi.org/10.1002/1097-4679\(199511\)51:6<768::AID-JCLP2270510607>3.0.CO;2-1](https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1)
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R.** (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science, 2*(4), 313–345. DOI: <https://doi.org/10.1111/j.1745-6916.2007.00047.x>
- Roberts, B. W., Walton, K. E., & Bogg, T.** (2005). Conscientiousness and health across the life course. *Review of General Psychology, 9*(2), 156–168. DOI: <https://doi.org/10.1037/1089-2680.9.2.156>

- Rossee, Y.** (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36. DOI: <https://doi.org/10.18637/jss.v048.i02>
- Saucier, G., & Srivastava, S.** (2015). What makes a good structural model of personality? Evaluating the Big Five and alternatives. *Handbook of Personality and Social Psychology*, 3, 283–305. DOI: <https://doi.org/10.1037/14343-013>
- Shiner, R. L., Masten, A. S., & Roberts, J. M.** (2004). Childhood personality foreshadows adult personality and life outcomes two decades later. *Journal of Personality*, 71(6), 1145–1170. DOI: <https://doi.org/10.1111/1467-6494.7106010>
- Simms, L. J.** (2008). Classical and Modern Methods of Psychological Scale Construction. *Social and Personality Psychology Compass*, 2(1), 414–433. DOI: <https://doi.org/10.1111/j.1751-9004.2007.00044.x>
- Tangney, J. P., Baumeister, R. F., & Boone, A. L.** (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, 72(2), 271–324. DOI: <https://doi.org/10.1111/j.0022-3506.2004.00263.x>
- Vautier, S., & Pohl, S.** (2009). Do balanced scales assess bipolar constructs? The case of the STAI scales. *Psychological assessment*, i(2), 187. DOI: <https://doi.org/10.1037/a0015312>
- Webb, T. L., & Sheeran, P.** (2003). Can implementation intentions help to overcome ego-depletion? *Journal of Experimental Social Psychology*, 39(3), 279–286. DOI: [https://doi.org/10.1016/S0022-1031\(02\)00527-9](https://doi.org/10.1016/S0022-1031(02)00527-9)
- Westfall, J., & Yarkoni, T.** (2016). Statistically controlling for confounding constructs is harder than you think. *PloS one*, 11(3), e0152719. DOI: <https://doi.org/10.1371/journal.pone.0152719>
- Whiteside, S. P., & Lynam, D. R.** (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and individual differences*, 30(4), 669–689. DOI: [https://doi.org/10.1016/S0191-8869\(00\)00064-7](https://doi.org/10.1016/S0191-8869(00)00064-7)
- Zimbardo, P. G., & Boyd, J. N.** (1999). Putting time in perspective: A valid, reliable individual-differences metric. *Journal of Personality and Social Psychology*, 77(6), 1271–1288. DOI: <https://doi.org/10.1037/0022-3514.77.6.1271>

Peer review comments

The author(s) of this paper chose the Open Review option, and the peer review comments are available at: <http://doi.org/10.1525/collabra.136.pr>

How to cite this article: Ludwig, R. M., Srivastava, S., & Berkman, E. T. (2018). Planfulness: A Process-Focused Construct of Individual Differences in Goal Achievement. *Collabra: Psychology*, 4(1): 28. DOI: <https://doi.org/10.1525/collabra.136>

Senior Editor: Brent Donnellan

Submitted: 06 February 2018

Accepted: 03 July 2018

Published: 24 July 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.