Theory & Psychology

Article

Theory & Psychology
2018, Vol. 28(6) 712–736

DOI: 10.1177/0959354318804670
journals.sagepub.com/home/tap
SAGE

Statistical positivism versus critical scientific realism. A comparison of two paradigms for motivation research: Part 1. A philosophical and empirical analysis of statistical positivism

Valery Chirkov
University of Saskatchewan, Canada

Jade Anderson
University of Saskatchewan, Canada

Abstract
In this two-part publication, we compare two paradigms—statistical positivism and critical scientific realism—in their application to research on academic motivation. In the first part, the propositions of statistical positivism and their applications to psychological research are presented. An empirical study in this part combines self-determination and achievement goal theories and builds a statistically integrated model of motivation of 385 college students using path analysis. This part ends with a critical analysis of this statistical model and the knowledge about motivation that it provides. In the second part, the propositions of critical scientific realism are articulated. An empirical study in Part 2 utilizes these propositions and initiates realist interviewing of 12 purposefully selected students. Using within- and between-case analyses, a model of a motivational mechanism of successful university students is proposed. The authors conclude that the continued use of statistical positivism generates minimal new knowledge about the mechanisms of academic motivation. This paradigm should be replaced with the realist one and a case-based methodology, which have a better chance to advance research and improve understanding of academic motivation.

Corresponding author:
Valery Chirkov, Department of Psychology, University of Saskatchewan, 9 Campus Dr. Saskatoon, S7N 5A5, Canada.
Email: v.chirkov@usask.ca

Keywords
academic motivation, case-based approach, critical scientific realism, statistical positivism, variable-based approach

The psychology of motivation searches for answers to the question, "Why do people act, experience, feel, and think the way they do?" Motivation researchers have taken diverse philosophical and methodological paths to answer this question. For instance, Freud used clinical interviews of patients with different forms of psychopathology to hypothesize the intrapsychic machinery of both maladjusted and relatively healthy minds. Behaviorists took a different approach. To discover general laws of human behavior, behaviorists went several steps back in the evolutionary ladder and focused on studying different species of animals in structured laboratory settings. By controlling and manipulating various external and internal stimuli, they meticulously recorded changes in organisms' behaviors to establish conjunctions between environmental conditions (schedules of reinforcement) and behavior outcomes.

The 1950s through 1970s was a period when intensive laboratory experiments were performed to examine the motivated behavior of humans. At this time, psychologists applied cognitive-behaviorist paradigms directly to humans. Under controlled conditions in psychological laboratories, social and motivational psychologists exposed their participants, mostly college students, to diverse situations and conditions to ascertain the general regularities of human social behavior (Danziger, 1992, 2000; Patnoe, 1988; Rodrigues & Levine, 1999). During this period, numerous motivational constructs, such as *goal, expectancy, valence, cognitive dissonance*, and *causal attribution* entered the vocabulary of psychologists. Many scholars consider these years to be the glory days of social and motivation psychology because numerous social-psychological and motivational theories constructed at that time still influence current research (Gilbert, Fiske, & Lindzey, 1996). Many others believe that this excessive experimentalism prevented researchers from understanding the complexity of the motivational mechanisms of human actions in real sociocultural environments and that it led to crises in social, personality, and motivation psychology (Elms, 1975; Pancer, 1997; Rosnow, 1981).

Behind the development of laboratory experimentation in social and motivation psychology, another trend of investigating human psychology has emerged. Namely, it is an approach of studying social behavior, personality, and motivation though self-report questionnaires analyzed with inferential and multivariate statistics with the aim of establishing reliable associations among the scores of these questionnaires (Cronbach, 1957; Danziger, 1985). With the advancement of various statistical techniques, this approach to the study of human psychology has become increasingly popular in many domains of modern psychology, and it is even frequently conflated with a quantitative methodology in psychological and social sciences (Halfpenny, 1982/2015; Toomela, 2010). Quantification in psychology has a long history (Michell, 2003; Stam, 2006) while quantitative studies in social and psychological sciences emerged as a distinct domain of these sciences near the end of the 19th century (Danziger, 1990; Halfpenny, 1982/2015). In modern times, despite the emergence of qualitative paradigms and methodologies, many researchers agree that the statistical approach, together with the revival of the neo-Galtonian model of experimentation (Danziger, 1990), has become the dominant paradigm associated with the scientific approach in psychology (Danziger, 1987; Polkinghorne, 1983).

The commitment of many psychologists to statistical methodology is rooted in the conviction that psychology can become a mature science only by emulating the natural sciences in their philosophy and methodology of research; specifically, by using enumeration and quantification of data followed by statistical analysis. Although most of the natural sciences have moved away from such a way of thinking about research, many psychologists still believe the myth that the quantification, measurement, and use of statistics bring real scientific vigor and objectivity to psychology (Bickhard, 1992, 2001; Rorer, 1991). Although these manipulations may be useful in scientific research, they do not define it, as real science uses diverse methodologies and techniques.

