FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

A COMPARISON AND ANALYSIS OF THE RELEVANT FEATURES IN REGRESSION MODELING WITH LATENT VARIABLES USING MULTI-ITEM MEASURES FOR DEVELOPMENTAL SCIENCE

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Christopher Clifford

2023

To: Dean Michael R. Heithaus
College of Arts, Sciences and Education

This dissertation, written by Christopher Clifford, and entitled A Comparison and Analysis of the Relevant Features in Regression Modeling with Latent Variables using Multi-Item Measures for Developmental Science, having been approved in respect to style and intellectual content, is referred to you for your judgement.

We have read this dissertation and recommend that it be approved.

Eliza Nelson

Nicole Fava

Stefany Coxe, Co-Major Professor

Timothy Hayes, Co-Major Professor

Leslie D. Frazier, Co-Major Professor

Date of Defense: June 16, 2023

The dissertation of Christopher Clifford is approved.

Dean Michael R. Heithaus
College of Arts, Sciences and Education

Andrés G. Gil
Vice President for Research and Economic Development
And Dean of the University Graduate School

Florida International University, 2023

ACKNOWLEDGEMENTS

I would like to thank my advisor and Major Professor, Dr. Leslie D. Frazier. Her support, patience and generosity have helped make the completion of this dissertation possible and I would not be where I am today without her. I can't thank you enough for helping to lift me up and make this degree a reality.

I would also like to thank my advisors and co-Major Professors, Dr. Stefany Coxe and Dr. Timothy Hayes. Both of them have been exceedingly generous with their time and helped me work through the technical and theoretical statistical portions of this dissertation. This topic and my interest in statistical methodologies and analytics would not have been achievable without them. Thank you both.

I would also like to thank fellow students and friends both within and outside my program for their assistance and company throughout my degree. Colleagues such as Zahra Nazarimehrvarani, Kureva Matuku, Ethan Nelson, Ibrahim Yakin, Nathan Sollenberger, Sarah Roberts and many more have all supported me in ways too numerous to count and have made this a joyful and memorable experience.

Finally, I would like to thank my family for their unwavering support and love. To my mother and father, thank you for all the love and faith you give me every day. I love you both with all my heart.

ABSTRACT OF THE DISSERTATION

A COMPARISON AND ANALYSIS OF THE RELEVANT FEATURES IN REGRESSION MODELING WITH LATENT VARIABLES USING MULTI-ITEM MEASURES FOR DEVELOPMENTAL SCIENCE

by

Christopher Clifford

Florida International University, 2023

Miami, Florida

Professor Leslie D. Frazier, Major Professor

Professor Stefany Coxe, Co-Major Professor

Professor Timothy Hayes, Co-Major Professor

In developmental science, a common and essential goal of research is to understand the relations that exist between constructs of interest. These constructs are not always directly observable and, in some cases, refer to abstract or theoretical factors. In those instances, the constructs end up being inferred from the results of measures or self-reported assessments taken across one or more items. The goal of this dissertation was to demonstrate the use of regression modeling methods for analyzing the relation between latent variables composed of multi-item measures.

This dissertation examined regression methods for multi-item measures in three separate studies. The first was an empirical study using a latent variable modeling framework to examine the roles of playfulness, stress, and coping. The second study used the data from the first study to compare four different approaches for estimating regression relationships with multi-item measures. The third study used a statistical

simulation to demonstrate how data characteristics (i.e., sample size, effect size, and reliability) differentially impact the accuracy of different methods. Together, this dissertation examined the practical and theoretical considerations in regression analysis using latent variables. The findings from this dissertation have implications for developmental scientists and provide empirically grounded guidance for the selection of regression estimation methods to use for assessing latent variable relations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION TO COLLECTED PAPERS

This dissertation examines the ways regression modeling with latent variables for multi-item measures may be applied within developmental science research. This dissertation will present three papers, each showcasing different applications and considerations for latent variable modeling. Paper I presented in Section 2 shows an empirical study that examines the relationship between playfulness, stress and coping through several measures in a latent variable model. Paper II in Section 3 examines the data analyzed in paper I in a pedagogical demonstration to compare the regression modeling technique used against others, showing the performance of four different regression modeling methods with real data. Finally, paper III in Section 4 uses a statistical simulation of data to demonstrate how specific characteristics of data (i.e., sample size, effect size, and reliability) impact the accuracy of the regression modeling techniques analyzed in paper II.

Topic and Aims

The overarching goal of this dissertation is to examine and compare quantitative methods for using latent variables from multi-item measures in regression modeling in developmental science. To this end the papers within this dissertation will provide specific practical and comparative considerations that address the constraints sometimes encountered with developmental data. Each of the three proposed papers provides information to allow researchers to evaluate the utility, value, and efficacy of different modeling techniques. Thus, this dissertation shows the practical use of these techniques in empirical research and contributes a much-needed methodological guide for developmental scientists on best practices for utilizing regression modeling techniques with latent variables for analyzing data.

Contributions from Collected Papers

The main contribution of the papers in this dissertation are: (1) to demonstrate a practical example of latent variable analysis in empirical research using personality research and data collected during the COVID19 pandemic to look at the relation between playfulness, stress and coping; (2) to present a pedagogical demonstration and comparison of how different regression techniques perform when working with latent variables composed from multi-item scales and provide considerations about their performance for developmental scientists and; (3) provide an in depth analysis of regression estimation techniques via a simulation study, using controlled parameters to specifically quantify the performance of these regression methods.

CHAPER II

Relationships among Adult Playfulness, Stress, and Coping during the COVID-19 Pandemic

This paper was published in Current Psychology, 2022

Christopher Clifford[1], Ellis Paulk[1], Qiyang Lin[2], Jeanne Cadwallader[3], Kathy Lubbers[3], and Leslie D. Frazier[1]

[1] Department of Psychology, Florida International University, Miami FL, USA

[2] Department of Counseling, Education Psychology and Special Education, Michigan State University, East Lansing, MI, USA

[3] PracticePlayfulness, Miami, FL, USA

Author Note

We have no known conflicts of interest to disclose.

This research was not funded.

The data presented in this manuscript are available upon reasonable request.

Correspondence concerning this article should be addressed to Christopher Clifford, M.S., Department of Psychology, Florida International University, 11200 S. W. 8th Street, Miami, FL 33199. Email: cclif003@fiu.edu

Abstract

The COVID-19 pandemic created high levels of stress that negatively affect mental health and well-being. The stress and coping process is influenced by individual difference factors, such as personality, that impact perceptual processes and emotional reactions. Adult playfulness is a personality characteristic that may lead to better mental and physical health outcomes. We test a theoretical model to determine whether the two factors of perceived stress, perceived self-efficacy (PSE) and perceived helplessness (PH), mediate the relationship among playfulness and coping in adults (N = 694). Scores on the Perceived Stress Scale were high indicating high levels of pandemic-related stress. The SEM model demonstrated direct effects of playfulness on PSE, PH, adaptive, maladaptive, and supportive coping. Both dimensions of perceived stress were partial mediators in the relationship among playfulness and coping outcomes. Findings illustrate the pathways by which adult playfulness can amplify or attenuate the impact of stress perceptions on coping strategies. The importance of building psychological resources such as playfulness to boost adaptive outcomes in stressful situations such as the COVID-19 pandemic is discussed.

keywords: adulthood, play, playfulness, stress, coping

Relationships among Adult Playfulness, Stress, and Coping during the COVID-19 Pandemic

Since reaching pandemic proportions in March 2020 (World Health Organization, 2020), the SARS-CoV-2 or COVID-19 pandemic has resulted in over 5.33 million deaths due to the disease and a total of 2672 million confirmed cases as of December 16, 2021 (World Health Organization Coronavirus (COVID-19) Dashboard, 2021). The overall impact of this global health crisis is unimaginable and is not confined to morbidity and mortality. At the same time, high levels of chronic stress related to the pandemic has resulted in declining mental and physical health and well-being (Brooks et al., 2020; Castelli et al., 2020; Di Giuseppe et al., 2020, Minahan et al., 2021; Ornell et al., 2020; Osimo et al., 2021; Rossi et al., 2020; Salari et al., 2020). These negative mental and physical outcomes are far-reaching, for example recent reports predict that the confluence of pandemic-related consequences will ultimately elevate suicide attempts and rates in future (Zalsman et al., 2020; see also De Berardis et al., 2018; Orsolini et al., 2020). One recent meta-analysis conducted during the first peak of the pandemic showed that worldwide levels of stress were as high as 29.6% (based on 9,074 participants across five studies; Salari et al., 2020). Another early study of the general public living in China, over 8% reported moderate to severe stress (Wang et al., 2020). Whereas in Italy, 27.2% of the population self-reported high or extremely high levels of stress (Mazza et al., 2020). In the nearly two years since the pandemic the prolonged stress associated with this global health crisis continues to impact mortality, mental and physical health outcomes, and quality of life worldwide (Liu et al., 2021).

The experience of stress and how we cope with it is influenced by individual difference factors such as personality. Personality impacts individuals' perceptual processes and emotional reactions to the stress of the pandemic (Osimo et al., 2021). Personality also influences how individuals cope with stress. Within the context of the pandemic, those with higher levels of resilience coped more successfully with pandemic-related stressors (Morales-Vives et al., 2020). One personality characteristic that may impact the stress and coping process is playfulness (Guitard et al., 2005; Proyer, 2013). Like resilience and optimism and other positive psychological attributes, adult playfulness could be a resource that leads to better mental and physical health outcomes (Farley et al., 2021; Proyer et al., 2018), especially important in the extraordinary context of coping with the COVID-19 pandemic. The present study examines the potentially beneficial role of adult playfulness as it impacts perceived stress and coping during the COVID-19 pandemic.

Perceived Stress

The subjective experience of stress is based upon a cognitive appraisal process through which (a) the event is assessed as threatening or demanding (e.g., primary appraisal), and (b) one's resources for managing are seen as lacking (e.g., secondary appraisal; Lazarus & Folkman, 1984). Since the cognitive appraisal process is driven by an individual's unique perceptions, interpretations, and experiences with a particular stressor, there are large individual differences in the perception of stress.

The most widely used measure of perceived stress is the Perceived Stress Scale (PSS) developed by Cohen and colleagues (Cohen et al., 1983). The PSS assesses the degree to which individuals perceive their lives as stressful over the last month. Several scholars (Golden-Kreutz et al., 2004; Hewitt et al., 1992; Khalili et al., 2017; Örücü & Demir, 2009; Roberti et al., 2006) have suggested a two-factor structure for the PSS: (a) perceived self-efficacy (the ability to manage and control the stressor); and (b) perceived helplessness (the perception that stress is outside one's control). Studies show that perceived self-efficacy relates to lower stress perceptions and better mental and physical health outcomes, whereas perceived helplessness shows the opposite pattern (Durak et al., 2010; Kaya et al., 2019). Understanding the individual difference factors that may impact the perceptual and subjective experience of stress is important and can shed light on how people cope.

Coping

Coping is conceptualized as the "cognitive and behavioral efforts to manage specific external and/or internal demands that are appraised as taxing or exceeding the resources of the person" (Lazarus & Folkman, 1984, p. 34). It represents the efforts individuals make to manage or adjust to the demands of the stressful situation as well as to regulate the emotional response to it (Barreto & Frazier, 2012).

Just as there are individual differences in the perception of stress, there are also individual differences in how individuals cope in different contexts (Carver et al, 1989). Coping aimed at directly altering the situation represents a form of problem-focused, planful problem-solving or active coping (Lazarus, 1980), and is useful when the individual has a high degree of perceived control and the self-efficacy to change the situation. Emotion-focused coping involves regulation of emotions through strategies such as distancing, venting, searching for meaning. It's aim is to change the emotional impact of the stressor without actually changing the situation (Frazier, 2002). Therefore, emotion-focused coping is useful when the individual has a relatively low amount of perceived control over the situation (Lazarus, 1992). Both problem-focused and emotion-focused coping are often used simultaneously (Lazarus, 1980).

Within health contexts, research on coping attempts to capture the more fine-grained strategies used to cope with stress (Carver et al., 1989). Specifically, disengagement, self-distraction, active coping, using emotional support from intimate partners, emotional support from others, relying on religion, humor, and substance use were found to represent meaningful coping strategies for managing illness (Fillion et al., 2002). Some researchers (e.g., Brown et al., 2005; Folkman & Lazarus, 1988; Jex et al., 2002) have argued that it is more useful to distinguish coping efforts based on whether they are harmful (e.g., ineffective, not reducing or increasing distress) or helpful (e.g., effective, reducing or ameliorating distress). Coping strategies that promote better adjustment are conceptualized as adaptive coping to refer to efforts to deal directly with the stressor by finding and implementing solutions (Parasuraman et al., 1987). Whereas maladaptive coping refers to coping through avoidance, self-criticism, and negative emotions (Kirby et al., 2011). Typically, adaptive coping leads to positive outcomes and maladaptive coping leads to negative outcomes (Parasuraman et al., 1987). This distinction could be particularly useful in health contexts.

Coping with the stress of the COVID-19 pandemic has been universally challenging. Research shows that in addition to extremely high levels of stress and experiences of PTSD, individuals who reported they avoided thinking about the stress and those who were unsure or unable to cope with it had greater levels of anxiety and depression (Kar et al., 2021). In one study of healthcare professionals during the pandemic, positive attitude – a functional coping strategy -- as well as turning to religion were adaptive and lead to better outcomes (Kar et al., 2021). Seeking social support and avoidance in a were coping strategies that led to negative outcomes such as higher distress (Babore et al., 2021). Avoidance and denial have traditionally been found to be dysfunctional ways to manage stress, especially in contexts such as a pandemic (Babore et al., 2021; Phua et al., 2005; Teasdale et al., 2012). These findings highlight the role of individual differences in the appraisal of stress and the situational conditions that impact coping efficacy (Biggs et al., 2017; Folkman & Moskowitz, 2004; Lazarus & Folkman, 1984).

Personality and the stress process

Personality, the lens through which we interpret our world, drives the individual differences seen in the stress and coping processes (Barreto & Frazier, 2012). Individuals assign meaning to a given situation through a dynamic, interactive, constructive process in which personality plays a key role in shaping emotional reactions, coping responses, and even one's health (Hooker et al., 1988; Lazarus, 1991). Personality may have direct effects on stress, coping, and health and can predispose people to interpret events in benign or threatening ways. Personality also impacts the stress process indirectly through its relationship with resources brought to bear in the coping process. Frazier and colleagues (Frazier, 2000; Hooker et al., 1994; Hooker et al., 1998), showed that personality is a powerful predictor of coping patterns and mental and physical health outcomes.

Positive personality characteristics may attenuate the experience of stress and lead to better coping outcomes (Carver & Scheier, 1991). For example, people who have positive expectations and optimism, may perceive less stress and cope better with stressful life events (Carver & Scheier, 1991). Personality predispositions, such as optimism, may represent the ability to remain flexible and may be key to resilience and thriving in extremely stressful situations (Carver, 1998). Due to the overlap between the traits that predict resilience and those that define playfulness, such evidence potentiates playfulness as an effective means to promote resilience and even thriving within the COVID-19 pandemic. Personality may also impact coping through social support (Cohen & Wills, 1985). The stress-buffering effects of interpersonal relationships may depend on an individual's personality. A playful person may elicit more social support, which may then, be an effective moderator of the stress process.

Playfulness as a dimension of personality

The concept of playfulness has been studied extensively in childhood. However, until recently, it has been largely overlooked in adulthood (Proyer, 2017; Smith, 1996). In adulthood playfulness is conceptualized as an individual difference variable that predisposes the way a person perceives or experiences situations (Proyer, 2012, 2017; Proyer et al., 2019). As Barnett states, playfulness is "the predisposition to frame (or reframe) a situation in such a way as to provide oneself (and possibly others) with amusement, humor, and/or entertainment" (Barnett, 2007, p. 955).

Research has advanced understanding of the importance of adult playfulness (Bowman, 1987; Glynn, 1992; Guitard et al., 2005; Martocchio & Webster, 1992), and demonstrated its benefit for better health, higher productivity, tension release, group cohesion, and improved workplace performance (Shen et al., 2014). Further research has linked adult playfulness to higher levels of creativity (Tegano, 1990), improved morale and motivation (Lyons, 1987), and increased adaptivity (Guitard et al., 2005).

Playfulness and Coping

Playfulness has been argued to be crucial to the process of coping (Chang et al., 2013; Hess & Bundy, 2003; Magnuson & Barnett, 2013; Saunders et al., 1999; Staempfli, 2007; Yarnal, 2011). As Magnuson and Barnett (2013) showed, playfulness mediates the stress-coping process through its influence on cognitive appraisal. As a coping strategy, playfulness may mediate the interpretation and experience of stress (Hess & Bundy, 2003; Staempfli, 2007). Playfulness may also guide reframing stressful situations in a way that facilitates flexibility, reduces perceived stress, and improves resilience (Barnett, 2007). In addition, highly playful individuals use adaptive or engagement coping strategies more frequently than their less playful counterparts, although both groups used the same types of coping strategies overall (Barnett, 2013). Higher levels of playfulness result in greater flexibility when dealing with difficult life events (Bundy 1993), supporting the notion that playfulness acts as a facilitator of stress and coping in adults (Yarnal, 2011). These findings suggest that playfulness may serve as an important coping resource that affords individuals with the capacity to cope more effectively with highly stressful situations than may otherwise cause psychological distress.