Despite this diversity of approaches, many essential questions pertaining to personality, social, and motivation research continue to trouble many scholars: Where is the individual in psychological research (Carlson, 1971; Valsiner, 1986a)? Why do psychologists rarely use case analysis (Barlow & Nock, 2009)? Regardless of psychologists' enormous efforts to be scientific, why do many people perceive the study of human behavior as unscientific (Ferguson, 2015; Lilienfeld, 2012)? In addition, they have more specific questions about motivation. What do motivation researchers ultimately try to discover: the reliable regularities among measured variables or the unobservable structures and mechanisms that drive and guide human actions? What are the best ways to make these discoveries: studying hundreds of participants in large samples by using statistical methods or investigating a small number of individual cases with deep and contextualized qualitative analyses? These concerns come to the ultimate question: Do current motivation studies move us to a more elaborate understanding of why people experience, feel, and think the way they do?

Many researchers will agree that motivation research is moving in the right direction; numerous motivational theories constitute significant contributions to modern psychology (e.g., the attribution theory of motivation and emotion, achievement motivation and achievement goal theories in their numerous forms, self-determination theory, self-efficacy theory, and various forms of social-cognitive theories) and guide successful applications. However, some inquisitive scholars remain unsatisfied, and they consider these to be mid-level theories that only present pieces of the puzzle of human psychological functioning; a puzzle that is far from being solved. These inquisitive scholars may further claim that the way modern empirical motivation research is progressing leaves fewer chances for making novel insights into the enigma of human psychology (Baumeister, Vohs, & Funder, 2007; Funder, 2009). These researchers are concerned with the shallow characters of quantitative/statistical studies and their inability to discover new facts, substantial empirical regularities, and the causal mechanisms of human motivation.

The authors of this article are among these concerned scholars. The goal of this article is to provide an analysis of the mainstream quantitative approach to studying motivation, to critically reflect on it, and, in Part 2, to present possible alternatives for thinking, theorizing, and conducting motivation research (Chirkov & Anderson, 2018). The authors' thesis is that the statistical approach, which has been labeled *statistical positivism* (Gigerenzer, 1987) or *statistism* (Lamiell, 2013), has exhausted its capacity, both philosophically and methodologically to generate discoveries, and it should thus be replaced with a realist thinking in the form of explanations through psychological mechanisms.

In the following, we will outline and justify several outstanding features of statistical positivism relating to empirical research. To demonstrate these features, we will report on a study of academic motivation conducted using achievement goals and self-determination theories. These theories were statistically combined to develop an integrated model of academic motivation. Critical reflections on this model will demonstrate the limitations and pitfalls of statistical positivism. We argue that the knowledge produced in this study has limited value and cannot be used to understand the motivational dynamics of human actions.

Positivism and statistical positivism in psychology

We will not provide a historical account of the different forms of positivism that are exhaustively analyzed elsewhere (Halfpenny, 2001, 1982/2015; Polkinghorne, 1983; Ray, 2000; Salmon, 2000; Smith, 1986; Tolman, 1992). Rather, our purpose is to highlight that statistical positivism entered the scientific scene with Francis Galton's invention of "anthropometric measurements" (Danziger, 1990); this was followed by biometrician and philosopher of science Karl Pearson, who advocated for the primary role statistics should play in empiricist scientific inquiries (Pearson, 1892/1957). Pearson's development of statistical techniques was followed by Sir Ronald Fisher, Jerzy Neyman, and Egon Pearson elaborating on experimental and inferential statistics (Gigerenzer, 1993; Halfpenny, 1982/2015; Halpin & Stam, 2006; John, 1992). The conquest of modern psychological research by statistical positivism occurred near the end of the 19th century and in the first third of the 20th century (Danziger, 1990). This conquest also penetrated other social sciences, where it was called "probabilistic," "inferential," or a "quantitative revolution" (Danziger, 1985, 1987; Kruger, Gigerenzer, & Morgan, 1987). Ultimately, the positivist hypothetico-deductive method, accompanied by the null-hypothesis significance testing, established itself as the standard scientific method in psychology (Rorer, 1991). Below, we summarize the propositions of positivism and its statistical version and provide comments on their applications in motivational psychology. This summary is based on multiple works (Carnap, 1966/1995; Halfpenny, 1982/2015; Hempel, 1942, 1966; Mandler & Kessen, 1964; Smith, 1986), which present the propositions of positivism, and on publications that provide critical reflections on this paradigm (Chirkov, 2016; Danziger, 1985, 1990; Hammersley, 2012; Manicas & Secord, 1983; Polkinghorne, 1983; Rorer, 1991; Znaniecki, 1934).

Main propositions of statistical positivism in psychology

The goal of science in general and of psychology, in particular, is to find stable empirical regularities among variables, events, and situational parameters that hold nearly universally across different populations. When empirically verified, these regularities acquire the status of scientific laws; when verbally arranged, they hold the status of scientific theories that allow researchers to explain and predict human behavior and experience.

Based on these regularities or laws, researchers design intervention programs to control and/or change people's behaviors. The following application of statistics transforms the positivist approach into the *statistical positivist paradigm* (Gigerenzer, 1987).

Positivist research has to be based exclusively on the empirical data that are generated by structured observations and/or by what is even more trustworthy, reliable measurements. For positivist researchers, *real* is what is empirically verifiable and measurable (Bhaskar, 1975/2008). They believe that all constructs that researchers study should be empirically provable, directly assessable, and represented by meaningful symbols, typically numbers. Theorizing and explanation should operate exclusively within these variables, and no entities can be proposed beyond these empirical facts and their statistical associations. References to unobservable and immeasurable factors and conditions should be avoided as they are seen as a deviation from the principles of the empiricist scientific inquiry. Such references are often called "pure/armchair/metaphysical speculations" (Mandler & Kessen, 1964). If a researcher proposes theoretical variables to explain empirical regularities, he or she must also operationalize them, develop means to measure them, and then test whether these measures relate to other constructs in the predicted directions. In addition, relationships among constructs should be as empirically explicit as possible with an opportunity for direct replication and quantification. Statistical programs' outputs provide good instances of such empirical verifiability of relations among motivational variables. Thus, the primary features of statistical positivism are *extreme empiricism*, *empirical/naïve/shallow realism*, and *strong anti-metaphysical attitudes*.

A direct consequence of extreme empiricism is a radical *operationalism* that requires researchers to define psychological constructs through the operations of their empirical instantiations (Chang, 2009; Koch, 1992; Langfeld, 1945; Petrie, 1971). Only through their operationalizations, hypothetical constructs are brought to life and can become "real" objects for positivist scientific research. The success of empirical motivational science depends on the availability of the operationalizations of different constructs in the form of scales, tests, and questionnaires. If there are no measurable operationalizations then no scientific research is possible.

The best way to obtain empirical data in positivist psychological research is to represent the objects of inquiries—the psychological hypothetical constructs—as a set of *variables*, like traits, needs, motives, cognitive attributes, emotions, and elements of socio-cultural environments. These variables must be observable and measurable. Thus, a *variable-based approach* is a crucial way of thinking about and structuring research in statistical positivism (Blumer, 1956; Danziger & Dzinas, 1997; Toomela, 2008). These variables are treated as components of the human psychological/motivational machinery that vary across individuals and can be subjected to multivariate statistical analysis. Thus, motivational mechanisms are represented by formal connections among their components—variables—through empirically verifiable statistical associations.

Statistical positivists search for stable empirical regularities through between-individual covariances among variables that are calculated on samples of participants. A multivariate statistical analysis establishes these regularities.[1] Behind this analysis, there is an implicit (and unjustified) assumption that the inter-individual covariances and, based on them, statistical associations are accurate representations of the relations and regularities that exist on the intra-individual level where actual psychological functioning unfolds

(Molenaar, 2004). This assumption is also based on a belief that if researchers are to investigate the motivational powers and forces that drive human behavior, they should look for them in the abstract associations of the operationalized constructs and interpret this "statistical play with symbols" (Znaniecki, 1934, p. 231) as a legitimate reflection of the dynamics of the real motivational forces inside individuals.

Sampling is an important step in any statistical investigation of motivation; however, it is also another controversial aspect of this research. Statistical positivists believe that sampling allows researchers to: (a) generalize their results to a larger population; (b) balance out individual differences and outliers in order to discover the most stable and universal trends in statistical regularities, and after being generalized to a larger population, these may become scientific laws and sources for discoveries and scientific theories; and (c) generate more reliable and, ultimately, more valid data compared to other methods of data generation.

When statistical associations are discovered in one sample and replicated in other samples, these associations may be generalized to the non-observed instances or cases of a whole population, and the status of these associations as scientific laws can be established. In the philosophy of science, such generalizations of empirical regularities from observed to non-observed instances are known as *enumerative generalization* (Chirkov, 2016). Positivists consider this form of generalization to be the inductive inference that should be pursued by empirical researchers.

As soon as statistical associations among psychological constructs/variables acquire the status of scientific laws by being replicated on numerous, diverse, and relatively large samples, they may be considered *causal relations* that constitute psychological *causal mechanisms*. This move from statistical covariances to causal relations is based on Hume's interpretation of causality: "when we've had many experiences of one kind of event constantly conjoined with another, we begin to think of them as cause and effect and infer the one from the other" (Morris & Brown, 2016, para. 12). Hume's interpretation of causality through conjoined events was followed by Mill's (1843/1965) application of such an understanding to psychology. This line of philosophical arguments led to Pearson's conviction that, "it is this conception of correlation between two occurrences embracing all relationships from absolute independence to complete dependence, which is the wider category by which we have to replace the old idea of causation" (1892/1957, p. 157). Based on these interpretations, statistical positivists conflate causality with conjunctions (covariance) of events: If B always follows A, and when A is absent B is also absent, then A is a cause of B. Statistical methodologists (Mulaik, 1987; Spadish, Cook, & Campbell, 2001) propagated the same understanding of causality when they announced that statistical models built through Structural Equation Modeling and/or path analysis are causal models.

According to positivism, an application of scientific knowledge happens when one employs scientific laws to explain particular instances of a phenomenon. Thus, the *deductive-nomological model of inference* (Hempel & Oppenheim, 1948) becomes the primary form of the positivist scientific explanation. For example, if motivation researchers were to discover a stable regularity that indicates that mastery achievement goals are positively associated with higher grades in high school students, they may deduce from this "scientific law" that if students in a high school have low grades, this happens because they do not strive to achieve mastery goals. In the positivist paradigm, testing this hypothesis and determining the level of students' mastery goals in that new sample constitutes the main logic of the scientific theory's verification and accumulation of knowledge. This is called the *hypothetico-deductive method* of theory verification. It is important to note that in all this hypothesizing, deducting, and verifying, no new knowledge is generated over and above the initial regularity of a positive association of mastery goals and high grades. Thus, positivists' enumerative generalization, a deductive-nomological model of explanation, and hypothetico-deductive mode of inference preclude these researchers from making scientific discoveries. This is one of the reasons that this paradigm has been rejected by natural scientists (Rorer, 1991).