Coping with the stress of the pandemic

Consider the stress of the COVID-19 pandemic. The stressors of the pandemic, while universal, dynamic, and unfolding over time, still impacts individuals differently leading to interindividual differences in how people cope with the stress of COVID-19. It bears mention that anxiety and distress are normal reactions to the unpredictable, ambiguous, and personally threatening nature of the COVID-19 pandemic. Nevertheless, ineffective coping may be maladaptive and may worsen negative mental and physical health outcomes. Whereas effective coping strategies are critical to mitigating the negative mental and physical health outcomes that have arisen due to the pandemic. Research has shown that playfulness acts as an effective facilitator in the stress process by reducing perceived stress, encouraging adaptive coping, decreasing negative emotions, and increasing positive emotions and life satisfaction. Despite their elevated risk of illness and death due to COVID-19, adults, and especially older adults, reported lower levels of negative affect and more agentic coping than younger adults (Young et al., 2021). Considering the benefits conveyed through playfulness, its use during the COVID-19 pandemic could prove to be an important, and even vital, resource to mitigate negative mental health outcomes.

The Present Study

The purpose of the present study was to examine the attributes of playfulness in adults and how they may interact with stress and coping within the context of the COVID-19 pandemic. Our goals were to determine what aspects of playfulness were most characteristic of our sample and to determine how playfulness impacts the perception of stress and the choice of coping strategy. Given the influence of playfulness on the cognitive appraisal process, we were specifically interested in the two-factor model of perceived stress (Durak et al., 2010; Kaya et al., 2019). The overarching goal is to identify a pattern of playfulness in adulthood that may serve as a valuable resource for effective stress management, especially in times of extreme stress. To achieve this goal, we test a theoretically grounded model (see Figure 1) in which the two-factor model of perceived stress mediates the influence of playfulness on coping outcomes.

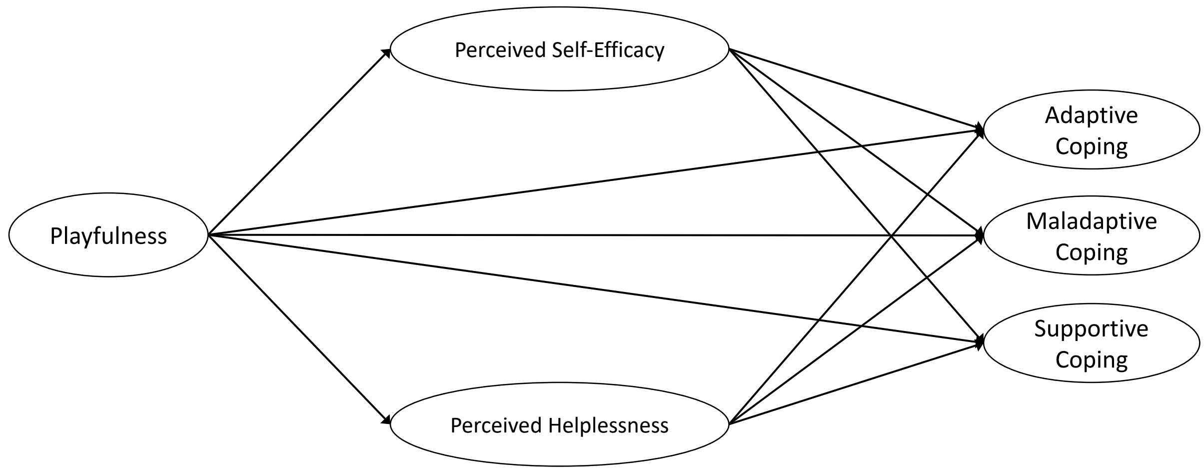


Figure 1. Mediation model displaying direct and indirect pathways of playfulness, perceived self-efficacy, perceived helplessness, and coping dimensions.

Consistent with earlier research (Chang et al., 2013; Magnuson & Barnett, 2013; Staempfi, 2007), we hypothesized that playful adults would experience less stress and cope more effectively with the stress of COVID-19. We expected playfulness to coact with perceived stress to either amplify or attenuate coping.

Methods

Participants

The participants in this study were 837 English-speaking adults (M_{age} = 23.52, *SD* = 6.28). The sample was predominantly female (88.05%). The sample self-reported race/ethnicity as African American/Black (13.52%), Asian American/Pacific Islander (1.33%), Asian (1.21%), Biracial/multiracial (1.70%), Hispanic/Latino (67.36%), Middle Eastern/Arab (0.73%), Native American/native Alaskan (0.12%), White/European American (8.40%), and other (5.6%). Whereas the most recent U.S. census shows that

18% of the population reports Hispanic as their racial/ethnic background, this study was conducted at the largest U.S. Hispanic-serving university in the nation. The largely Hispanic student body is representative of the population in which this research took place. Socioeconomic status was measured on a Likert-type scale that ranged from an annual household income of less than $4,000 to a high of over $150,000 or more. The most frequently self-reported annual income categories were $20,000-$34,000 and $35,000-$49,000. This is lower than the U. S. Department of Household and Labor statistics national median family income for 2021 that was $79,000 (Richardson, 2021).

Measures

Playfulness

Playfulness was measured using the Short Measure of Adult Playfulness (SMAP; Proyer, 2012), a five-item questionnaire using a four-point Likert-type scale with answer choices ranging from 1 (strongly disagree) to 4 (strongly agree). Questions ask participants to self-assess how playful they are both as identified by themselves ("I am a playful person"), and by others ("Good friends would describe me as a playful person"), as well as the frequency and level of engagement they normally experience involving play. High scores indicate greater playfulness. In the present sample, scores ranged from a low of five to a high of 20 (M = 15.76; SD = 2.92). The Cronbach's alpha (α = 0.84) indicated strong internal consistency and reliability.

Stress

The Perceived Stress Scale (PSS; Cohen et al., 1983) is a10-item Likert-type scale that assesses the perception of stress during the last month, and how often respondents thought or felt a certain way. Answer choices range from 0 (never) to 4 (very often). Questions such as "In the last month, how often have you felt nervous and 'stressed'?" or "In the last month, how often have you felt that you were on top of things?" After item reversals a total score is gleaned. High scores indicate higher perceptions of stress. Consistent with prior research, we conducted a confirmatory factor analysis (CFA; Lance & Vandenberg, 2002) and landed on two distinct factors: perceived helplessness and perceived self-efficacy. In the present sample, scores ranged from a low one of to a high of 40 (*M* = 20.08; *SD* = 7.42). The Cronbach's alpha ($\alpha = 0.88$) for the total score in the current sample indicates strong internal consistency and reliability.

Coping

Coping strategies were assessed using 28-item brief COPE scale (Carver, 1997). Coping is assessed across 14 dimensions: active coping, planning, positive reframing, acceptance, humor, religion, using emotional support, using instrumental support, self-distraction, denial, venting, substance use, behavioral disengagement, and self-blame. Responses, on a Likert-type scale, ranged from 1 ("I haven't been doing this at all") to 4 ("I've been doing this a lot"). We conducted confirmatory factor analysis (CFA; Lance & Vandenberg, 2002) to assess the factor structure of the brief COPE scale in the present sample. We found coping strategies cluster onto three distinct variables: adaptive coping focused (e.g., strategies that emphasized active engagement and planning); maladaptive coping (e.g., strategies focused on disengaging from a problem or distracting oneself including denial, disengagement, substance use); and supportive coping (e.g., focused on receiving support from other including emotional support, religion). In the present sample, the Cronbach's alpha (α = .86) for the total scale demonstrated strong internal consistency and high reliability.

Procedure

Recruitment took place through online SONA systems software of a large public urban university in the southern United States. SONA systems is an online system for announcing, scheduling, and awarding research credit for university students who participate in research. Additionally, we recruited through social media platforms where flyers were posted. All recruitment channels provided inclusion/exclusion criteria required for participation. Inclusion criteria were: (a) age 18 years old or older; and (b) competence reading and writing in English. Participants who did not meet these criteria were excluded from participation in the prescreening questions in SONA systems and within the survey itself. After providing informed consent, participants were administered the scales as part of a larger study on the mental and physical health outcomes during the COVID-19 pandemic. The survey was administered fully online through Qualtrics survey software. Participants recruited through the University were awarded research credit for participation.

Results

A total score of 13 on the PSS represents a normal level of stress and scores of 20 or higher represent high levels of stress which require therapeutic intervention (Cohen & Williamson, 1988). Mean scores on the PSS across many studies with many samples range from 12 to 14.7 (Lee, 2012). The average score for the present sample was 20.08

(*SD* = 7.42). Given that respondents indicated their perceptions of stress over the last month, during the pandemic, these high levels of stress reflect the stress of the pandemic.

The data set did not have a significant amount of missing data. Outliers were minimal due to the Likert-type data collected. The analyses were performed on ordinal categorical data, thus missing data techniques such as FIML were not used, and missing cases were dropped resulting in a total of 694 participants. We opted to use `lavaan` for the Structural Equation Modeling on our ordered categorical data.

Based on prior research (Ng, 2013) and CFA results confirming overall model fit, the PSS was treated as two latent variables: perceived helplessness and perceived self-efficacy. Next, we ran a parallel mediation model (See Figure 1) with two mediators computing both the direct and indirect pathways using the `lavaan` package (Rosseel, 2012) in `R` R Core Team, 2022). The final fit indices showed the model had an acceptable fit for the variables ($\chi^2(764) sig: < 0.001$, TLI – 0.91, CFI – 0.91, RMSEA – 0.08, SRMR – 0.10). See Figure 1.

The Model

The SEM model results showed playfulness was a significant predictor of perceived self-efficacy (β = 0.315, p < 0.001), and perceived helplessness (β = -0.141, p = 0.001). Higher levels of playfulness were related to higher levels of perceived self-efficacy while higher levels of playfulness were related to lower levels of perceived helplessness. The specifics of these and further comparisons in the main model can be found in SEM regression estimates on Table 1.

Table 1: SEM Regression Estimates Table

Parameter Estimates	Estimate	Standardized Estimate (β)	*SE*	*95% CI*		*p*
				LB	*UB*	
Direct Effects of Playfulness						
Adaptive Coping	0.177	0.277	0.025	0.129	0.226	<0.001
Maladaptive Coping	0.046	0.116	0.014	0.02	0.073	0.001
Supportive Coping	0.132	0.175	0.03	0.073	0.191	<0.001
Perceived Self-Efficacy	0.242	0.315	0.032	0.179	0.306	<0.001
Perceived Helplessness	-0.116	-0.141	0.033	-0.182	-0.050	0.001
Mediator Effects						
Perceived Self-Efficacy predicting						
Adaptive Coping	0.485	0.585	0.055	0.378	0.593	<0.001
Maladaptive Coping	-0.066	-0.127	0.027	-0.119	-0.012	0.017
Supportive Coping	0.370	0.377	0.054	0.263	0.476	<0.001
Perceived Helplessness predicting						
Adaptive Coping	0.284	0.366	0.045	0.195	0.372	<0.001
Maladaptive Coping	0.280	0.581	0.038	0.206	0.353	<0.001
Supportive Coping	0.292	0.319	0.048	0.199	0.385	<0.001
Indirect Effects						
Playfulness through Perceived Self-Efficacy						
Adaptive Coping	0.118	0.184	0.02	0.078	0.157	<0.001
Maladaptive Coping	-0.016	-0.04	0.007	-0.029	-0.002	0.021
Supportive Coping	0.090	0.119	0.017	0.056	0.123	<0.001
Playfulness through Perceived Helplessness						
Adaptive Coping	-0.033	-0.051	0.011	-0.055	-0.011	0.003

Maladaptive Coping	-0.032	-0.082	0.010	-0.052	-0.013	0.001
Supportive Coping	-0.034	-0.045	0.012	-0.057	-0.011	0.003
Total Effects of Playfulness						
Adaptive Coping	0.262	0.410	0.024	0.216	0.308	<0.001
Maladaptive Coping	-0.002	-0.005	0.016	-0.033	0.028	0.894
Supportive Coping	0.187	0.248	0.028	0.133	0.242	<0.001

Perceived self-efficacy was a significant predictor of adaptive coping (β = 0.585, $p < 0.001$), maladaptive coping (β = -0.127, $p = 0.017$), and supportive coping (β = 0.319, $p < 0.001$). Increases in self-efficacy showed increases for adaptive and supportive coping. However, increasing self-efficacy decreased maladaptive coping. Perceived helplessness similarly had significant relations between all three styles of coping: adaptive (β = 0.336, $p = 0.001$), maladaptive (β = 0.581, $p < 0.001$), and supportive (β = 0.377, $p < 0.001$). Increases in perceived helplessness showed across the board increases in each of the three styles of coping.

Examining the total indirect pathways for mediation, the pathway of playfulness through perceived self-efficacy was significant for adaptive (β = 0.184, $p < 0.001$), maladaptive (β = -0.040, $p = 0.021$), and supportive coping (β = 0.119, $p < 0.001$). Increases in playfulness predicted increases in adaptive and supportive coping through perceived self-efficacy. In contrast, increases in playfulness predicted a decrease in maladaptive coping when it was through self-efficacy. For the indirect pathway examining how playfulness predicted coping through perceived helplessness the adaptive (β = -0.051, $p = 0.003$), maladaptive (β = -0.082, $p = 0.001$), and supportive coping (β = -0.045, $p = 0.003$) indirect paths were once again all significant. For the indirect pathway using the perceived helplessness mediator, all three coping outcomes showed that an increase in playfulness would predict a decrease in coping along perceived helplessness.

Finally, the direct effects of playfulness on coping showed significant results for the three coping styles: adaptive ($\beta = 0.277$, $p < 0.001$), maladaptive ($\beta = 0.116$, $p = 0.001$) and supportive ($\beta = 0.175$, $p < 0.001$). Higher levels of playfulness were directly related to higher levels of adaptive coping, higher levels of maladaptive coping and higher levels of supportive coping. The total effects pathways were significant in the case of adaptive ($\beta = 0.410$, $p < 0.001$) and supportive coping ($\beta = 0.248$, $p < 0.001$) but not significant in the case of maladaptive coping ($\beta = -0.005$, $p = 0.894$).

Due to the results showing maladaptive coping not having a significant total effect, an additional examination of the mediators was run by comparing the results of single-mediator models. In these models, the perceived self-efficacy and perceived helplessness were each given their own model where they were used as the sole mediator in the play-stress-coping mediation model. These results can be seen in Appendix: Supplemental Tables 1A and 1B.

Of note, these single mediation models show the total effects of maladaptive coping being insignificant in both cases. The only notable difference in the single mediator models was in the perceived helplessness single-mediator model, where the indirect effects of playfulness through both adaptive and supportive coping were insignificant. This may indicate that perceived self-efficacy has a confounding effect on perceived helplessness as a mediator (MacKinnon et al., 2000).

In summary, playfulness predicted adaptative and supportive coping completely, both through the direct, indirect, and total effects. Increases in playfulness positively predicted adaptive and supportive coping directly as well as indirectly through perceived self-efficacy. Higher levels of playfulness were related to lower levels of adaptive and supportive coping through perceived helplessness. Finally, playfulness predicted maladaptive coping through direct effects and through the indirect effects of perceived self-efficacy and perceived helplessness. Higher playfulness linked to higher maladaptive coping directly but lower maladaptive coping through perceived self-efficacy and perceived helplessness. Further, playfulness was a significant predictor for both perceived self-efficacy and perceived helplessness; a positive predictor in the case of self-efficacy and a negative predictor in the case of perceived helplessness.

Discussion

This study demonstrates how playfulness predicts stress and coping in adults during the COVID-19 pandemic. Our findings show that, in the context of the pandemic, our participants were highly stressed. Playful individuals -- those who consider themselves playful and are identified as such by others -- perceive less stress and use more adaptive coping strategies to lessen the distress. Specifically, playfulness, a beneficial psychological resource, related to perceived stress in predictable ways. Higher levels of playfulness were positively related to higher levels of perceived self-efficacy for stress. Lower levels of playfulness related to higher levels of perceived helplessness. Playfulness, defined as the predisposition to perceive and interpret situations in a way that provides oneself and others with amusement, humor, and/or entertainment (Proyer & Ruch, 2011), is linked with how stressful situations are perceived as either something within or outside one's control. Our findings support prior research that shows that playfulness reduces stress (Magnuson & Barnett, 2013). Our findings also show that more playful people are more likely to use adaptive and social support coping – both effective means at reducing distress. We also found the opposite pattern in which lower levels of playfulness related to higher levels of maladaptive coping. These findings are consistent with prior research that shows that playful individuals tend to use beneficial, adaptive, and stressor-focused coping strategies while less playful individuals tend to rely on negative, avoidant, escape-oriented, maladaptive strategies (Magnuson & Barnett, 2013). Our findings suggest that especially in the high stress context of the COVID-19 pandemic, playfulness can be a personal resource that provides a strong adaptive advantage for stress perception and coping efficacy.

In this study we have extended prior research to examine the role of playfulness in stress and coping at a more detailed level by examining the two factors of perceived stress, and how those perceptions may mediate the role of playfulness on coping. We found direct effects of playfulness on both dimensions of perceived stress, and all dimensions of coping. As well, both dimensions of perceived stress showed direct effects on each of the coping outcomes. Perceived self-efficacy was related to higher levels of adaptive and support coping and lower levels of maladaptive coping. Whereas, higher levels of perceived helplessness was related to lower adaptive and support coping and higher maladaptive coping. We found that perceived self-efficacy partially mediates the impact of playfulness on all three forms of coping. That is, the coactive influence of higher playfulness and higher self-efficacy is positively related to more adaptive and support coping, and reduces the likelihood of maladaptive coping. We also found that perceived helplessness partially mediates the role of playfulness such that the positive impacts of playfulness on adaptive coping remain, and the negative impact of perceived helplessness on adaptive coping are offset. Similarly, the interactive effects of playfulness and perceived helplessness relate to lower levels of maladaptive coping.

Some may see our reliance on cross-sectional data to test our mediational hypotheses as a potential limitation. One could certainly argue that by modeling the manner in which relationships among variables unfold over time, longitudinal mediation designs (Maxwell et al., 2011; Maxwell & Cole, 2007), may be better equipped than cross-sectional designs to speak to causality and causal ordering. However, it bears emphasizing that the temporal precedence captured by these models is a necessary, but not a sufficient, condition for establishing causality.

In the context of our present research questions, we set out to conduct a broad assessment of the influence of playfulness on perceived stress and coping. We assessed these individual differences factors using participants' responses to standard, well-validated, widely used scales.

Therefore, rather than viewing our design as a convenient, possibly distorted (cf. Maxwell et al., 2011; Maxwell & Cole, 2007) cross-sectional substitute for an ostensibly superior longitudinal one, we, in fact, believe this to be a valid, defensible, appropriate design for answering our specific research questions. Other possible limitations are the self-report nature of the data and the reliance on a university participant pool. Nevertheless, the large sample size and extensive analyses validated the measures and provided illuminating pathways that may provide blueprints for future larger population-based, longitudinal, or intervention-type studies.

Niall Bolger (1990) said that “coping is personality in action under stress” (p. 525). In the present study we have shown that playfulness, a disposition that influence how individuals perceive, interpret, and engage with their lives, has both direct and indirect effects on coping – and therefore may be an important malleable personal resource that can be cultivated to help people perceive less stress, perceive more control over the stress they have, and to use more adaptive coping strategies to reduce the distress. Although we have measured playfulness in terms of how people currently see themselves, and we have shown how those self-perceptions are beneficial in the stress-coping process, playfulness can be learned at any age (Andreopoulou et al., 2019; McMillan, 2017; Rice, 2009; Tanis, 2012). There are many ways to be playful from engaging with ideas, being spontaneous, lightening the mood with humor or silliness, playing games, to framing a situation in a positive light; playfulness provides an adaptive benefit that leads to better outcomes for mental and physical well-being (Proyer, 2017). Our findings demonstrate the importance of perceiving self-efficacy for managing stress as a direct and indirect influence on coping. The combination of playfulness and perceive self-efficacy interact to positively bolster adaptive and supportive coping strategies. The take home messages from our findings are the potential useful pathways through which adults can maximize their playfulness to achieve more optimal outcomes even within the highly stressful context of the COVID-19 pandemic.

References

Aldwin, C. M., Levenson, M. R., Spiro, A., & Bossé, R. (1989). Does emotionality predict stress? Findings from the Normative Aging Study. *Journal of Personality and Social Psychology*, *56*(4), 618.

Andreopoulou, P., & Moustakas, L. (2019). Playful Learning and Skills Improvement. *Open Journal for Educational Research*, *3*(1), 25-38.

Babore, A., Trumello, C., Lombardi, L., Candelori, C., Chirumbolo, A., Cattelino, E., ... & Morelli, M. (2021). Mothers' and Children's Mental Health During the COVID-19 Pandemic Lockdown: The Mediating Role of Parenting Stress. *Child Psychiatry & Human Development*, 1-13.

Barnett, L. (2007). The Nature of Playfulness in Young Adults. *Personality and Individual Differences*, 949-958.

Barreto, M. L., & Frazier, L. D. (2012). Coping with Life Events through Possible Selves. *Journal of Applied Social Psychology*, *42*(7), 1785–1810. https://doi.org/10.1111/j.1559-1816.2012.00918.x.

Biggs A., Brough, P., & Drummond, S. (2017). Lazarus and Folkman's psychological stress and coping theory. In C.L. Copper, and J.C. Quick (Eds.), *The Handbook of Stress and Health: Guide to Research and Practice.* John Wiley & Sons, Hoboken, NJ.

Bolger, N., & Schilling, E. A. (1991). Personality and the Problems of Everyday Life: The Role of Neuroticism in Exposure and Reactivity to Daily Stressors. *Journal of personality*, *59*(3), 355-386.

Bowman, J. (1987). Making Work Play. In G. A. Fine (Eds), *Meaningful play, playful meanings*, p, 61-71, Champaign, IL: Human Kinetics.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The Psychological Impact of Quarantine and How to Reduce It: Rapid Review of the Evidence. *The Lancet*, *395*(10227), 912-920.

Brown, S. P., Westbrook, R. A., & Challagalla, G. (2005). Good Cope, Bad Cope: Adaptive and Maladaptive Coping Strategies Following a Critical Negative Work Event. *Journal of Applied Psychology*, *90*(4), 792.

Bundy, A. (1993). Assessment of Play and Leisure: Delineation of the Problem. *American Journal of Occupational Therapy*, *47*, 217-222.

Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the brief COPE. *International Journal of Behavioral Medicine*, *4*(1), 92–100. https://doi.org/10.1207/s15327558ijbm0401_6.

Carver, C. (1998). Resilience and Thriving: Issues, Models, and Linkages. *Society for the Psychological Study of Social Issues*, 54:2, 245-266.

Carver, C. S., & Scheier, M. F. (1991). Self-regulation and the Self. In *The Self: Interdisciplinary Approaches* (pp. 168-207). Springer, New York, NY.

Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology, 56(2*), 267–283. https://doi.org/10.1037/0022-3514.56.2.267.

Castelli, L., Di Tella, M., Benfante, A., & Romeo, A. (2020). The spread of COVID-19 in the Italian population: anxiety, depression, and post-traumatic stress symptoms. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, *65*(10), 731.

Chang, P. J., Qian, X., & Yarnal, C. (2013). Using playfulness to cope with psychological stress: Taking into account both positive and negative emotions. *International Journal of Play*, *2*(3), 273–296. https://doi.org/10.1080/21594937.2013.855414.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*(2), 310.

Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. *Measuring stress: A guide for health and social scientists*, *10*(2), 1-2.

De Berardis, D., Fornaro, M., Valchera, A., Cavuto, M., Perna, G., Di Nicola, M., Serafini, G., Carano, A., Pompili, M., Vellante, F., Orsolini, L., Fiengo, A., Ventriglio, A., Yong-Ku, K., Martinotti, G., Di Giannantonio, M., & Tomasetti, C. (2018). Eradicating Suicide at Its Roots: Preclinical Bases and Clinical Evidence of the Efficacy of Ketamine in the Treatment of Suicidal Behaviors. *International Journal of Molecular Sciences*, *19*(10), 2888. https://doi.org/10.3390/ijms19102888

Di Giuseppe, M., Gemignani, A., & Conversano, C. (2020). Psychological resources against the traumatic experience of COVID-19. *Clinical Neuropsychiatry: Journal of Treatment Evaluation, 17*(2), 85–87.

Durak, M., Senol-Durak, E., & Gencoz, T. (2010). Psychometric properties of the satisfaction with life scale among Turkish university students, correctional officers, and elderly adults. *Social Indicators Research*, *99*, 413-429.

Falk, G., Carter, J. A., Carter, I.A., Nyhof, E. C., & Romero, P. D. (2020) Unemployment rates during the COVID-19 Pandemic: In brief (Rep. No. R46554). Congressional Research Service, Retrieved December 14, 2020.

Farley, A., Kennedy-Behr, A., & Brown, T. (2021). An investigation into the relationship between playfulness and well-being in Australian adults: An exploratory study. *OTJR: Occupation, Participation and Health*, *41*(1), 56-64.

Fillion, L., Kovacs, A. H., Gagnon, P., & Endler, N. S. (2002). Validation of the shortened cope for use with breast cancer patients undergoing radiation therapy. *Current Psychology*, *21*(1), 17–34. https://doi.org/10.1007/bf02903157.

Flesia, L., Monaro, M., Mazza, C., Fietta, V., Colicino, E., Segatto, B., & Roma, P. (2020). Predicting perceived stress related to the covid-19 outbreak through stable psychological traits and machine learning models. *Journal of Clinical Medicine*, *9*(10), 3350. https://doi.org/10.3390/jcm9103350.

Folkman, S., & Lazarus, R. S. (1988). The relationship between coping and emotion: Implications for theory and research. *Social Science & Medicine*, *26*(3), 309-317.

Folkman, S., & Moskowitz, J. T. (2004). Coping: Pitfalls and promise. *Annual Review of Psychology*, *55*(1), 745–774. https://doi.org/10.1146/annurev.psych.55.090902.141456.

Frazier, L. D. (2002). Stability and change in patterns of coping with Parkinson's disease. *The International Journal of Aging and Human Development*, *55*(3), 207-231.

Glynn, M.A. (1992). The adult playfulness scale: An initial assessment. *Psychological Reports*, *71*(5), 83. https://doi.org/10.2466/pr0.71.5.83-103.

Golden-Kreutz, D. M., Browne, M. W., Frierson, G. M., & Andersen, B. L. (2004). Assessing stress in cancer patients: a second-order factor analysis model for the Perceived Stress Scale. *Assessment*, *11*(3), 216-223.

Guitard, P., Ferland, F., & Dutil, É. (2005). Toward a better understanding of playfulness in adults. *OTJR: Occupation, Participation and Health*, *25*(1), 9–22. https://doi.org/10.1177/153944920502500103.

Hess, L. M., & Bundy, A. C. (2003). The Association Between playfulness and coping in adolescents. *Physical & Occupational Therapy In Pediatrics*, *23*(2), 5–17. https://doi.org/10.1080/j006v23n02_02.

Hewitt, P. L., Flett, G. L., & Mosher, S. W. (1992). The Perceived Stress Scale: Factor structure and relation to depression symptoms in a psychiatric sample. *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 247-257.

Hooker, K., Frazier, L. D., & Monahan, D. J. (1994). Personality and coping among caregivers of spouses with dementia. *The Gerontologist*, *34*(3), 386-392.

Hooker, K., Monahan, D. J., Bowman, S. R., Frazier, L. D., & Shifren, K. (1998). Personality counts for a lot: Predictors of mental and physical health of spouse caregivers in two disease groups. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53*(2), P73-P85.

Johns Hopkins University (2021). *Mortality and Morbidity of COVID-19*. Retrieved from JohnHopkinsMedicine.org.

Kar, N., Kar, B., & Kar, S. (2021). Stress and coping during covid-19 pandemic: Result of an online survey. *Psychiatry Research*, *295*, https://doi.org/10.1016/j.psychres.2020.113598.

Kaya, C., Tansey, T. N., Melekoglu, M., Cakiroglu, O., & Chan, F. (2019). Psychometric evaluation of Turkish version of the Perceived Stress Scale with Turkish college students. *Journal of Mental Health*, *28*(2), 161-167.

Khalili, R., Ebadi, A., Tavallai, A., & Habibi, M. (2017). Validity and reliability of the Cohen 10-item Perceived Stress Scale in patients with chronic headache: Persian version. *Asian Journal of Psychiatry*, *26*, 136-140.

Kirby, R., Shakespeare-Finch, J., & Palk, G. (2011). Adaptive and maladaptive coping strategies predict posttrauma outcomes in ambulance personnel. *Traumatology*, *17*(4), 25-34.

Lance, C. E., & Vandenberg, R. J. (2002). Confirmatory factor analysis. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 221–254). Jossey-Bass.

Lazarus, S. F. (1980). An Analysis of Coping in a Middle-Aged Community Sample. *Journal of Health and Social Behavior*, 219-239.

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist*, *46*(4), 352-367.

Lazarus, R. S. (1992). Coping Theory and Research: Past, Present, and Future. *Psychosomatic Medicine*, 234-247.

Lazarus, R. S., & Folkman, S. (1984). *Stress, Appraisal, and Coping*. New York, NY: Springer.

Lee, E. H. (2012). Review of the psychometric evidence of the perceived stress scale. *Asian Nursing Research*, *6*(4), 121-127.

Litman, J. A. (2006). The cope inventory: Dimensionality and relationships with approach and avoidance motives and positive and negative traits. *Personality and Individual Differences*, *41*(2), 273–284. https://doi.org/10.1016/j.paid.2005.11.032.

Liu, S., Lithopoulos, A., Zhang, C. Q., Garcia-Barrera, M. A., & Rhodes, R. E. (2021). Personality and perceived stress during COVID-19 pandemic: Testing the mediating role of perceived threat and efficacy. *Personality and Individual differences*, *168*, 110351.

Lyons, M. (1987). A taxonomy of playfulness for use in occupational therapy. *Australian Occupational Therapy Journal*, 152-156.

Magnuson, C. D., & Barnett, L. A. (2013). The playful advantage: How playfulness enhances coping with stress. *Leisure Sciences*, *35*(2), 129–144. https://doi.org/10.1080/01490400.2013.761905.

Martocchio, J., & Webster, J. (2006). Effects of feedback and cognitive playfulness on performance in microcomputer software training. *Personnel Psychology*, *45*(3), 553–578. https://doi.org/10.1111/j.1744-6570.1992.tb00860.x.

Mazza, C., Ricci, E., Biondi, S., Colasanti, M., Ferracuti, S., Napoli, C., & Roma, P. (2020). A nationwide survey of psychological distress among Italian people during the COVID-19 Pandemic: Immediate psychological responses and associated factors. *International Journal of Environmental Research and Public Health*, *17*(9), 3165. https://doi.org/10.3390/ijerph17093165.

McGhee, P. (2003). *Personality Traits.* Cambridge, UK: Cambridge University Press.

McMillan, D. (2017). Towards the playful professional. *Playful Teaching and Learning*, 198-212.

Minahan, J., Falzarano, F., Yazdani, N., & Siedlecki, K. L. (2021). The COVID-19 pandemic and psychosocial outcomes across age through the stress and coping framework. *The Gerontologist*, *61*(2), 228-239.

Morales-Vives, F., Dueñas, J. M., Vigil-Colet, A., & Camarero-Figuerola, M. (2020). Psychological variables related to adaptation to the COVID-19 lockdown in Spain. *Frontiers in Psychology*, *11*, 2438.

Natvig, G. A. (2003). Associations between psychosocial factors. *International Journal of Nursing Practice*, 166-175.

Ng, S. M. (2013). Validation of the 10-item Chinese perceived stress scale in elderly service workers: one-factor versus two-factor structure. *BMC psychology*, *1*(1), 1-8.

Ornell, F., Halpern, S. C., Kessler, F. H. P., & Narvaez, J. C. D. M. (2020). The impact of the COVID-19 pandemic on the mental health of healthcare professionals. *Cadernos de Saude Publica*, *36*, e00063520.

Orsolini, L., Latini, R., Pompili, M., Serafini, G., Volpe, U., Vellante, F., ... & De Berardis, D. (2020). Understanding the complex of suicide in depression: from research to clinics. *Psychiatry Investigation*, *17*(3), 207.

Örücü, M. Ç., & Demir, A. (2009). Psychometric evaluation of perceived stress scale for Turkish university students. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *25*(1), 103-10.

Osimo, S. A., Aiello, M., Gentili, C., Ionta, S., & Cecchetto, C. (2021). The influence of personality, resilience, and alexithymia on mental health during COVID-19 pandemic. *Frontiers in Psychology*, *12*, 341.

Parasuraman, S., & Hansen, D. (1987). Coping with work stressors in nursing: Effects of adaptive versus maladaptive strategies. *Work and Occupations*, *14*(1), 88-105.

Phua, D. H. (2005). Coping responses of emergency physicians and nurses to the 2003 severe acute respiratory syndrome outbreak. *Academic Emergency Medicine*, *12*(4), 322–328. https://doi.org/10.1197/j.aem.2004.11.015.

Proyer, R. T. (2012). Development and initial assessment of a short measure for adult playfulness: The SMAP. *Personality and Individual Differences*, 989-994.

Proyer, R. T. (2013). The well-being of playful adults: Adult playfulness, subjective well-being, physical well-being, and the pursuit of enjoyable activities. *The European Journal of Humour Research*, *1*(1), 84-98.

Proyer, R. T. (2017). A new structural model for the study of adult playfulness: Assessment and exploration of an understudied individual differences variable. *Personality and Individual Differences*, *108*, 113–122. https://doi.org/10.1016/j.paid.2016.12.011.

Proyer, R. T. (2020). Assessing Other-Directed, Light hearted, Intellectual, and Whimsical Playfulness in Adults. *European Journal of Psychological Assessment*, 624-634.

Proyer, R. T., Gander, F., Bertenshaw, E. J., & Brauer, K. (2018). The positive relationships of playfulness with indicators of health, activity, and physical fitness. *Frontiers in Psychology*, *9*, 1440.

R Core Team. (2022). R: A Language and Environment for Statistical Computing. https://www.r-project.org/

Rice, L. (2009). Playful learning. *Journal for Education in the Built Environment*, *4*(2), 94-108.

Roberti, J. W., Harrington, L. N., & Storch, E. A. (2006). Further psychometric support for the 10-item version of the perceived stress scale. *Journal of College Counseling*, *9*(2), 135-147.

Rossi, R., Socci, V., Pacitti, F., Di Lorenzo, G., Di Marco, A., Siracusano, A., & Rossi, A. (2020). Mental health outcomes among frontline and second-line health care workers during the coronavirus disease 2019 (COVID-19) pandemic in Italy. *JAMA Network Open*, *3*(5), e2010185-e2010185.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1-36.

Salari, N., Hosseinian-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulpoor, S., Mohammadi, M., ... & Khaledi-Paveh, B. (2020). Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Globalization and Health*, *16*(1), 1-11.