Positivist scientists are "research workers" (Fisher, 1970; Mandler & Kessen, 1964) who meticulously collect empirical regularities, i.e., statistical associations among variables, and add them to the body of scientific knowledge through relentless publications. New scientific discoveries are made by a selected few who, by the power of their creativity, generate new insights into old regularities; in this manner, discoveries just happen, and they cannot be taught and managed. Thus, the creative and thrilling enterprise of scientific investigation, known as the *context of discovery* (Reichenbach, 1958), is replaced by a tedious algorithmic process of applying statistical programs to operationalized variables—the process where, by definition, discoveries can never happen.

After presenting these propositions of statistical positivism, we move to the empirical study where we applied this research paradigm and empirically instantiated the consequences of its propositions.[2]

Empirical arguments

In the empirical studies, we used Achievement Goal Theory (AGT) and Self-Determination Theory (SDT), theories that are widely used in motivation, education, and other domains of psychology (Elliot & Hulleman, 2017; Ryan & Deci, 2017; Senko, Hulleman, & Harackiewicz, 2011). Their synthesis has been a focus of several investigations (Ciani, Sheldon, Hilpert, & Easter, 2011; Duda, Chi, & Newton, 1995; Dyrlund, 2009; Ntoumanis, 2001; Standage, Duda, & Ntoumanis, 2003). In the current empirical study, we utilized a variable-based approach guided by statistical positivism. In the empirical study in Chirkov and Anderson (2018), we used a case-based methodology supported by the realist paradigm. Our purpose was to compare these two paradigms and their underlying methodologies in an attempt to discover the motivational mechanism of academic activity by integrating AGT and SDT.[3]

Empirical study. This study followed a conventional variable-based procedure of collecting and analyzing data. It included two constructs from AGT: mastery-approach and performance-approach goals, as well as some from SDT: competence and autonomy needs satisfaction and the five forms of motivational regulation: extrinsic, introjected, identified, integrated, and intrinsic. These five types of regulation were combined in three larger motivational constructs: controlled regulation: extrinsic + introjected; autonomous regulation: identified + integrated; and intrinsic motivation (Chirkov, Ryan, & Willness, 2005). Three outcome variables were used in this study: students' well-being, academic dedication (the intention to continue education), and academic achievement (students' grades). There is no consensus among motivation researchers regarding the nature of relations between these two theories.

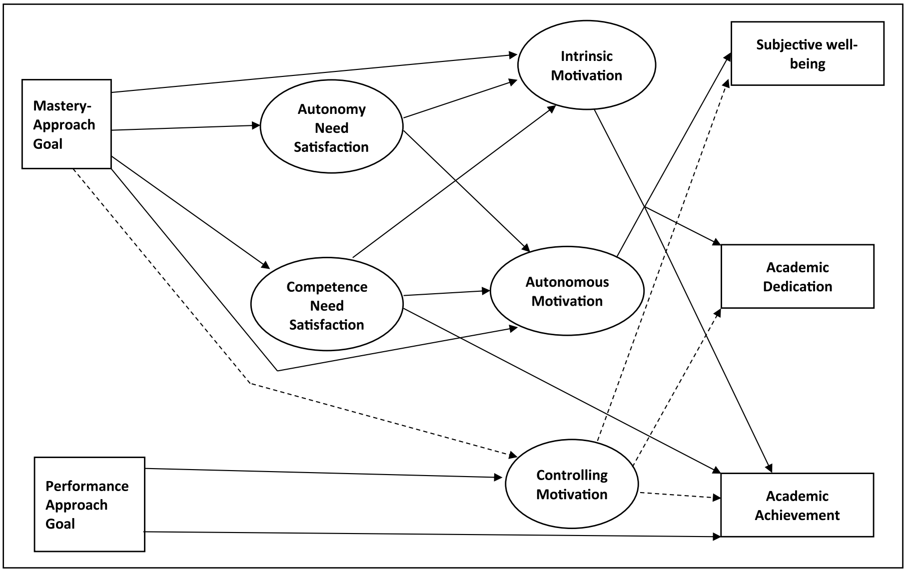

Figure 1. A proposed theoretical integrated model. Solid lines indicate positive and dashed lines – negative relations.

Ciani et al. (2011) treated the SDT constructs as the mechanisms moderating the emergence of different achievement goals, whereas Standage et al. (2003) and Dyrlund (2009) proposed that achievement goals statistically predict the SDT constructs and, through these constructs, predict outcomes. Only Standage et al. (2003) included a single outcome variable: intention to partake in leisure-time physical activity. In contrast, the other two studies did not include outcome variables, which substantially weakens their integrative syntheses. We followed the logic of Standage et al. (2003) and Dyrlund (2009) and proposed our theoretical integrated model, depicted in Figure 1, which we then empirically tested.[4]

Method

Participants

Students (n = 457) enrolled in first-year courses at a mid-size Canadian university completed the survey to attain extra credits. Data from 62 students who self-identified as non-Canadian and 10 additional students who did not complete the survey were removed from the sample, and 385 participants were included in the analysis. Among those included, 71% were female, and the mean age was 19.6, with a range of 17 to 45 years.