Saunders, I., Sayer, M., & Goodale, A. (1999). The relationship between playfulness and coping in preschool children: A pilot study. *American Journal of Occupational Therapy*, *53*(2), 221–226. https://doi.org/10.5014/ajot.53.2.221.

Shen, X. S., Chick, G., & Zinn, H. (2014). Adult playfulness trait scale. *PsycTESTS Dataset*. https://doi.org/10.1037/t61751-000.

Smith, B. S. (1966). Piage on Play: A critique. *Psychological Review*, 104-110.

Staempfli, M. B. (2007). Adolescent playfulness, stress perception, coping, and well-being. *Journal of Leisure Research*, 393-412.

Teasdale, E., Yardley, L., Schlotz, W., & Michie, S. (2012). The importance of coping appraisal in behavioural responses to pandemic flu. *British Journal of Health Psychology*, 17(1), 44-59. doi: 10.1111/j.2044-8287.2011. 02017.x.

Tegano, D. (1990). Relationship of tolerance of ambiguity and playfulness to creativity. *Psychological Reports*, 1047-1056.

Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., McIntyre, R. S., Choo, F. N., Tran, B., Ho, R., Sharma, V. K., & Ho, C. (2020). A longitudinal study on the mental health of general population during the COVID-19 epidemic in China. *Brain, Behavior, and Immunity*, *87*, 40–48. https://doi.org/10.1016/j.bbi.2020.04.028.

World Health Organization. (2020, December 28). *Listings of WHO's response to COVID-19*. https://www.who.int/news/item/29-06-2020-covidtimeline.

Yarnal, X. L. (2011). The role of playfulness in the leisure-stress coping process among emerging adults: An SEM analysis. *Leisure*, 191-209.

Zalsman, G., Stanley, B., Szanto, K., Clarke, D. E., Carli, V., & Mehlum, L. (2020). Suicide in the Time of COVID-19: Review and Recommendations. *Archives of Suicide Research*, *24*(4), 477-482.

Appendix

Supplemental Tables

Supplemental Table 1 – Single Mediator Regression Estimates

Table 1.A – Single Mediation Model Estimates Using Perceived Self-Efficacy as the Mediator

Parameter Estimates	Estimate	Standardized Estimate (β)	*SE*	*95% CI*		*p*
				LB	*UB*	
Direct Effects of Playfulness						
Adaptive Coping	0.195	0.303	0.024	0.148	0.241	<0.001
Maladaptive Coping	0.052	0.154	0.014	0.025	0.08	<0.001
Supportive Coping	0.149	0.198	0.029	0.092	0.206	<0.001
Perceived Self-Efficacy	0.255	0.316	0.034	0.189	0.322	<0.001
Mediator Effects						
Perceived Self-Efficacy predicting						
Adaptive Coping	0.268	0.338	0.032	0.206	0.330	<0.001
Maladaptive Coping	-0.221	-0.526	0.029	-0.277	-0.165	<0.001
Supportive Coping	0.150	0.161	0.037	0.078	0.222	<0.001
Indirect Effects						
Playfulness through Perceived Self-Efficacy						
Adaptive Coping	0.068	0.107	0.012	0.045	0.092	<0.001
Maladaptive Coping	-0.056	-0.166	0.010	-0.076	-0.037	<0.001
Supportive Coping	0.038	0.051	0.010	0.018	0.059	<0.001
Total Effects of Playfulness						
Adaptive Coping	0.263	0.410	0.024	0.217	0.309	<0.001
Maladaptive Coping	-0.004	-0.012	0.014	-0.031	0.023	0.764
Supportive Coping	0.187	0.249	0.028	0.133	0.241	<0.001

Table 1.B – Single Mediation Model Estimates Using Perceived Helplessness as the Mediator

Parameter Estimates	Estimate	Standardized Estimate (β)	*SE*	*95% CI*		*p*
				LB	*UB*	
Direct Effects of Playfulness						
Adaptive Coping	0.257	0.409	0.023	0.211	0.303	<0.001
Maladaptive Coping	0.035	0.086	0.014	0.008	0.062	0.01
Supportive Coping	0.195	0.258	0.028	0.140	0.251	<0.001
Perceived Helplessness	-0.114	-0.137	0.034	-0.179	-0.048	0.001
Mediator Effects						
Perceived Helplessness predicting						
Adaptive Coping	0	0	0.027	-0.053	0.053	0.999
Maladaptive Coping	0.327	0.665	0.033	0.261	0.392	<0.001
Supportive Coping	0.074	0.081	0.035	0.007	0.142	0.031
Indirect Effects						
Playfulness through Perceived Helplessness						
Adaptive Coping	0	0	0.003	-0.006	0.006	0.999
Maladaptive Coping	-0.037	-0.091	0.011	-0.060	-0.015	0.001
Supportive Coping	-0.008	-0.011	0.005	-0.018	0.001	0.084
Total Effects of Playfulness						
Adaptive Coping	0.257	0.409	0.023	0.211	0.302	<0.001
Maladaptive Coping	-0.002	-0.005	0.016	-0.034	0.029	0.894
Supportive Coping	0.187	0.247	0.028	0.132	0.241	<0.001

CHAPTER III

How Should You Analyze Regression Relations between Multi-Item Scales? Measurement Error and Other Considerations

Christopher Clifford, Timothy Hayes, Stefany Coxe, and Leslie D. Frazier

Department of Psychology

Florida International University

Author Note

We have no known conflicts of interest to disclose.

This research was not funded.

The data presented in this manuscript are available upon reasonable request.

Correspondence concerning this article should be addressed to Christopher Clifford, M.S., Department of Psychology, Florida International University, 11200 S. W. 8th Street, Miami, FL 33199. Email: cclif003@fiu.edu

Abstract

In this study, we perform a pedagogical demonstration of different regression method performance on a previously used structural model from Clifford et al. (2022). The model included five latent variables, each of which were measured from multi-item scales. We compare the performance of the estimates, standard errors, and semi-partial correlations of the structural equation model (SEM) against regression estimates taken from: 1) summed scores, 2) factor scores regressions, and 3) a novel two-stage estimation procedure called structural after measurement (SAM). Results confirm the performance and preference of SEM for regression modeling with latent variables using multi-item scales, we also demonstrate that SAM is a viable alternative under certain circumstances in the data.

keywords: regression, latent variable, structural equation modeling, sum score, factor score regression, structural after measurement

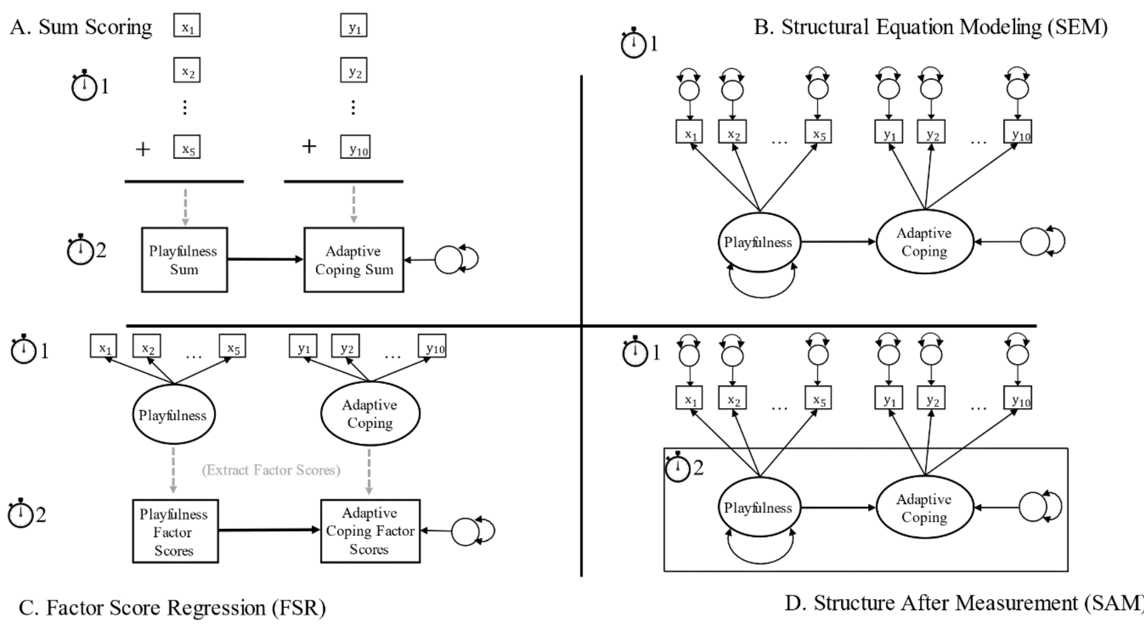
How Should You Analyze Regression Relations between Multi-Item Scales? Measurement Error and Other Considerations

Psychology involves the study of complex, multifaceted constructs that are not directly observable yet often influence behavior, such as individuals' internal affective and cognitive states. Such constructs are frequently measured indirectly using multi-item scales with each item capturing an aspect of the construct. For example, Clifford et al. (2022) were interested in predicting three types of coping behaviors (y_1– y_3) from playfulness (x), perceived self-efficacy (m_1) , and perceived helplessness (m_2) in a parallel mediation model, with each construct measured using a multi-item scale. To illustrate the indirect nature of psychological measurement, consider the measure of adaptive coping, which asks participants whether they tend to take concrete steps to improve difficult situations, engage in active acceptance of them, and express, rather than hold back, their negative feelings about negative situations. All of these items reflect aspects of adaptive coping, although not all participants engage in all of them. Although participants' adaptive coping levels are not directly observable, they can be inferred from their responses to these items.

Researchers are typically interested in the relationships among constructs rather than simply the constructs themselves. Multiple statistical approaches are available that could be used to estimate these relationships, though often researchers may overly rely on the method they are most familiar with. In this paper, we will discuss the advantages and limitations of four approaches to estimating regressions among the constructs measured by multi-item scales: 1) using total scale scores, 2) using latent factors in a structural equation model (SEM), 3) using factor scores, and 4) using a novel two-stage estimation procedure recently introduced in the methodological literature called Structural After Measurement (SAM, Rosseel & Loh, 2022). As we will discuss, while the most commonly used approaches (e.g., sum scoring) may not be the most accurate, the most accurate approaches (e.g., latent variable SEM) may not be the most feasible. A primary goal of this paper is to highlight newly developed methods designed to provide both high accuracy and high feasibility in a variety of realistic scenarios.

Figure 1.

A Visual Comparison of Sum Scoring, Structural Equation Modeling, Factor Score Regression and Structural After Measurement Regression Estimation Approaches



Although we will demonstrate each approach with the full model from Clifford et al. (2022) in a subsequent section, in the discussion that follows we will restrict our attention to a simplified illustrating scenario. Imagine that a developmental researcher is interested in predicting an outcome, *y*, adaptive coping behaviors (proactive, solution-based responses to stressors) from a single predictor (*x*), individuals' dispositional playfulness (a construct capturing individuals' responses to situations using humor, amusement, and positivity). Figure 1 depicts four different statistical approaches using this illustrative scenario.

Arguably the most common method for modeling regressions among multi-item measures is sum scoring. Sum scoring involves adding a participant's scores across all items in each scale and using those sums as items in a regression analysis (for more on these measurement issues, see Curran et al., 2016; Kuhfeld & Soland, 2022). As seen in Figure 1A, sum scoring proceeds in two steps. A researcher would first sum participants' scores on all playfulness items (*x*) and on all coping items (*y*). Then, an observed variable regression would be performed using these total scores. The advantages of sum scoring include the straightforward nature of computing scale scores and the ease with which sum score composites can be incorporated into Ordinary Least Squares (*OLS*) analyses that are familiar to virtually all researchers working in psychology. Simply stated, *anybody* can use sum scores in their regression models without the need for specialized, advanced statistical training.

Unfortunately, while the simplicity of sum scoring is appealing, the validity of this method is undercut by a key limitation: it treats participants' responses as perfect, error-free measures of their true levels of each construct. More concretely, this approach treats participants' self-reported responses to playfulness and coping items as if they were perfect proxies for participants' true levels of playfulness and coping. This idealistic assumption is unlikely to be met in practice, however, both because even carefully constructed questionnaire items are rarely perfectly clear and unambiguous to all participants and because participants' responses may be affected by chance fluctuations (e.g., a loud noise in the testing environment that distracts a participant's attention from carefully reading an item; for more on sources of error variability, see van Bork et al., 2022). In the psychometric literature, participants' true levels of playfulness and coping are described as their *true scores* on these constructs; differences between observed questionnaire responses and these true scores (e.g., caused by item ambiguities or chance factors affecting responding) are referred as *measurement error*.

A key assumption of linear (OLS) regression models is that all model variables[1] are perfectly measured, with no measurement error (see Cohen et al., 2003, p. 119). When this assumption is not met (as is nearly always true in practice), a variety of negative consequences result, including attenuated and magnified estimates of model regression coefficients and increased Type I error rates (for further detail, see Cole & Preacher, 2014). Stated more concretely, depending on the amount of measurement error in each variable and the complexity of a given model, regression coefficient estimates might be too small or too large, as will key measures of effect size (for example R^2). In such situations the results of significance tests will be untrustworthy.

These harmful effects of measurement error are concerning because of their implications for both validity and replicability and because the effect sizes reported in a given study are often used as the basis for a priori power analyses that inform sample size

[1] Particularly the predictors.

decisions in subsequent studies. If, for example, the ΔR^2 effect size due to a key predictor in an observed variable regression is artificially magnified due to unaddressed measurement error, a power analysis based on this effect size will suggest that adequate power (e.g., of .80 or .95) could be achieved with a smaller sample size than is actually necessary, leaving follow-up studies based on this analysis critically underpowered.

To avoid these deleterious effects, researchers can switch from an observed variable regression analysis to a latent variable structural equation model (*SEM*) that strategically separates true score from error variation in each construct, as shown in Figure 1B. In brief, a foundational assumption of the latent variable SEM approach is that participants' responses to survey items (small squares, e.g., x_1, x_2, …, x_5) covary *systematically* because the items are tapping the same construct (large circles, playfulness and coping, with one-headed arrows pointing at the items measuring each construct). Although participants' responding may be error-prone and imperfect (for reasons mentioned above and expanded upon below), all else being equal, participants higher in their true levels of playfulness should be expected to select higher response options for all playfulness items in Figure 1B than individuals lower in their true levels of playfulness, and the same should be true of the coping items.

Whereas the large circles in Figure 1B depict latent common factors–participants' true levels of playfulness and coping–that cause the covariation among the items (squares), the small circles in Figure 1B represent random sources of measurement error caused by item ambiguities and chance factors such as brief distractions while reading some items. SEM assumes that all relationships among observed variables are due to the common latent factor (i.e., the large circle with one-headed arrows pointing at all items);

there are no direct relationships between measurement errors (small circles) across items. This follows directly from the assumption that measurement error represents chance (random) fluctuations. If, for example, a loud noise in the room distracts participants from carefully reading an item, this should exert no effect on their ability to focus on reading the other items if the room is subsequently quiet.

Because SEM models regression relationships between true score latent factors (circles, in Figure 1B) rather than between error prone observed variables (squares, in Figure 1A), it avoids the attenuation and magnification of regression coefficients seen in sum scoring methods that fail to separate true scores from measurement error. As a result, the estimates from latent variable SEMs will be more accurate than those from observed variable regressions using sum scores and power analyses based on these estimates will be more trustworthy. For these reasons, SEM is considered a gold standard method for modeling regression relationships using multi-item scales.

Despite its many advantages, SEM's sample size requirements present a practical limitation. SEM requires large samples, typically greater than n = 200 and often exceeding n = 500 (Kline, 2023). These sample size requirements can prove prohibitive for developmental scientists studying specialized or vulnerable populations, such as infants, or time-intensive methods, such as EEG. How might developmental researchers address measurement error in their analyses when their sample sizes are often smaller than recommended cutoffs for SEM?

A primary reason that sample size recommendations for SEMs are so large is that latent variable regression models like the one in Figure 1B are composed of multiple parts that must all be estimated simultaneously. More precisely, this model includes both a measurement portion (i.e., a factor analysis portion, in which each latent variable is measured by a set of observed items, depicted as the large circles pointing to the squares with one-headed arrows) and a structural portion (i.e., a regression portion, in which the latent outcome, coping, is regressed on the latent predictor, playfulness, depicted by the large circles connected by a one-headed arrow). Rather than estimating both portions of the model simultaneously, as occurs with SEM, they could be estimated sequentially. This would serve to break the larger model into a series of smaller submodels, each with more modest sample size requirements than the full SEM.

Among researchers trained in SEM or factor analysis methods, an intuitive sequential approach would be to first estimate factor scores from each latent variable and then run a regression between those factor scores in a subsequent step. This approach has been called Factor Score Regression (FSR, see Croon, 2002; Devlieger et al., 2016; Devlieger & Rosseel, 2017; Hayes & Usami, 2020a, 2020b; Hoshino & Bentler., 2013; Skrondal & Laake, 2001) and is depicted for the playfulness model in Figure 1C. Like SEM, the intuition behind this approach begins with the fact that the observed scores in one's dataset are imperfect measures of their true scores (e.g., participants' self-reported responses to playfulness questions may not perfectly reflect their true levels of dispositional playfulness). As described, SEM approaches this problem by using the relationships between observed variables to infer the relationships among participants' true scores in a model that accounts for measurement error–for example, estimating the regression coefficient describing the relationship between true score playfulness and true score coping in Figure 1B. Importantly, SEM does not explicitly estimate true scores;

rather, it infers (e.g., regression) relationships among true score latent variables based on the covariances of the observed variables.

By contrast, the FSR approach begins by attempting to explicitly estimate participants' true scores on each construct (e.g., playfulness and coping). Conceptually, this is accomplished in a manner analogous to calculating predicted *y*-hat values in a regression (in fact, the so-called "regression method" for estimating factor scores does exactly this, see Thompson, 1934, for details, 1946; see also Thurstone, 1935, 1947). In contrast to familiar *y*-hat values from linear regression, which predict an *observed outcome* from the values of a set of *observed predictors*, factor score estimation approaches seek to predict the values of an *unobserved (latent) outcome* from a set of *observed items*. Once these predicted factor scores (factor score *y*-hat values) are calculated, they can be saved as entirely new variables in one's dataset. This is visualized by the dashed arrows from each factor model (circle) to each observed variable (square) linking Steps 1 and 2 in Fig. 1C. Once this is accomplished, the estimated factor scores can be used in a standard linear regression analysis, treated as any other observed variables.

The primary advantage of FSR is that its multi-step approach to estimation requires considerably smaller sample sizes than SEM. At the first step, this is because factor scores can be "extracted" from each factor model one at a time, with the sample size requirements for each separate factor model being much lower than for the full SEM of Figure 1B. Then, on the second step, sample size requirements for a simple linear regression are relatively modest (see Cohen, 1988). The primary disadvantage of FSR stems from the fact that because factor scores are estimates of participants' true scores, not the true scores themselves, regression estimates computed from these factor scores will not be as accurate as regression estimates performed using the SEM approach of Fig 1B. Conceptually, this is because factor score estimates do not fully incorporate measurement error (for more on this topic, see Steiger & Schönemann, 1978).

What is needed is an approach that captures both the accurate true score estimation of SEM while incorporating the sample size benefits of FSR's sequential estimation approach. Only recently has such an approach been developed and implemented into software via what Rosseel and Loh (2022) have termed a "Structural After Measurement" (*SAM*) approach. As the name suggests, SAM involves computing the regression between the constructs (called the 'structural' part of the model in SEM; shown in Step 2 of Figure 1D) *after* the estimating the factor models (called the 'measurement' part of the model in SEM; shown in Step 1 of Figure 1D). Although the details of the SAM algorithm are complex, conceptually this approach breaks SEM-style estimation into a sequence of steps, each requiring a lower sample size than the full latent variable regression model of Figure 1B. In terms of the running example, this is accomplished by first estimating the factor models for playfulness and coping separately (the large circles labeled playfulness and coping pointing at the square boxes in Figure 1D) and then estimating the regression between the factors (as shown in the Step 2 box of Figure 1D, depicting coping and playfulness connected by a one-headed arrow), within a single multi-stage model.

By accurately estimating both the true score and error in the model while retaining the multi-step aspects of FSR, SAM incorporates the advantages of both methods. Critically, for developmental researchers, SAM is able to give comparable results to SEM with smaller sample sizes. As implemented by Rosseel and Loh in the `lavaan` (Rosseel, 2012) package in `R` (R Core Team, 2022) using SAM requires minimal changes from traditional SEM syntax, with multi-step estimation handled behind-the-scenes. With concrete benefits and a strong ease of use, SAM is poised to become a powerful alternative to SEM for moderate-to-low sample sizes. In contrast to the full SEM of Figure 1B, SAM only requires a large enough sample size to fit each single factor model one at a time.

Methods

Overview of the Empirical Analysis Model

For the remainder of the paper, we will be using an extended empirical example using data from Clifford et al., (2022) that incorporates more variables and a more realistic model–specifically, indirect effects in a two mediator model, as diagrammed in Figure 2. In addition to adaptive coping (defined above), maladaptive coping and supportive coping serve as second and third outcome variables. Maladaptive coping refers to avoidant, self-critical, or destructive tendencies in response to adversity whereas supportive coping involves coping with adversity using external social or spiritual support systems. Clifford et al., (2022) hypothesized that the relationship between playfulness and coping would be mediated by current perceived stress, as measured by self-efficacy and helplessness. Self-efficacy refers an individual's belief in their capacity to manage or control a stressor (Bandura, 1978). By contrast, helplessness refers to feelings of a lack of agency and inability to manage or control a stressor (Peterson et al., 1993). Perceived helplessness and perceived self-efficacy are a two factor structure for the Perceived Stress Scale (PSS, Cohen et al., 1983), a structure which has been suggested by several scholars (Golden-Kreutz et al., 2004; Hewitt et al., 1992; Khalili et al., 2017; Örücü & Demir, 2009; Roberti et al., 2006).

Each construct can be conceptualized as a latent variable (circle in Figure 2, in line with the prior discussion of Figure 1) measured by a multi-item scale (with items indicated by grayscale squares in Figure 2, in line with the prior discussion of Figure 1). Playfulness was measured by the Short Measure of Adult Playfulness (SMAP; Proyer et al. 2012). The SMAP is a 5 item Likert-questionnaire with item responses ranging from 1 (strongly disagree) to 4 (strongly agree). Perceived Self-Efficacy (PSE) and Perceived Helplessness (PH) were measured using their respective subscales from the Perceived Stress Scale (PSS; Cohen et al., 1983; see also Ng, 2013), with item response options ranging from 0 (never) to 4 (very often). Adaptive, Maladaptive, and Supportive Coping were measured via their respective subscales from the COPE-18 scale (Carver, 1997), an 18 question Likert-type scale with response options ranging from 1 ("I haven't been doing this at all") to 4 ("I've been doing this a lot"). For further details see Clifford et al. (2022).

The full sample consists of $N = 694$ undergraduate students participating in a study of playfulness and coping strategies for course credit[2]. The study had full IRB approval and met the APA JARs standards for quantitative data. A detailed description of the sample characteristics can be found in Clifford et al. (2022). As described above, our

[2] Note that this sample size, and therefore the model results reported later, differs slightly from that reported in Clifford et a. (2022). For the present demonstration, we used only observations with complete cases (no missing data). Although complete case analysis is not generally recommended, we employed this approach in order to compare analysis methods without the additional confounding due to different missing data handling techniques. Importantly, whereas conventional SEM allows missing data handling using full information maximum likelihood (FIML) estimation, SAM currently requires complete cases.

gold standard method, SEM, produces extremely accurate estimates, but requires a large sample size. An additional benefit of this large sample is that it allows us to compare the performance of each alternative data analysis method to gold standard SEM in a context where discrepancies between them cannot be attributed to sample size. Note, however, that this is a larger sample than is often feasible to collect in developmental psychology, for the reasons mentioned above. In order to demonstrate how each of the methods compares to SEM in a more realistic sample, we therefore also replicated all analyses in a random subsample of $N = 150$.

Results

Table 1 presents reliability estimates (Cronbach's alpha; Cronbach, 1951) for all measures in the full sample and the subsample. Tables 2 and 3 and Tables 4 and 5 show model results using each of the different data analysis methods in the full sample and subsample, respectively. In Table 1, we present standardized regression (path) coefficient estimates, standard errors, model R^2 values, and squared semipartial correlations. Standardized regression coefficients reflect the change in a given outcome, in standard deviation units, expected to result from a standard deviation increase in a predictor, controlling for any other predictors in the model. Using standardized coefficients allows us to more easily compare values across data analysis methods that differ in their default scaling (e.g., the units of factor scores in an FSR analysis are naturally different than those of sum scores or latent variables estimated via SEM). The standard errors associated with each coefficient reflect the estimated sampling variability of each estimator. Importantly, note that FSR and sum scoring methods should have smaller standard errors because they ignore measurement error and its associated uncertainty (for a more extended discussion of tradeoffs between observed and latent variable approaches, see Ledgerwood & Shrout, 2011). The model R^2 values quantify the proportion of variance in a given outcome explained by all predictors of that outcome. By contrast, each squared semipartial correlation reflects the proportion of variance in an outcome explained by a given predictor, after partialing out overlapping variance shared with the other predictors of that outcome. Thus, whereas the model R^2 value represents an overall effect size for a regression model, each squared semipartial correlation represents an effect size for a specific predictor-to-outcome relationship.

Table 2 highlights the bias in each method compared to SEM as a percentage difference. For example, in the case of the regression coefficient for sum scoring, this would be quantified by the formula $\left[\left(\beta_1^{(\text{Sum Scoring})} - \beta_1^{(\text{SEM})}\right)/\beta_1^{(\text{SEM})}\right] \times 100$. The formula functions the same way for our other data analysis methods and other quantities (e.g., standard errors, semipartial correlation) by substituting in the appropriate quantities. Percentage bias is a commonly used metric in the quantitative and statistical literature because it provides useful context to help clarify the magnitude by which raw estimates differ from a gold standard (here, SEM). Importantly, negative values of percentage bias reflect scenarios in which a data analysis method results in estimates that are too small compared to SEM whereas positive values reflect scenarios in which a data analysis method results in estimates that are too large compared to SEM. Percentage bias estimates greater than 10 in absolute value are widely considered problematic in the methodological literature (see Enders & Bandalos, 2001; Wheaton et al., 1977).

Full Sample Results

Sum Scoring Results

As described above, Table 3 shows percent bias in all estimates compared to SEM in the full sample. Recall that we compared each method to SEM, the gold standard. Several trends are evident in these percent bias estimates. As expected based on previous methodological literature (Bollen, 1989; Cole & Preacher, 2014; Ledgerwood & Shrout, 2011; Westfall & Yarkoni, 2016) estimates based on sum scores exhibit both positive biases (i.e., they are too large) and negative biases (i.e., they are too small) in different parts of the mediation model. Specifically, paths involving only a single predictor to an outcome, such as those between playfulness (x) and each mediator (m_1 and m_2, perceived self-efficacy and perceived helplessness), exhibit negative biases due to *attenuation of correlation* (Spearman, 1987). For example, the sum score estimate of the regression coefficient of playfulness predicting perceived helplessness exhibits a percentage bias of -18.88. That is, this estimate (-0.12 in Table 2) was 18.88% smaller than the corresponding estimate from the latent variable SEM (–0.14 in Table 2).

Paths involving multiple predictors to an outcome (e.g., those in which x, m_1, and m_2 predict each y variable–the coping subscales in the substantive example), by contrast, exhibit positive as well as negative biases. In this analysis, the sum score estimate of the regression coefficients of playfulness predicting each coping style consistently too large. For example, the sum score estimate of the regression coefficient of playfulness predicting adaptive coping exhibits a percentage bias of 28.33. That is, this estimate (0.30 in Table 2) was 28.33% larger than the corresponding estimate from the latent variable SEM (0.23 in Table 2). Although percentage bias estimates ranged from negative to positive, sum scoring generally resulted in problematic levels of bias in comparison to gold standard SEM in most relationships in the model.

In addition to interpreting the magnitude of percentage bias estimates in Table 3 and examining the raw estimates in Table 2, it is also helpful to examine the squared semipartial correlation values in Table 2, as these correspond to a familiar effect size metric. For example, examining the regression of adaptive coping on perceived self-efficacy, SEM suggests that self-efficacy uniquely accounts for eighteen percent of the variance in adaptive coping, controlling for all other predictors, whereas sum scoring suggests that self-efficacy uniquely accounts for only 10 percent of the variance in adaptive coping. Although the differences are not always so stark, sum scoring generally resulted in squared semipartial correlation values that were different from those of gold standard SEM in most relationships in the model.

Finally, as expected, the standard errors resulting from the sum score analysis were generally smaller than those resulting from latent variable SEM. As described above, this is because linear regression methods assume all predictors are measured without error whereas latent variables SEMs explicitly incorporate error into the model, adding addition sources of uncertainty to the standard errors.

Factor Score Regression (FSR) Results

We now turn to the FSR columns in Tables 2 and 3. As expected based on the previous methodological literature (Croon, 2002; Devlieger et al., 2016, 2019; Devlieger & Rosseel, 2017; Hayes & Usami, 2020a, 2020b; Hoshino & Bentler., 2013; Skrondal & Laake, 2001), estimates based on factor scores also exhibit both positive biases (i.e., they are too large) and negative biases (i.e., they are too small) in different parts of the mediation model, although not in the same pattern as sum scores. Like sum scoring, FSR generally resulted in problematic levels of bias in comparison to gold standard SEM in most relationships in the model. For example, the FSR estimate of the regression coefficient of self-efficacy predicting adaptive coping exhibits a percentage bias of –31.85. That is, this estimate (0.40 in Table 2) was 31.85% smaller than the corresponding estimate from the latent variable SEM (0.58 in Table 2). Examining the corresponding squared semipartial correlation effect sizes in Table 2, SEM suggests that self-efficacy uniquely accounts for 18% of the variance in adaptive coping, controlling for all other predictors, whereas sum scoring suggests that self-efficacy uniquely accounts for only 11 percent of the variance in adaptive coping. Unsurprisingly, FSR generally resulted in squared semipartial correlation values that were different from those of gold standard SEM in most relationships in the model as well as standard errors that were predictably smaller than gold standard SEM.

Structural After Measurement (SAM)

In contrast to both sum scoring and FSR, estimates based on the structural after measurement (SAM) approach exhibit negligible bias throughout the model. That is, SAM estimates in this analysis were near-identical to those from SEM. For example, even the most biased SAM estimate, the coefficient of helplessness predicting maladaptive coping, was only 7.25% smaller than the corresponding SEM estimate (–7.25 in Table 3). Turning to the raw estimates in Table 2, we see that this corresponds to an estimate of .60 from the SEM analysis and an estimate of .59 from the SAM analysis. The squared semipartial correlations corresponding to these estimates were also near-identical: .21 and .20, respectively. Finally, as expected, the standard errors returned by SAM were near-identical to those returned by SEM.

Full Sample Results Summary

In sum, in the full sample of $N = 694$, participants estimated regression coefficients based on sum scoring and FSR analyses exhibited substantial biases in both the positive and negative directions. Moreover, the standard errors of both methods were predictably smaller than those of SEM (which explicitly accounts for uncertainty in the estimation of the latent variables in the model). By contrast, both the estimated regression coefficients and the estimated standard errors based on SAM analysis were comparable to those of gold standard SEM (because both models explicitly account for uncertainty in the estimation of the latent variables in the model, with the primary difference being in whether this is done in a single step or a multi-step estimation process).

N = 150 Subsample Results

As described above, Table 4 displays the estimates, standard errors, model R-squared values, and squared semipartial correlations in the small subsample of $N = 150$ participants. Correspondingly, Table 5 displays percent bias in all estimates compared to SEM in the subsample. Examining these Tables, the same trends are evident as in the full sample. Once again, estimates based on sum scores exhibited both positive and negative biases in different parts of the mediation model. Again, attenuation of correlation caused estimates of paths involving only a single predictor to an outcome (e.g., paths from x to each mediator) to be too small compared to SEM whereas estimates of paths involving multiple predictors to an outcome (e.g., those in which x, m_1, and m_2 predict each y

variable) were both too large and too small compared to SEM in different parts of the model. FSR exhibited similar results as well, with problematic levels of positive and negative bias in the majority of pathways. As in the full sample, sum scoring and FSR results returned smaller standard errors compared to those of SEM. Finally, SAM produced near-identical results to gold standard SEM.

In sum, in the subsample of $N = 150$ participants, estimated regression coefficients based on sum scoring and FSR analyses again exhibited substantial biases in both the positive and negative directions. Moreover, the standard errors of both methods were once again smaller than those of SEM. Once again, even in this smaller sample setting, both the estimated regression coefficients and the estimated standard errors based on SAM analysis were comparable to those of gold standard SEM.

Discussion

Developmental scientists are often interested in estimating regressions among constructs of theoretical interest that are measured with error. When these constructs are measured using multi-item scales, it is possible to decouple participants' true levels of the constructs (their true scores) responses from measurement error. The gold standard approach to this task is to run latent variable regressions in the SEM framework. However, this approach is not always feasible in smaller samples. To address this limitation, the current paper described and demonstrated three alternative approaches that can be used with smaller sizes: observed variable regression using sum scores, observed variable regression using estimated factor scores, and a multi-stage latent variable regression approach called structural after measurement (SAM). We compared these methods' performance to that of gold-standard SEM in both a large sample dataset of N = 694 participants and in a smaller subsample of N = 150 participants.

As expected from the methodological literature (Roseel & Loh, 2022), the SAM approach performed comparably to gold-standard SEM in both samples whereas sum scoring and factor score regression exhibited bias in both their estimates and standard errors. One noteworthy aspect of this demonstration was that coefficients estimated using sum scoring were both magnified (too large) and attenuated (too small) in different parts of the model. The reality that coefficients may be inflated as well as deflated bears emphasizing because it goes against the commonly held intuition that *all* effects might be deflated (weakened; attenuated) by measurement error, such that significance tests represent 'conservative tests.' After all, if true score variation represents 'signal' and measurement error represents 'noise,' commonsense suggests that the increased noise in sum score composites would make it harder to detect the signal, not easier, resulting in weaker coefficients less likely to reach the threshold for statistical significance. The reality, however, is more complex (see Cole & Preacher, 2014 for a thorough discussion of these issues), as demonstrated here.

Relatedly, another troubling aspect of this demonstration was that the biases evident in regression using sum scores and factor scores remained problematically large even in the full sample of N = 694 participants–a sample size substantially larger than many researchers are typically able to collect in practice. Once again, this finding is in line with the predictions from the methodological literature, which suggest that biases due to unaddressed measurement error will remain present no matter how large the sample (see e.g., Bollen, 1989; Cole & Preacher, 2014; Westfall & Yarkoni, 2016). In contrast to these methods, SAM performed near-identically to gold-standard SEM. This is again unsurprising, as both of these methods explicitly and fully model measurement error.