Measures

We used validated and published measures for all the constructs. The wording of some items was modified to meet the nature of our study and the level of our participants. The

Table 1. The Scales, the Number of Items, and the Cronbach's Alpha Reliability Coefficient.

Scale	Number of items	Cronbach's alpha
AGQ-R: Mastery-Approach Goals	3	.729
AGQ-R: Performance-Approach Goals	3	.816
BPNS: Autonomy satisfaction scale	7	.612
BPNS: Competence satisfaction scale	6	.736
SRQ-A: Intrinsic Motivation sub-scale	3	.799
SRQ-A: Autonomous Motivation sub-scale	6	.746
SRQ-A: Controlling Motivation sub-scale	6	.867
Academic Commitment Scale (used in 2013)	5	.726
SWLS	5	.819
PANAS	20	.879

following measures were used: The Achievement Goal Questionnaire – Revised (AGQ-R; Elliot & Murayama, 2008), the Basic Psychological Needs Support Scale (BPNS; Vansteenkiste, Duriez, Simons, & Soenens, 2006), the Self-Regulation Questionnaire – Academic (SRQ-A; Chirkov et al., 2005), The Academic Commitment Scale (Vallerand, Fortier, & Guay, 1997), the Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985), and the Positive and Negative Affect Scale (PANAS; Watson, Tellegen, & Bartol, 1988). Age, year of study, and self-reported Grade Point Average (GPA) were also recorded. Table 1 presents the number of items and Cronbach's alpha reliability coefficients of all scales.

Procedure

Participants were recruited using an online research recruitment program, and they completed the survey on paper in a classroom. A maximum of 20 students were surveyed at a time. A research assistant was present at all times and debriefed participants after the study. The entire session took approximately 20 minutes, and it was approved by the university Research Ethics Board.

Results

After checking for completeness of data and primary descriptive statistics, the intercorrelations among all variables were calculated (Table 2).

Using EQS 6.2 software, a path analysis was conducted to test the theoretical model depicted in Figure 1. Before testing the model, the path analysis assumptions were verified. There were 38 parameters to be estimated in the tested model, and the sample size was sufficient for testing models with up to 38 parameters. There was no evidence of multicollinearity in the data as the determinant of the inputted variance-covariance matrix was greater than 0.0001 (0.19360D-03). There was also no evidence of a specification error as the residuals of the variance-covariance matrix were small and centered around zero.

Table 2. Means, Standard Deviations, and Intercorrelations of the Variables.

Measure	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	19.64	1.01	-												
2. Years in university	1.38	0.83	.264***	-											
3. Mastery-Approach Goal	4.06	0.67	*n.s.*	*n.s.*	-										
4. Performance-Approach Goal	3.89	0.89	*n.s.*	*n.s.*	.298***	-									
5. Performance-Avoidance Goal	3.72	1.07	*n.s.*	*n.s.*	*n.s.*	.489***	-								
6. Autonomy Satisfaction	3.38	0.55	*n.s.*	.131*	.178**	*n.s.*	−.104*	-							
7. Competence Satisfaction	3.52	0.65	*n.s.*	.168**	.397***	.157**	*n.s.*	.579***	-						
8. Intrinsic Motivation	2.22	0.95	.125*	*n.s.*	.374***	*n.s.*	*n.s.*	.343***	.333***	-					
9. Autonomous motivation	4.12	0.63	*n.s.*	*n.s.*	.349***	.138**	*n.s.*	.227***	.284***	.267***	-				
10. Controlling Motivation	2.65	1.03	−.129*	*n.s.*	−.212***	*n.s.*	.129*	−.285***	−.224***	−.236**	−.318***	-			
11. Psychological Well-being	3.94	0.41	*n.s.*	.143*	.267***	*n.s.*	*n.s.*	.402***	.486***	.208***	.357***	−.256***	-		
12. Subjective Well-being	3.49	0.55	*n.s.*	.162*	.341***	*n.s.*	−.154**	.554***	.665***	.342***	.343***	−.215***	.697***	-	
13. Academic Dedication	4.39	0.67	*n.s.*	*n.s.*	.282***	.160**	*n.s.*	.316**	.448***	.170**	.295***	−.213***	.327***	.365***	-
14. Academic Achievement	3.17	1.01	−.105*	*n.s.*	149**	.292***	*n.s.*	*n.s.*	.241***	*n.s.*	.104*	*n.s.*	*n.s.*	.114*	.104*

Note: *p< .05, **p < .01, ***p <.001; n.s. - non-significant.