Whereas regressions using sum scores or factor scores performed poorly even when the sample size was large, the SAM approach performed well regardless of sample size. Although SEM is generally discouraged in smaller samples, the SAM approach was explicitly designed as a small sample method. Recent simulation research by Rosseel and Loh (2022) suggests that SAM outperforms SEM when sample sizes are small, particularly when the reliability of the measures employed in an analysis is low. Under these conditions, SEMs exhibit higher rates of convergence failures (i.e., instances of the models not successfully running) and noteworthy levels of small sample bias in their estimates. Despite the high reliability of the scales obtained in the Clifford et al., (2022) data (as shown Table 1) and the fact that both SEM and SAM models successfully converged (ran) in our illustrative example even with $N = 150$, we observed analogous performance differences in some cases. For example, examining the pathway from playfulness to perceived self-efficacy in Tables 2 and 4, we see that the SAM estimate of .274 in the $N = 150$ subsample comes relatively closer to reproducing the large sample SEM estimate of .286 than the corresponding estimate using SEM in the small 150 participant subsample (.267). Treating the large sample SEM estimate as an accurate estimate of the population value, this suggests that SAM performed better in recapturing this population value for the playfulness → self-efficacy pathway.

Taken together, these results demonstrate several key points. First, even when scales have higher reliability (i.e., small amounts of measurement errors, as suggested by the high alpha values observed in the example data), and a large sample size, regression using sum scores or factor scores should be avoided due to serious and unpredictable biases. Second, SAM performs comparably to SEM, making it a viable alternative to the gold standard approach. Third, SAM is actually superior to SEM in common situations, such as research involving lower reliability measures and smaller samples. Finally, SAM should be an option considered by developmental scientists especially in research situations with smaller sample sizes. More broadly, SAM may be a viable alternative for any social scientist using multi-item, survey data.

Because the SAM method was only recently developed and proposed in the methodological literature (Rosseel & Loh, 2022), many researchers may not yet be aware of it. One goal of this paper was to draw attention to this method as a promising go-to technique for estimating latent variable regressions. Not only is the method comparable to SEM in its performance, it is also comparable in its implementation: researchers familiar with using the `lavaan` package in `R` to run SEM analyses can run equivalent SAM models with only minor adjustments to their model syntax(e.g., using the `sam()` rather than the `sem()` function, while keeping the remaining model syntax the same). Though throughout our paper we emphasized the utility of the SAM approach in small samples, the method also carries additional benefits. For example, because SAM takes a multi-stage approach to estimation, fitting separate measurement models (CFA models) one-by-one before estimating the latent variable regressions, the consequences of misspecification in any single part of the model (e.g., omitting a cross-loading or correlated residual in a CFA model) are mitigated and localized when using this method. This stands in contrast to SEM, which estimates all parts a model (i.e., both the CFA and latent variable regressions) at once, with the consequence that misspecification anywhere in a larger model can lead to bias throughout.

In sum, when estimating regressions among constructs measured by multi-item scales, researchers face a choice between multiple potential approaches. Researchers comfortable and familiar with sum scoring may use this method as a default, but our review of the methodological literature suggests that this method should rarely be used due to its inability to account for measurement error. FSR faces similar concerns. Instead, those lucky enough to obtain large sample sizes should consider using latent variable SEM or multistage SAM to estimate their models. Conversely, researchers working with small-to-moderate sample sizes as is often the case in developmental science, should consider SAM as a first choice option, especially as additional features (e.g., missing data handling methods, categorical estimation) are implemented.

References

Bandura, A. (1978). Reflections on Self-Efficacy. *Advances in Behavior Research and Therapy*, *1*(4), 237–269.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley.

Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the brief COPE. *International Journal of Behavioral Medicine*, *4*(1), 92–100. https://doi.org/10.1207/s15327558ijbm0401_6

Clifford, C., Paulk, E., Lin, Q., Cadwallader, J., Lubbers, K., & Frazier, L. D. (2022). Relationships among adult playfulness, stress, and coping during the COVID-19 pandemic. *Current Psychology*, *0123456789*. https://doi.org/10.1007/s12144-022-02870-0

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Inc.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates Publishers.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983a). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(3), 385–396. https://doi.org/10.2307/2136404

Cohen, S., Kamarck, T., & Mermelstein, R. (1983b). Perceived Stress Scale. *Journal of Health and Social Behavior*, *24*, 386–396. http://www.mindgarden.com/products/pss.htm

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–315. https://doi.org/10.1037/a0033805

Cronbach, J. L. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/https://doi.org/10.1007/BF02310555

Croon, M. (2002). Using predicted latent scores in general latent structure models. In G. A. Marcoulides & I. Moustaki (Eds.), *Latent Variable and Latent Structure Models* (1st ed., pp. 195–223). Lawrence Erlbaum.

Curran, P. J., Cole, V., Bauer, D. J., Hussong, A. M., & Gottfredson, N. (2016). Improving factor score estimation through the use of observed background characteristics. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(6), 827–844. https://doi.org/https://doi.org/10.1080/10705511.2016.1220839

Devlieger, I., Mayer, A., & Rosseel, Y. (2016). Hypothesis Testing Using Factor Score Regression: A Comparison of Four Methods. *Educational and Psychological Measurement*, *76*(5), 741–770. https://doi.org/10.1177/0013164415607618

Devlieger, I., & Rosseel, Y. (2017). Factor score path analysis: An alternative for SEM? In *Methodology* (pp. 31–28). https://doi.org/https://doi.org/10.1027/1614-2241/a000130

Devlieger, I., Talloen, W., & Rosseel, Y. (2019). New Developments in Factor Score Regression: Fit Indices and a Model Comparison Test. *Educational and Psychological Measurement*, *79*(6), 1017–1037. https://doi.org/10.1177/0013164419844552

Enders, C. K., & Bandalos, D. L. (2001). The Relative Performance of Full Information Maximum Likelihood Estimation for Missing Data in Structural Equation Models. *Structural Equation Modeling*, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803

Golden-Kreutz, D. M., Browne, M. W., Frierson, G. M., & Andersen, B. L. (2004). Assessing stress in cancer patients: A second-order factor analysis model for the perceived stress scale. *Assessment*, *11*(3), 216–223. https://doi.org/10.1177/1073191104267398

Hayes, T., & Usami, S. (2020a). Factor Score Regression in Connected Measurement Models Containing Cross-Loadings. *Structural Equation Modeling*, *27*(6), 942–951. https://doi.org/10.1080/10705511.2020.1729160

Hayes, T., & Usami, S. (2020b). Factor Score Regression in the Presence of Correlated Unique Factors. In *Educational and Psychological Measurement* (Vol. 80, Issue 1). https://doi.org/10.1177/0013164419854492

Hewitt, P. L., Flett, G. L., & Mosher, S. W. (1992). The Perceived Stress Scale: Factor structure and relation to depression symptoms in a psychiatric sample. *Journal of Psychopathology and Behavioral Assessment*, *14*(3), 247–257. https://doi.org/10.1007/BF00962631

Hoshino, T., & Bentler., P. M. (2013). Bias in Factor Score Regression and a Simple Solution. In A. R. de Leon & K. C. Chough (Eds.), *Analysis of Mixed Data: Methodos & Applicatioons* (1st ed., pp. 43–61). Chapman & Hall. https://doi.org/https://doi.org/10.1201/b14571

Khalili, R., Sirati nir, M., Ebadi, A., Tavallai, A., & Habibi, M. (2017). Validity and reliability of the Cohen 10-item Perceived Stress Scale in patients with chronic headache: Persian version. *Asian Journal of Psychiatry*, *26*, 136–140. https://doi.org/10.1016/j.ajp.2017.01.010

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.

Kuhfeld, M., & Soland, J. (2022). Avoiding bias from sum scores in growth estimates: An examination of IRT-based approaches to scoring longitudinal survey responses. *Psychological Methods*, *27*(2), 234–260. https://doi.org/10.1037/met0000367

Ledgerwood, A., & Shrout, P. E. (2011). The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology*, *101*(6), 1174–1188. https://doi.org/10.1037/a0024776

Ng, S. M. (2013). Validation of the 10-item chinese perceived stress scale in elderly service workers: One-factor versus two-factor structure. *BMC Psychology*, *1*(1), 1–8. https://doi.org/10.1186/2050-7283-1-9

Örücü, M. Ç., & Demir, A. (2009). Psychometric evaluation of perceived stress scale for Turkish university students. *Stress and Health*, *25*(1), 103–109. https://doi.org/10.1002/smi.1218

Peterson, C., Maier, S. F., & Seligman, M. E. (1993). *Learned Helplessness: A Theory for the Age of Personal Control*. Oxford University Press.

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Roberti, J. W., Harrington, L. N., & Storch, E. A. (2006). Further Psychometric Support for the 10-Item Version of the Perceived Stress Scale. *Journal of College Counseling*, *9*(2), 135–147. https://doi.org/10.1002/j.2161-1882.2006.tb00100.x

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. R package version 0.5-12. *Journal of Statistical Software*, *48*(2), 1–36. https://econpapers.repec.org/article/jssjstsof/v_3a048_3ai02.htm%0Ahttp://www.jstatsoft.org/v48/i02/

Rosseel, Y., & Loh, W. W. (2022). A structural after measurement approach to structural equation modeling. *Psychological Methods*. https://doi.org/10.1037/met0000503

Skrondal, A., & Laake, P. (2001). Regression among factor scores. *Psychometrika*, *66*(4), 563–575. https://doi.org/10.1007/bf02296196

Spearman, C. (1987). The Proof and Measurement of Association between Two Things. *The American Journal of Psychology*, *100*(3), 441–471.

Steiger, J. H., & Schönemann, P. H. (1978). A history of factor indeterminacy. *Theory Construction and Data Analysis*, 136–178.

Thompson, G. H. (1934). The meaning of "i" in the estimate of "g." *British Journal of Psychology*, *25*(1), 92–99.

Thompson, G. H. (1946). *The factorial analysis of human ability* (2nd ed.). University of Lonodon Press.

Thurstone, L. L. (1935). *The vectors of mind: Multiple-factor analysis for the isolation of primary traits*. University of Chicago Press. https://doi.org/https://doi.org/10.1037/10018-000

Thurstone, L. L. (1947). *Multiple-Factor Analysis: A Development and Expansion of The Vectors of Mind*. University of Chicago Press.

van Bork, R., Rhemtulla, M., Sijtsma, K., & Borsboom, D. (2022). A Causal Theory of Error Scores. *Psychological Methods*. https://doi.org/10.1037/met0000521

Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLoS ONE*, *11*(3), 1–22. https://doi.org/10.1371/journal.pone.0152719

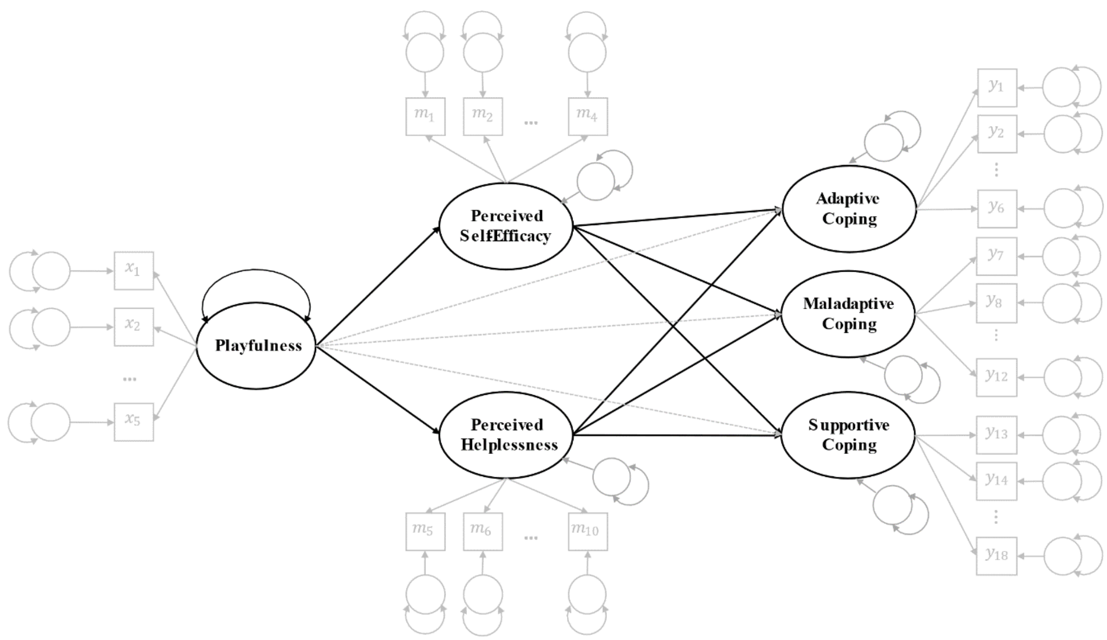
Wheaton, B., Muthén, B., Alwin, D. F., & Summers, G. F. (1977). Assessing Reliability and Stability in Panel Models. In *Sociological Methodology* (Vol. 8, pp. 84–136). Wiley.

Appendices

Appendix A

Figure 2

A Full Structural Equation Model of the Complete Model Used in Analysis for our Baseline Regression Model



Appendix B

Table 1

Reliability Measures of the Variables Used in Our Complete Model in Full and N = 150 Subsample

	SMAP	PSE	PH	ACOPE	MCOPE	SCOPE
N = 694	0.85	0.80	0.89	0.84	0.81	0.80
N = 150	0.83	0.79	0.90	0.83	0.81	0.85

Note: SMAP – Short Measure of Adult Playfulness; PSE – Perceived Self-Efficacy; PH – Perceived Helplessness; ACOPE – Adaptive Coping; MCOPE – Maladaptive Coping; SCOPE – Supportive Coping

Table 2 - Regression Estimates Results Table

	Standardized Estimates				Standard Error				*Semipartial Correlation*			
	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM
Outcome: Perceived Self-Efficacy (PSE)												
Play	0.286	0.255	0.245	0.285	0.041	0.035	0.035	0.041				
PSE R^2	0.082	0.065	0.06	0.081								
Outcome: Perceived Helplessness (PH)												
Play	-0.143	-0.116	-0.13	-0.143	0.041	0.037	0.037	0.041				
PH R^2	0.021	0.013	0.017	0.02								
Outcome: Adaptive Coping												
Play	0.233	0.299	0.237	0.23	0.04	0.032	0.034	0.041	0.05	0.083	0.053	0.049
Self-Efficacy	0.584	0.38	0.398	0.579	0.059	0.039	0.039	0.061	0.184	0.096	0.106	0.182
Helplessness	0.343	0.201	0.191	0.341	0.057	0.039	0.04	0.059	0.068	0.029	0.026	0.067
A-Coping R^2	0.308	0.235	0.202	0.302								
Outcome: Maladaptive Coping												
Play	0.085	0.114	0.066	0.084	0.037	0.031	0.032	0.036	0.007	0.012	0.004	0.007
Self-Efficacy	-0.138	-0.139	-0.139	-0.128	0.056	0.037	0.037	0.055	0.01	0.013	0.013	0.009
Helplessness	0.602	0.534	0.506	0.592	0.047	0.032	0.033	0.047	0.209	0.201	0.179	0.203
M-Coping R^2	0.475	0.376	0.343	0.453								
Outcome: Supportive Coping												
Play	0.165	0.195	0.168	0.165	0.043	0.036	0.037	0.04	0.025	0.036	0.026	0.025
Self-Efficacy	0.37	0.275	0.244	0.369	0.064	0.043	0.043	0.06	0.074	0.05	0.04	0.074
Helplessness	0.344	0.204	0.228	0.343	0.059	0.042	0.043	0.055	0.068	0.029	0.037	0.068
S-Coping R^2	0.136	0.112	0.089	0.136								

Note: SEM = structural equation modeling, Sum = path analysis using sum score composite variable, FSR = factor score regression, SAM = structural after measurement

Table 3 – Percent Bias (%) Comparison to SEM Regression Estimate Results

	Standardized Estimates				Standard Error				*Semipartial Correlation*			
	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM
Outcome: Perceived Self-Efficacy (PSE)												
Play	~	-10.9	-14.46	-0.44	~	-15.22	-14.65	-0.04				
PSE R^2	~	-20.61	-26.83	-0.87								
Outcome: Perceived Helplessness (PH)												
Play	~	-19.12	-9.41	-0.32	~	-8.99	-9.38	-0.03				
PH R^2	~	-34.58	-17.93	-0.64								
Outcome: Adaptive Coping												
Play	~	28.33	1.76	-1.03	~	-19.78	-16.5	1.98	~	68.14	6.37	-1.96
Self-Efficacy	~	-34.97	-31.75	-0.8	~	-34.59	-33.77	1.96	~	-47.81	-42.41	-1.39
Helplessness	~	-41.3	-44.18	-0.36	~	-31.85	-30.14	1.87	~	-57.94	-62.24	-0.57
A-Coping R^2	~	-23.69	-34.25	-1.79								
Outcome: Maladaptive Coping												
Play	~	34.06	-22.13	-0.73	~	-16.25	-13.95	-1.22	~	83.48	-37.71	-1.37
Self-Efficacy	~	0.42	0.08	-7.23	~	-34.95	-33.36	-1.23	~	24.46	23.83	-13.76
Helplessness	~	-11.32	-15.95	-1.6	~	-32.34	-29.52	-1.04	~	-3.99	-14.39	-3.03
M-Coping R^2	~	-20.98	-27.82	-4.81								
Outcome: Supportive Coping												
Play	~	17.85	1.3	-0.11	~	-15.3	-13.88	-6.8	~	41.79	5.42	-0.13
Self-Efficacy	~	-25.69	-33.92	-0.25	~	-33.65	-32.5	-6.81	~	-31.86	-46.02	-0.3
Helplessness	~	-40.53	-33.52	-0.14	~	-29.17	-28.26	-6.66	~	-56.83	-46.44	-0.14
S-Coping R^2	~	-17.41	-34.66	-0.34								

Table 4 - Regression Estimates Results Table in N=150 Subsample

	Standardized Estimates				Standard Error				*Semipartial Correlation*			
	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM
Outcome: Perceived Self-Efficacy (PSE)												
Play	0.267	0.255	0.245	0.275	0.09	0.035	0.035	0.089				
PSE R^2	0.071	0.065	0.06	0.075								
Outcome: Perceived Helplessness (PH)												
Play	-0.105	-0.116	-0.13	-0.104	0.089	0.037	0.037	0.089				
PH R^2	0.011	0.013	0.017	0.011								
Outcome: Adaptive Coping												
Play	0.193	0.299	0.237	0.185	0.083	0.032	0.034	0.086	0.034	0.083	0.053	0.031
Self-Efficacy	0.813	0.38	0.398	0.788	0.115	0.039	0.039	0.12	0.367	0.096	0.106	0.344
Helplessness	0.46	0.201	0.191	0.446	0.117	0.039	0.04	0.12	0.125	0.029	0.026	0.118
A-Coping R^2	0.5	0.235	0.202	0.471								
Outcome: Maladaptive Coping												
Play	0.08	0.114	0.066	0.078	0.074	0.031	0.032	0.072	0.006	0.012	0.004	0.006
Self-Efficacy	-0.136	-0.139	-0.139	-0.153	0.112	0.037	0.037	0.108	0.01	0.013	0.013	0.013
Helplessness	0.681	0.534	0.506	0.651	0.091	0.032	0.033	0.088	0.275	0.201	0.179	0.252
M-Coping R^2	0.588	0.376	0.343	0.563								
Outcome: Supportive Coping												
Play	0.154	0.195	0.168	0.149	0.088	0.036	0.037	0.082	0.022	0.036	0.026	0.02
Self-Efficacy	0.493	0.275	0.244	0.495	0.129	0.043	0.043	0.121	0.135	0.05	0.04	0.136
Helplessness	0.388	0.204	0.228	0.388	0.12	0.042	0.043	0.113	0.089	0.029	0.037	0.089
S-Coping R^2	0.202	0.112	0.089	0.203								

Note: SEM = structural equation modeling, Sum = path analysis using sum score composite variable, FSR = factor score regression, SAM = structural after measurement

Table 5 – Percent Bias (%) Comparison to SEM Regression Estimate Results in N = 150 Subsample

	Standardized Estimates				Standard Error				*Semipartial Correlation*			
	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM	SEM	Sum	FSR	SAM
Outcome: Perceived Self-Efficacy (PSE)												
Play	0.286	0.255	0.245	0.285	0.041	0.035	0.035	0.041				
PSE R^2	0.082	0.065	0.06	0.081								
Outcome: Perceived Helplessness (PH)												
Play	-0.143	-0.116	-0.13	-0.143	0.041	0.037	0.037	0.041				
PH R^2	0.021	0.013	0.017	0.02								
Outcome: Adaptive Coping												
Play	0.233	0.299	0.237	0.23	0.04	0.032	0.034	0.041	0.05	0.083	0.053	0.049
Self-Efficacy	0.584	0.38	0.398	0.579	0.059	0.039	0.039	0.061	0.184	0.096	0.106	0.182
Helplessness	0.343	0.201	0.191	0.341	0.057	0.039	0.04	0.059	0.068	0.029	0.026	0.067
A-Coping R^2	0.308	0.235	0.202	0.302								
Outcome: Maladaptive Coping												
Play	0.085	0.114	0.066	0.084	0.037	0.031	0.032	0.036	0.007	0.012	0.004	0.007
Self-Efficacy	-0.138	-0.139	-0.139	-0.128	0.056	0.037	0.037	0.055	0.01	0.013	0.013	0.009
Helplessness	0.602	0.534	0.506	0.592	0.047	0.032	0.033	0.047	0.209	0.201	0.179	0.203
M-Coping R^2	0.475	0.376	0.343	0.453								
Outcome: Supportive Coping												
Play	0.165	0.195	0.168	0.165	0.043	0.036	0.037	0.04	0.025	0.036	0.026	0.025
Self-Efficacy	0.37	0.275	0.244	0.369	0.064	0.043	0.043	0.06	0.074	0.05	0.04	0.074
Helplessness	0.344	0.204	0.228	0.343	0.059	0.042	0.043	0.055	0.068	0.029	0.037	0.068
S-Coping R^2	0.136	0.112	0.089	0.136								

CHAPTER IV

A Comparison of Regression Estimation Methods in Multi-Item Scales for Developmental Scientists: A Simulation Study

Christopher Clifford, Timothy Hayes, Stefany Coxe, and Leslie D. Frazier

Department of Psychology

Florida International University

We have no known conflicts of interest to disclose.

This research was not funded.

The data presented in this manuscript are available upon reasonable request.

Correspondence concerning this article should be addressed to Christopher Clifford, M.S., Department of Psychology, Florida International University, 11200 S. W. 8th Street, Miami, FL 33199. Email: cclif003@fiu.edu

Abstract

Comparing the efficacy that different methods of regression analyses have in a multi-item latent variable model is difficult in real data because the true population parameters can never be known due to the nature of sampling from population data. In this study statistical simulation methods were used to demonstrate how data characteristics (i.e., sample size, effect size, and reliability) may differentially impact the accuracy of regression approaches. Specifically, we simulated data with parameters based on paper I (Clifford et al., 2022) and then analyzed this model with four different approaches to regression estimation: Structural Equation Modeling (SEM), Structural After Measurement (SAM), Sum Scoring, and Factor Score Regression (FSR). The results showed that SEM was the best performing method for modeling regression estimating between latent factors from multi-item scales under typical data circumstances. We also showed SAM to be a comparable, and in some cases preferable, alterative under less ideal circumstances such as smaller sample size and lower reliability. Finally, our results also showed sum scoring and FSR regression estimation methods to have serious bias risks associated with their use and recommend avoid using them in favor of better performing and more thorough modeling methods.

keywords: simulation, regression modeling, structural equation modeling, structural after measurement, sum scoring, factor score regression, latent variables

An Analysis of Regression Estimation Methods in Simulated Multi-Item Scales for Developmental Scientists

In traditional research, inferences about a construct and its relationships to other constructs of interest are made via a sample taken from a larger population. In paper I, we did just this; examining how playfulness predicted coping in a three-variable latent mediation model (adaptive, maladaptive and supportive coping), exploring the role that stress (two factors, perceived self-efficacy and perceived helplessness) served as a mediator. By contrast, paper II employed a simplified model to compare estimates returned by several different regression methods. That paper showed that different methods for estimating regressions among constructs measured using multi-item scales often produced starkly different estimates. Specifically, the results of the 'gold standard' analysis, SEM, was used as a basis for evaluating the accuracy of the other approaches. Yet, because sample data are never a perfect recreation of the population of interest (due to factors such as the imperfect nature of measurement instruments, sampling error and the limitations of data collection), the extent to which the gold standard SEM analysis accurately reproduced the population parameters remains unknown. Therefore, for demonstrative purposes, in paper II the estimates produced by SEM were treated as if they were proxies for 'true' population values despite that these population values were, in fact, unknown.

Because the demonstrative analyses in paper II, strictly speaking, only allow us to compare regression estimates against each other (but not to the true, unknown population values), paper III uses simulation methods to compare the estimates of these methods to a set of population values that set within the simulation and, therefore, known. In a simulation, the exact population values can be specified and, as such, the statistical analysis methods can be compared to these values, not just to each other. Furthermore, by repeatedly sampling datasets in a simulation many times, simulations allow for the direct assessment of the methods' performance, on average, under repeated sampling.

Thus, the primary aim of paper III was to replicate the demonstration from paper II using simulation methodology that affords direct comparison to population values. We were particularly interested in how data characteristics (i.e., sample size, effect size, and reliability) may differentially impact the accuracy of these regression approaches. Specifically, the aims of paper III are to show that: (1) SEM and SAM would result in more accurate model estimates (compared to the true population values set in the simulation) than regressions based on sum scores or factor scores; (2) estimates from sum scoring and FSR analyses might exhibit both magnification and attenuation under different circumstances (Cole & Preacher, 2014); (3) the biases present in estimates based on analyses of FSR and sum scores are not eliminated with larger samples; (4) these biases will only become worse, however, as the reliability of each measure decreases, indicating a greater proportion of variance due to measurement error; (5) based on recent work by Rosseel and Loh (2022), it was anticipated that SAM should exhibit strong performance in smaller samples whereas SEM, using simultaneous estimation, may be prone to increased instances of model nonconvergence; and (6) when a population value is truly zero, the biases found in sum scoring and FSR may lead to increased false positive (Type I error) rates (Cole & Preacher, 2014; Ledgerwood & Shrout, 2011).

Methods

This paper compares the statistical estimates between four different regression estimation methods in a simple mediation model with three latent factors composed of multi-item indicators. To do this, a statistical simulation in R (R Core Team, 2022) was conducted. The core parameters of this simulation were based on the estimates from three different variables in the mediation triangle used in paper I: *x* (Playfulness), *m* (Self-Efficacy), and *y* (Adaptive Coping; See Figure 2). The population values for the ***a***, ***b***, and ***c'*** paths in the simulation were set to 0.286, 0.584, and 0.233 respectively (the standardized estimates for these constructs returned in paper 1). All variables (latent and observed) in the simulation were generated using standard normal distributions.

Simulation Design

Factors Varied the Simulation

In this simulation, the following factors were varied. First, to assess the performance of the four regression estimators in both small and large samples, the sample size was set either to $n = 150$ or $n = 500$. 2) Second, to assess the each method's performance under scenarios with relatively more or less measurement error, reliability, as specified by Cronbach's α (Cronbach, 1951) was set to either 0.5 or 0.8 for all variables using the method reported in Hayes and Usami (2020)[3]. Third, to demonstrate the potential for both magnified and attenuated regression coefficients (see Cole &

[3] Specifically, all (standardized) factors were held equal, with values set using the formula $\lambda = \sqrt{\left[\frac{\alpha}{(p-(p-1)*\alpha}\right]}$, where p was the number of indicators (5) and α was desired value of Cronbach's alpha.

Preacher, 2014), when some measures are characterized by more measurement error than others, a condition was run in which reliability was set to 0.8 for the latent x and y measurement models. Finally, to replicate findings from the methodological literature demonstrating increased Type 1 error rates when using sum scoring, the *c*’ path was held to 0 in a final set of conditions. In this set of conditions, we would expect significant results only 5% of the time for a well-behaved estimation method.

In sum, we ran a fully crossed 2 (N = 150 vs. 500) × 3 (α = .5 for all measures vs. .8 for all measures vs. .5 for *m* and .8 for *x* and *y*) × 2 (*c*’ = 0 vs. nonzero) factorial design resulting in 2*3*2 = 12 unique simulation cells. We generated 1000 simulated datasets per unique simulation cell (experimental condition) resulting in 12,000 datasets generated.

Analyses Performed on Simulated Data

For each dataset, four different regression estimation models were run via the `lavaan` package in `R` (Rosseel, 2012) and analyzed, similar to the four different regression estimation approaches in paper II (see Figure 1, paper II): SEM, SAM, sum scoring, and FSR. All analyses were run using default settings in `lavaan`. SAM analyses were conducted using connected estimation (see Rosseel & Loh, 2022).

Simulation Outcome

The simulation was designed to assess the following outcomes. First, to compare the average estimates from each regression method to the true population value in each condition, we captured (a) the average estimate returned be each method in each cell and

(b) percentage bias, defined here in relation to the true population values (rather than to the estimates from SEM as in paper II) using the following formula:

$$\% \text{ bias} = \frac{\left(\bar{\hat{\beta}}_j - \beta_j\right)}{\beta_j} \times 100$$

where β_j is the population value of the *j*th regression coefficient in the model (i.e., the *a*, *b*, or *c*' path in our latent mediation model) set in the simulation and where $\bar{\hat{\beta}}_j$ is the average estimate of β_j returned by a particular method (SEM, SAM, sum scoring, or FSR) in a given simulation condition.

Additionally, to further assess the methods' performance, particularly in smaller samples, (and particularly for SEM and SAM), we recorded the proportion of analyses (simulated datasets) that resulted in of estimation issues (convergence failures, improper solutions) produced in each condition. Finally, to assess Type I error rates in conditions where *c*' was set to 0, we recorded the proportion of analyses that resulted in significant *t*-tests of the coefficient.

Results

Bias in Sum Scoring and FSR

Tables 1 and 2 present average estimates and percent bias returned by each regression method, stratified by simulation condition at small sample sizes (*N* = 150, Table 1) and large sample sizes (*N* = 500, Table 2), respectively. For the first aim, SEM and SAM did result in more accurate regression estimates than sum scoring or factor scores. By contrast, sum scoring regression and factor score regression resulted in problematic levels of bias in the majority of simulation conditions (see bolded entries in Tables 1 and 2). As expected, this bias was more severe when reliability was lower (indicating a larger proportion of variance due to measurement error).

Alarmingly, as suggested by Cole and Preacher (2014), these problematic levels of bias were, at times, magnified rather than attenuated. Specifically, magnification can be seen in the c' path in the conditions in which reliability of the mediator (*m*) was set to 0.5 whereas the reliability of *x* & *y* were set to 0.8. This showcases what happens when measurement error is improperly handled in a multiple predictor model (Cole & Preacher, 2014).

In the third aim, it can then be seen that while SEM and SAM average estimates remained unbiased across sample sizes, the bias in sum score and factor score estimates did not improve with increased sample size. This is indicative of the extent to which the bias resulting from using these techniques is inherent in their inability to account for error variance, making a powerful case for not using sum score or factor score estimation in regression modeling.

Convergence and Improper Solutions in Small Samples ($N = 150$)

Table 3 displays the percentage of convergence failures/improper solutions in small samples ($N = 150$) by simulation condition and regression method. As seen in Table 3, convergence rate issues were low across all methods and conditions when there were issues at all. Nonetheless, SEM did result in higher rates of convergence failures and/or estimation issues (improper solutions, etc.) than the other methods. The low rates of convergence failures for sum scoring and FSR, while desirable, do not offset the severe biases described above and Type I error rates described below. By contrast, SAM performed equivalently to sum scoring and FSR, producing minimal convergence issues in the simulation. Taken together with this method's minimal bias and adequate Type I error control, these results are encouraging and replicate those reported by Rosseel and Loh (2022). Additionally, these results support the fifth aim of the paper, showing that SEM is more prone to convergence issues at smaller sample size and low reliability, but SAM does not suffer from these issues to the same extent.

Type I Error Rates in Sum Scoring and FSR

With respect to the last aim, for sum score and factor score regression estimation there was a notable increase in Type I error when the true population value of the *c*' path was set to zero (0). As shown in Table 4, Type I error was even more pronounced for sum score and factor score estimation at large sample sizes. Once again, this corroborates the prior literature suggesting that when methods ignore measurement error, estimates that are truly zero in the population may become biased away from zero, resulting in increased Type I error rates (Ledgerwood & Shrout, 2011). Examining Table 4, it is clear that both sum scoring and FSR resulting in Type I error rates that were substantially higher than the nominal value (5%) across all simulation conditions. By contrast, SEM and SAM estimates were at or below the nominal rate in all cases.

Discussion

In line with the methodological literature, this simulation demonstrated (1) that regression estimates from SEM and SAM regression are more accurate than those based on sum or factor scores; (2) that sum and factor score analyses exhibit both magnification and attenuation under different circumstances; (3) that large samples sizes do not eliminate the biases in sum score and factor score regression estimates–rather, the biases produced by these methods are an inherent feature of their failure to address measurement error; (4) that these biased estimates in sum and factor score estimation become worse with decreased reliability; (5) that SAM has strong performance in small samples and may be preferable to SEM in these instances as can be seen by SAM's increased convergence proportion; and finally (6) that Type I error rates of truly zero paths was problematically high in sum score and factor score regression analyses.

The takeaways of this simulation lead to the recommendation that developmental researchers avoid both sum scoring and FSR because these methods produce bias and inflated Type I error rates even in large samples. Instead, developmental researchers are encouraged to use statistical techniques, such as SEM (in larger samples) or SAM (in either larger or smaller samples), that appropriately account for measurement error. Additionally, given ongoing concerns about replication, the increased Type I error rates that result from sum scores and factor scores are especially noteworthy. For those unable to use SEM in their research due to issues regarding sample size, the results from this simulation corroborate those from Rosseel and Loh (2022) and show that SAM may be an effective compromise. This method accounts for measurement error in the same manner as SEM but lessens sample size needs by estimating the model piece-by-piece in a multi-step process. In sum, this paper shows the importance of using estimation techniques that account for measurement error in multi-item scales and clearly demonstrates the risks of ignoring measurement error. For these reasons, these findings are important for developmental scientists.