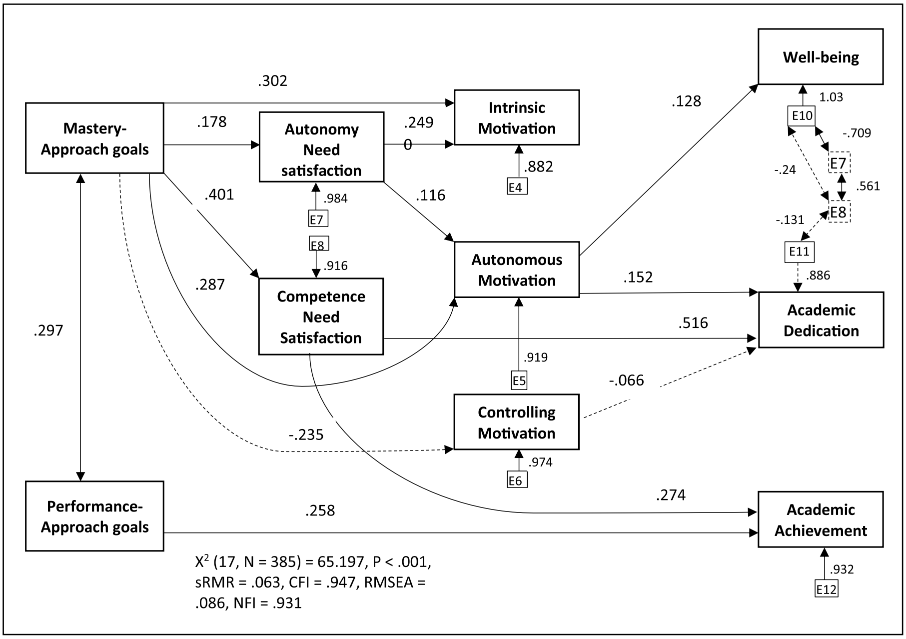

Figure 2. The statistical integrated model. Note: the standardized path coefficients are shown; all paths are significant at p< .05 unless otherwise indicated.

The theoretical model presented in Figure 1 had a poor fit to the data: X^2 (20, N = 385) = 262.45, $p < .001$, sRMR = .136, CFI = .732, RMSEA = .178, NFI = .723. After a series of model modifications supported by the Lagrange multiplier and Wald tests, the best-fitting model was found by: (a) allowing the error variances of highly correlated variables to correlate in the model, which included subjective well-being (SWB) and competence, competence and autonomy, autonomy and SWB, and competence and dedication (the variables correlated at r = .66, .58, .55, and .45 and the error variances correlated at r = -.24, .561, -.709, and -.131, respectively); (b) adding paths from performance-approach goals directly to achievement (theoretically supported by considerable research) and from mastery-approach goals directly to controlling motivation; and (c) removing paths from autonomous motivation to achievement, from controlling motivation to SWB, and from autonomy satisfaction to dedication. This modified model (Figure 2) provided a good fit to the data, X^2 (17, N = 385) = 65.197, $p < .001$, sRMR = .063, CFI = .947, RMSEA = .086, NFI = .931 (standardized path coefficients presented in Figure 2).

As hypothesized, the mastery-approach goals positively statistically predicted autonomy and competence need satisfaction (.178 and .401, respectively, at $p < .05$), with 3.2% of the variance in autonomy and 16.1% of the variance in the competence accounted for. Also, as expected, the mastery goals and autonomy need satisfaction positively statistically predicted intrinsic (.302 and .249, respectively, at $p < .05$) and autonomous motivation (.287 and .116, respectively, at $p < .05$). These associations mean that students who adopt mastery-approach goals in their classes feel their basic psychological needs are satisfied; also, those who have adopted mastery-approach goals and who have experienced satisfaction of a need for autonomy are intrinsically and autonomously motivated toward their learning. Contrary to our predictions, however, competence satisfaction did not significantly predict intrinsic and autonomous motivation (.069 and .104, respectively, at $p > .05$, not shown on the figure). Together, mastery goals, autonomy, and competence satisfaction predicted 22.1% of the variance in intrinsic and 15.6% of the variance in autonomous motivation. These associations indicate that mastery goals statistically predict self-determined academic motivation via need satisfaction (primarily the need for autonomy) and that self-determination theory constructs serve as mechanisms and outcomes of the achievement goal variable.

Further contradicting the hypothesized integration of SDT and AGT, the performance-approach goals did not statistically predict controlling motivation (.079 at $p > .05$, not shown on the figure), the association that supports the theoretical integrative model. However, the mastery-approach goals negatively statistically predicted controlling motivation (-.235 at $p < .05$). Together, mastery- and performance-approach goals accounted for 5.1% of the variance in controlling motivation. First, this set of relations only partially corresponds to the theoretically predicted relations. Second, it indicates that students with mastery goals tend to feel less controlled in their motivation; their acceptance of performance goals did not associate with their motivation. Autonomy satisfaction and intrinsic motivation did not statistically predict any outcome variables. Competence satisfaction predicted achievement (.274 at $p < .05$) and dedication (.516 at $p < .05$), while controlling motivation negatively predicted dedication (-.066 at $p < .05$). Autonomous motivation positively predicted both SWB and dedication (.128 and .152 respectively, at $p < .05$). These associations mean that the motivation of students has an important impact on their academic outcomes. Specifically, autonomously motivated individuals experience more well-being and dedication. Those who feel that their competence is fulfilled are more dedicated, and they have high achievements, while those who are more pressured and controlled in their studies demonstrate low dedication toward continuing their studies.

The tested model provides a relatively consistent picture of the relations among constructs of the two theories that correspond to the results presented in the literature. Although some associations did not fit the theoretical predictions, they were considered "new" discoveries to be explored in subsequent studies. This is a typical example of the type of statistics-based positivist study that constitutes the mainstream of motivation research. A reflection on this model follows.

Critical reflections on the statistics-based integrated model

What does this statistical model represent? What kind of psychological reality does it reflect? What information does it provide to researchers? Are our conclusions valid and justified?

According to statistical positivism, the presented statistical model that was calculated on a sample of students reflects the near universal relations among variables and displays causal motivational mechanisms of academic behavior. Statistical positivists have several reasons to think this way. First, such a model is theoretically driven. Researchers assume that the postulated theoretical associations reflect substantial and causal relationships that constitute real motivational mechanisms. They also assume that every student is motivated by the relationships described by this model. To test these theoretical assumptions, we used statistical inferences based on the covariances among the variables. The primary instrument for decision making regarding the scientific importance of the results was the Null-Hypothesis Significance Testing (NHST) procedure of statistical significance of the model as a whole as well as specific relations in particular. The established statistical model served the purpose of verifying the postulated theoretical model, thus equating theoretical reasoning with statistical inferences. The established significant associations in the predicted directions were interpreted as real and substantial causal relations among the motivational constructs.

This assumption represents a fundamental confusion of the objective reality that exists independently of our knowledge about it with researchers' conceptual schemes and abstractions about this reality. Theories of motivation predict that an individual student who adopts, for example, a mastery-approach goal will have self-determined motivation, whereas, in our study, we established covariances of the scores on the goal measure with the scores on the motivation measure. By their very nature, the theoretical predictions and statistical covariances are not symmetrical. They address different levels of reality. The theoretical model reflects real motivational mechanisms that function within and determine the behavior of individual students. The statistical model is an abstract and formal visual representation of the statistical associations among the variables that researchers have constructed, operationalized, and measured on the conveniently sampled self-selected students. Positivist researchers assume that this representation is a valid depiction of psychological reality, and they believe it can be used to theorize about human psychology. It is well known that the arrows in the constructed model and the corresponding regression coefficients indicate the proportions of variance that one variable accounts for in other variables. These accounts have little or no relevance to the substantial and causal relations that exist in real individuals. We believe this is the major illusion that maintains the attraction to and use of inferential statistics in motivation research. It is much easier to use formal statistical models under the impression that they represent intra-individual motivational functioning than it is to study the same mechanisms where they actually operate. This latter study must be done by meticulously investigating motivated individuals in different conditions and at different times in their lives and extracting knowledge of the causal dynamics of motivation by the power of the researcher's mind, not merely relying on the power of statistical programs (Meehl, 2006; Valsiner, 1986b; Znaniecki, 1934).

Another wrongful assumption of the above interpretation is that NHST is used to confirm theoretically driven relations (Cohen, 1994; Lykken, 1968; Neckerson, 2000; Rozeboom, 1960). The level of statistical significance, which indicates the probability of discovering the statistical associations by chance, is interpreted by statistical positivists as an indicator of the presence or absence of essential and real relations among constructs and as corroboration of theoretical predictions (Gigerenzer, 1993; John, 1992; Lambdin, 2012). According to this logic, non-significant associations (associations believed to be discovered by chance) should be removed from a theoretical discussion because they do not reflect real and meaningful relations among the variables. The use of statistical significance as a primary decision-making tool can be illustrated by our attempts to modify the proposed model to increase its fit to the empirical data. Although we started with theoretical assumptions, ultimately, it was the program that strongly influenced what path we added or removed, what errors were set to correlate, and which not to correlate in order to achieve a well-fitted model. Some of these requirements went against the proposed theoretical model, and some were of purely technical or statistical importance. Although handbooks on statistical analysis remind researchers that model modifications should be theoretically meaningful, we executed these modifications with one goal in mind: to get a good fit model—without reaching this fit, the research cannot be published. This is important because unpublished research does not contribute to scientific knowledge, which means that statistically insignificant relations are scientifically useless. Eventually, the statistical program decided what constituted an important scientific (and publishable) contribution and what did not. Knowledge construction was left to an algorithmic statistical routine, while the creative and insightful aspects of real scientific discovery were removed from this process. Following Bourdieu (2004), we may say that the desubstantialization of psychology by statistics is happening in positivist studies. As Danziger (1985) noted,

> Psychology appears to be unique in the degree to which statistical inference has come to dominate the investigation of *theoretically* postulated relationships. In this discipline it is generally assumed without question that the only valid way to test theoretical claims is by the use of statistical inference. This assumption is associated with an implicit belief in the theory-neutrality of the techniques employed. (p. 3)

Also, sociologist Znaniecki, a furious critic of the use of statistical methods in social sciences, noted:

> By making the study of facts subservient in advance to its final purpose of a mathematical play with symbols, not only does it fail to stimulate progress in the analysis of these facts, but actually obstructs it. Thus, the worst mistake of mediaeval scholasticism is here repeated: juggling with concepts instead of investigating reality has to be again accepted as the essence of science. (1934, p. 231)

Furthermore, such an assumption constitutes an *ecological fallacy*, a wrongful belief that the results obtained on a sample level represent and can be interpreted at the level of individuals (Robinson, 1950). However, there is no justification that this model, which is based on nearly 400 participants, represents any participant included in the sample let alone individuals outside of it. Individual participants are “dissolved” into the aggregated data, the resulting de-individualized and a-contextual model represents nobody; thus, it is difficult, perhaps impossible, to apply these statistical associations to real people in real situations (see also similar critiques in Hammersley, 2012; Toomela, 2010).