References

Clifford, C., Paulk, E., Lin, Q., Cadwallader, J., Lubbers, K., & Frazier, L. D. (2022). Relationships among adult playfulness, stress, and coping during the COVID-19 pandemic. *Current Psychology*, *0123456789*. https://doi.org/10.1007/s12144-022-02870-0

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–315. https://doi.org/10.1037/a0033805

Cronbach, J. L. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/https://doi.org/10.1007/BF02310555

Hayes, T., & Usami, S. (2020). Factor Score Regression in the Presence of Correlated Unique Factors. In *Educational and Psychological Measurement* (Vol. 80, Issue 1). https://doi.org/10.1177/0013164419854492

Ledgerwood, A., & Shrout, P. E. (2011). The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology*, *101*(6), 1174–1188. https://doi.org/10.1037/a0024776

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. R package version 0.5-12. *Journal of Statistical Software*, *48*(2), 1–36. https://econpapers.repec.org/article/jssjstsof/v_3a048_3ai02.htm%0Ahttp://www.jstatsoft.org/v48/i02/

Rosseel, Y., & Loh, W. W. (2022). A structural after measurement approach to structural equation modeling. *Psychological Methods*. https://doi.org/10.1037/met0000503

Appendix

Appendix A

Figure 1. A theoretical model showing the relationship between Playfulness, Self-Efficacy and Coping

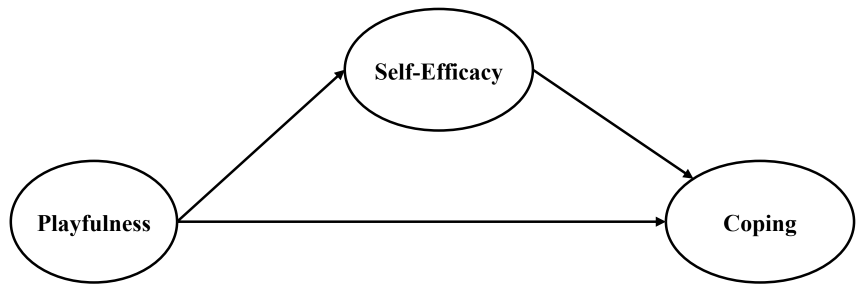


Figure 2. A full theoretical model showing the basis of the relation between our three latent variables as a result of the multi-item indicators

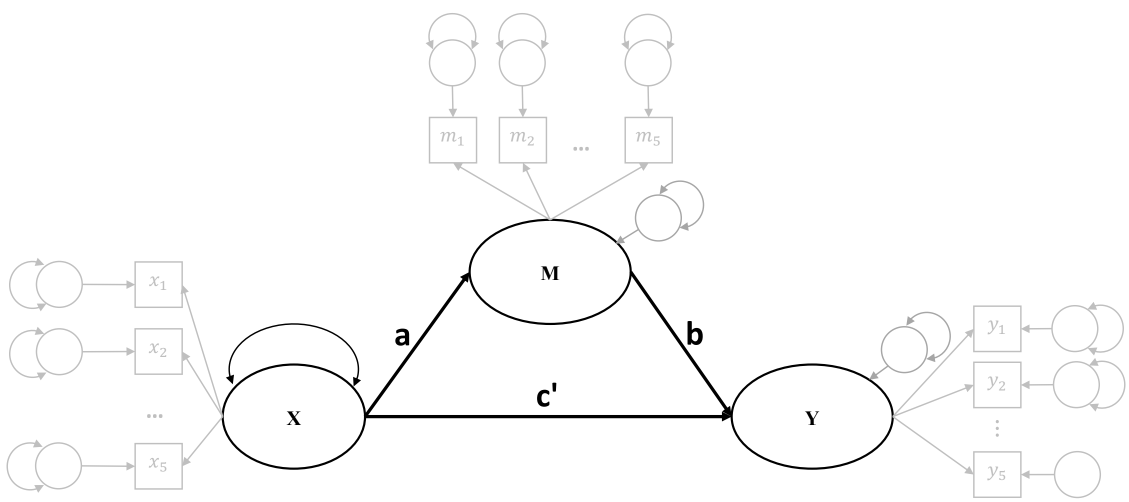


Figure 2 showcases the model upon which the simulation data is based. Each latent variable and the relationship to the other two variables are based on results of the parameters calculated in Paper I between Playfulness (X), Perceived Self-Efficacy (M) and Adaptive Coping (Y) respectively.

Appendix B

Reliability	X To Y Path	Path	True Path Value	β Estimate				Percent Bias (%) in β Estimate			
				SEM	SAM	Sum Scoring	FSR	SEM	SAM	Sum Scoring	FSR
0.5	0	a	0.286	0.284	0.282	0.142	0.132	-0.7	-1.4	**-50.3**	**-53.8**
0.5	0	b	0.584	0.593	0.599	0.287	0.264	1.5	2.6	**-50.9**	**-54.8**
0.5	0	c'	0	0.002	-0.002	0.046	0.045	n/a	n/a	n/a	n/a
0.5	regular	a	0.286	0.279	0.278	0.139	0.127	-2.4	-2.8	**-51.4**	**-55.6**
0.5	regular	b	0.584	0.589	0.599	0.305	0.282	0.9	2.6	**-47.8**	**-51.7**
0.5	regular	c'	0.233	0.228	0.228	0.154	0.145	-2.1	-2.1	**-33.9**	**-37.8**
0.5 M, X & Y 0.8	0	a	0.286	0.288	0.288	0.182	0.174	0.7	0.7	**-36.4**	**-39.2**
0.5 M, X & Y 0.8	0	b	0.584	0.591	0.59	0.356	0.338	1.2	1	**-39**	**-42.1**
0.5 M, X & Y 0.8	0	c'	0	-0.005	-0.005	0.070	0.075	n/a	n/a	n/a	n/a
0.5 M, X & Y 0.8	regular	a	0.286	0.286	0.287	0.182	0.174	0	0.3	**-36.4**	**-39.2**
0.5 M, X & Y 0.8	regular	b	0.584	0.585	0.584	0.361	0.345	0.2	0	**-38.2**	**-40.9**
0.5 M, X & Y 0.8	regular	c'	0.233	0.236	0.234	0.258	0.260	1.3	0.4	**10.7**	**11.6**
0.8	0	a	0.286	0.284	0.284	0.228	0.226	-0.7	-0.7	**-20.3**	**-21**
0.8	0	b	0.584	0.586	0.586	0.461	0.458	0.3	0.3	**-21.1**	**-21.6**
0.8	0	c'	0	-0.001	-0.001	0.028	0.029	n/a	n/a	n/a	n/a
0.8	regular	a	0.286	0.29	0.29	0.233	0.231	1.4	1.4	**-18.5**	**-19.2**
0.8	regular	b	0.584	0.585	0.585	0.47	0.467	0.2	0.2	**-19.5**	**-20**
0.8	regular	c'	0.233	0.229	0.229	0.209	0.209	-1.7	-1.7	**-10.3**	**-10.3**

Table 1 - *n* = 150 Sample Size Simulation Results

Reliability	X To Y Path	Path	True Path Value	β Estimate				Percent Bias (%) in β Estimate			
				SEM	SAM	Sum Scoring	FSR	SEM	SAM	Sum Scoring	FSR
0.5	0	a	0.286	0.283	0.283	0.141	0.139	-1	-1	**-50.7**	**-51.4**
0.5	0	b	0.584	0.584	0.584	0.285	0.280	0	0	**-51.2**	**-52.1**
0.5	0	c'	0	0.001	0.001	0.043	0.043	n/a	n/a	n/a	n/a
0.5	regular	a	0.286	0.287	0.287	0.144	0.141	0.3	0.3	**-49.7**	**-50.7**
0.5	regular	b	0.584	0.586	0.586	0.302	0.296	0.3	0.3	**-48.3**	**-49.3**
0.5	regular	c'	0.233	0.234	0.234	0.157	0.155	0.4	0.4	**-32.6**	**-33.5**
0.5 M, X & Y 0.8	0	a	0.286	0.282	0.282	0.178	0.176	-1.4	-1.4	**-37.8**	**-38.5**
0.5 M, X & Y 0.8	0	b	0.584	0.585	0.584	0.355	0.352	0.2	0	**-39.2**	**-39.7**
0.5 M, X & Y 0.8	0	c'	0	0	0	0.069	0.070	n/a	n/a	n/a	n/a
0.5 M, X & Y 0.8	regular	a	0.286	0.284	0.284	0.179	0.177	-0.7	-0.7	**-37.4**	**-38.1**
0.5 M, X & Y 0.8	regular	b	0.584	0.586	0.586	0.365	0.361	0.3	0.3	**-37.5**	**-38.2**
0.5 M, X & Y 0.8	regular	c'	0.233	0.232	0.232	0.254	0.254	-0.4	-0.4	9	9
0.8	0	a	0.286	0.287	0.287	0.229	0.229	0.3	0.3	**-19.9**	**-19.9**
0.8	0	b	0.584	0.583	0.583	0.459	0.458	-0.2	-0.2	**-21.4**	**-21.6**
0.8	0	c'	0	0.001	0.001	0.029	0.029	n/a	n/a	n/a	n/a
0.8	regular	a	0.286	0.288	0.288	0.231	0.230	0.7	0.7	**-19.2**	**-19.6**
0.8	regular	b	0.584	0.583	0.583	0.471	0.469	-0.2	-0.2	**-19.3**	**-19.7**
0.8	regular	c'	0.233	0.235	0.235	0.213	0.213	0.9	0.9	-8.6	-8.6

Table 2 - n = 500

Sample Size Simulation Results

Table 3 - Model Convergence Issues by Condition

Sample Size	Reliability	X To Y Path	% of Models Run with Convergence Issues			
			SEM	SAM	Sum Scoring	FSR
150	0.5	0	3.2	0.4	0.4	0.4
150	0.5	regular	3.8	0.3	0.3	0.3
150	0.5 M, X & Y 0.8	0	0.2	0.1	0.1	0.1
150	0.5 M, X & Y 0.8	regular	0.5	0.1	0.1	0.1

Table 4 - Type I Error Rate

Sample Size	Reliability	Type I Error Rate (%)			
		SEM	SAM	Sum Scoring	FSR
150	0.5	1.1	1.1	10.9	9.6
150	0.5 M, X & Y 0.8	4.9	4.8	15	16.8
150	0.8	4.5	4.5	7.5	7.2
500	0.5	3.7	3.7	18.2	18.4
500	0.5 M, X & Y 0.8	5.0	5.0	36.9	38
500	0.8	4.7	4.7	11.9	12.1

CHAPTER V

General Conclusions

This dissertation presents three papers–one empirical study, one pedagogical demonstration, and one statistical simulation study–focusing on how regression modeling with latent variables for multi-item measures should be applied within developmental science research. The first study applied a latent variable modeling framework to personality and behavioral data examining the relationships between playfulness, stress, and coping. The second study used the same data to demonstrate and compare four different approaches to estimating regression relationships among constructs measured using multi-item scales. In the third paper, statistical simulation methods were used to demonstrate how data characteristics (i.e., sample size, effect size, and reliability) differentially impacted the accuracy of these same approaches. Importantly, in contrast to the demonstration in paper II, this simulation approach allowed comparison of each method's performance to a set of known (true) population values set in the simulation.

The first study produced a paper (Clifford et al., 2022) that examined the relationships between stress, adult playfulness, and coping during the COVID-19 pandemic using a parallel mediation model utilizing latent variables to represent each construct estimated with structural equation modeling. The primary aim of this paper was to investigate how adult playfulness influenced the stress-coping process. Developmentally, play research has historically focused on children; yet for adults playfulness may be a personality resource that may lead to better mental and physical health outcomes (Clifford et al., 2022). We tested a theoretical model to determine whether the two factors of perceived stress, perceived self-efficacy (PSE) and perceived helplessness (PH), mediate the relations between playfulness and coping in adults. Scores on the Perceived Stress Scale (PSS; Cohen et al., 1983) were high, likely indicating high levels of pandemic-related stress at the time of data collection. Results from the SEM model demonstrated direct effects of playfulness on PSE, PH, and adaptive, maladaptive, and supportive coping. Both dimensions of perceived stress were partial mediators of the relations among playfulness and coping outcomes. These findings illustrate the pathways through which adult playfulness can amplify or attenuate the impact of stress perceptions on coping strategies and reinforce the practical application of applying regression modeling to complex relations involving latent variables.

The second paper compared different modeling techniques that might be used to perform a regression analysis with latent variables measured by multi-item scales. This paper looked at four different approaches to regression estimation with multi-item scales using: (1) sum scores; (2) latent factors in a structural equation model (SEM); (3) factor score regression; and (4) a novel two-stage estimation procedure recently introduced in the methodological literature called Structural After Measurement (SAM, Rosseel & Loh, 2022). Our results conformed with expected results and previous literature, showing the stark divergence of estimates resulting from sum score and factor score regression and those resulting from "gold standard" SEM. Further, this paper showcased SAM as a viable alternative compared to SEM.

The final paper took the comparative analyses from paper II one step further by simulating data based on the mediation model from paper I and exploring how each of the four estimation techniques in paper II performed when scrutinized against a set of known population values. I successfully accomplished all six aims for this paper, showing that: (1) SEM and SAM regression estimates are more accurate than those of sum scores or factor scores; (2) sum and factor score analyses exhibit magnification and attenuation under different circumstances; (3) large samples sizes do not eliminate the biases in sum score and factor score regression estimates; (4) these biased estimates in sum and factor score estimation become worse with decreased reliability (increased measurement error); (5) SAM performs well in small samples and may be preferable to SEM in these instances, as can be seen by SAM's lower rates of convergence failures; and, finally, (6) Type I error rates were alarmingly high in sum score and factor score regression estimation methods when a true population pathway was equal to zero. These results suggest that SEM and SAM should be the preferred approaches for developmental scientists who are estimating regression relations between multi-item scales.

These three papers successfully accomplished the overarching goals of this dissertation. Specifically, I successfully examined and compared a variety of quantitative methods for estimating regressions among latent variables measured by multi-item scales that might be considered by developmental scientists. Whereas SEM remains the "gold standard" approach for researchers able to collect large samples (as in paper I), SAM represents a promising alternative approach that addresses the constraints sometimes faced by developmental scientists studying difficult to reach or hard to collect populations. Furthermore, each of the three papers in this dissertation highlights a different academic skillset. Paper I focuses on applying state of the art data analysis methods to empirical data. Paper II focuses on using a real dataset as the basis for a pedagogical demonstration and comparison of four methods. Finally, paper III focuses on the use of statistical simulation methods to compare the performance of these methods against a known population model. In these ways, this dissertation achieves each of the goals it set out to achieve and contributes to the existing body of methodological guidance for developmental scientists on best practices for running regressions using latent variables.

References

Clifford, C., Paulk, E., Lin, Q., Cadwallader, J., Lubbers, K., & Frazier, L. D. (2022). Relationships among adult playfulness, stress, and coping during the COVID-19 pandemic. *Current Psychology*, *0123456789*. https://doi.org/10.1007/s12144-022-02870-0

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of health and social behavior*, 385-396

Rosseel, Y., & Loh, W. W. (2023). A structural after measurement approach to structural equation modeling. *Psychological Methods*.

VITA

CHRISTOPHER CLIFFORD

2012-2015	B.A., Psychology University of New Mexico Albuquerque, New Mexico
2016-2020	M.S., Psychology Florida International University Miami, Florida
2020-2023	Ph.D., Psychology Florida International University Miami, Florida

PUBLICATIONS AND PRESENTATIONS

Nazarimehrvarani, Z., Lee, Y. T., Clifford, C., & Frazier, L. D. (in preparation). Emotion regulation, emotional eating, and disordered eating in college women.

Paulk, E., Bazo Perez, M., Clifford, C., Fava, N. M., & Frazier, L. D. (in preparation) Risk and Resilience for Eating Disorders in Military Personnel: Linkages Among Sexual Trauma, Muscle Dysmorphia and Risk for Disordered Eating.

Lubbers, K. G., Cadwallader, J. L., Clifford, C., Lin, Q., & Frazier, L. D. (Revise & Resubmit). Adult play and playfulness: A qualitative exploration of its meaning and importance. *Journal of Play in Adulthood.*

Clifford, C., Paulk, E., Lin, Q., Cadwallader, J., Lubbers, K. G., & Frazier, L. D. How does adult playfulness impact stress and coping during the COVID-19 pandemic? *Journal of Happiness Studies*. https://doi.org/10.1007/s12144-022-02870-0

Yakin, I., Clifford, C., Frazier, L. D., Meral, Y. & Çarboğa, G. *COVID-19 Concerns and Generalized Anxiety Disorder: The Role of Mindfulness, Cognitive Flexibility, and Cognitive Control.* Paper presented to the International Convention of Psychological Science (ICPS). March 9 – 11, 2023, Brussels Belgium.

Nazarimehrvarani, Z., Clifford, C., & Frazier, L. D. *Emotion Regulation, Orthorexia Nervosa, and Their Role in Disordered Eating*. To be presented at the Association for Psychology Science (APS) Annual Convention, May 25-28, 2023, Washington, D.C.

Clifford, C., Paulk, E., Lin, Q., Cadwallader, J., Lubbers, K. & Frazier, L. D. *How does adult playfulness impact stress and coping during the COVID-19 pandemic?* Paper submitted for consideration at the 113th Annual Meetings of the Southern Society for Philosophy and Psychology, March 31-April 2, 2022, Mobile, Alabama.

Frazier, L. D., & Clifford, C. (2022) *Playfulness and eudaimonia: The importance of play for psychological well-being.* Paper submitted for consideration at the 113th Annual Meetings of the Southern Society for Philosophy and Psychology, March 31-April 2, 2022, Mobile, Alabama.

Clifford, C., Paulk, E., Lin, Q., Wang, T., & Frazier, L. D. *The coactive influence of playfulness on stress and coping.* Paper presented at the 7th Biennial Meeting of the Association for Research in Personality, online, July 9, 16, & 23, 2021.