All constructs in our statistical model were operationalized and measured (see Table 1). All the covariance-based associations among these variables are printed in the program outputs, and they could be given to other researchers who might question the model. As such, the model of psychological mechanisms and its components and structures were made empirically explicit and visible. Therefore, here we have a clear representation of the extreme empiricist claim that the psychological reality is what is empirically verifiable. This empiricist assumption strongly contradicts the fundamental thesis of any science, that the empirical regularities that are observed by researchers are manifestations of deep unobservable structures and mechanisms that generate these regularities. These mechanisms have to be inferred by the reasoning and creative imagination of researchers, not be presented by formal statistical models (Bhaskar, 1975/2008; Bunge, 2004). For example, Kurt Lewin "argued that the reference of scientific theory was the lawfulness of the genotypical level of events and not the fluctuating phenotypical conjunctions through which this lawfulness manifested itself in the empirical world" (Danziger, 1985, p. 7).

The operationalization of psychological and motivational constructs is another point of serious critique of the empiricism of this approach (Bickhard, 2001; Chirkov, 2016; Feest, 2005; Grace, 2001; Green, 1992; Leahey, 1980). The main arguments against these operationalizations are: (a) the operationalizations in psychology produce the phenomena of interest rather than measure pre-existing entities[5]; (b) they represent very restrictive forms of psychological concepts/constructs; (c) when these operationalizations have been accepted by scientific communities (e.g., when a particular scale becomes widely accepted as a representation of a motivational construct), they start dictating the understanding and interpretation of complex theoretical ideas; and (d) some concepts that refer to deep components of the motivational machinery cannot and should not be operationalized; for example, "evolution," "unconsciousness," "libido," "basic psychological needs," and many others.

One reason why extreme empiricism and operationalism have been attractive to researchers is because these requirements comply with the positivist thesis that "good science" is free from any type of "subjectivism" in devising scientific propositions (Halfpenny, 1982/2015). Reasoning, intuition, insight, and, finally, the creative discovery of new ideas about the subject matter are considered too "subjective," too dependent on the peculiarities of a researcher's mind and, thus, too "soft" to be a solid base for real science. To eliminate this subjectivism, according to positivism, reason has to be submitted to empirically verifiable experiences. Real science is about confirmable facts and empirical regularities, the discovery of which should be independent of the whims and fancies of researchers' temperament and intelligence.

We used a relatively large sample of participants and tried to solve several problems by using this sampling methodology. We needed such a sample because only in samples with diverse participants we may get between-individual variability and, hence, use multivariate statistics. Sampling allowed us to calculate aggregated data that created an impression that what we discovered are relatively stable regularities that can be generalized as scientific laws of human motivation. Danziger (1990) commented on this purpose of sampling in the social sciences:

> The major methodological implication of these [statistics-based investigation of social issues] highly effective demonstrations was that the inherent lawfulness of human conduct would become apparent only if observations on a large number of individual cases were combined. This led to an infatuation with large samples, for only through them it seemed could the laws governing human action be made manifest. (p. 76)

By using a large sample, we aimed to secure the generalizability of our results to a larger population of university students. In addition, by using a large sample, we tried to balance out individual differences and various outliers to secure extraction of the associations that are expected to be nearly universally valid for college students. Following a conventional practice, we did not use probability sampling because we had the unjustified belief that using a relatively large number of participants would enable our results to be generalizable to a population of university students, regardless of the fact that we used convenience sampling.

Unfortunately, all these assumptions and beliefs are wrong. In order to generalize from a sample to a population and to justify a claim of discovering nearly universal laws of motivation, researchers need to have probability and representative samples (Thompson, 2012). Because our model was based on a convenience sample of participants and was not representative of other Canadian college students (as is typical for motivation studies), and because it may be biased, the discovered relations cannot be generalized to a larger population. Many researchers try to compensate for this violation of the requirement for statistical generalizations by replicating the regularities on different samples or by using a meta-analysis of different studies to reach knowledge generalization. Still, an implicit assumption here is that the higher the number of participants involved in an inquiry, the more reliable and more generalizable are the regularities among the variables established on these samples.

The recent "crisis of replicability" occurred when replications of empirical regularities across different samples failed. These failures have produced a loud outcry among researchers that psychology as a science is in crisis because its results are not replicable (Gilbert, King, Pettigrew, & Wilson, 2016; Open Science Collaboration, 2015). Characteristically, this crisis exists predominantly for positivists, who see the goal of science to be discovering empirical regularities. For critical scientific realists, a search for replicable empirical associations is a futile endeavor because these regularities depend on a multiplicity of underlying mechanisms and contextual conditions, which are very difficult (if even possible) to replicate (Bhaskar, 1975/2008; Sayer, 1992; Tsang & Kwan, 1999). According to realists, efforts of researchers should instead be aimed at discovering causal psychological mechanisms and establishing stable/universal and changeable/specific aspects and conditions for their functioning by using *conceptual replication* (Crandall & Sherman, 2016; Schaller, 2016). Replications of the hypothesized mechanisms can be done, but only under the conditions of closed systems; for example, in highly controlled laboratory experiments.

Also, by eliminating outliers and deviant cases, we "polished" our sampled psychological reality to make this reality meet the requirements of statistical analysis. Znaniecki (1934) commented on this matter:

> Science is reason challenging experience forcing it into a rational order. An exception is a revolt of experience against reason. Statistical science, faced with such revolt, passively relinquishes its claims and withdraws from the struggle into the realm of pure mathematical concepts. (p. 233)

Finally, by using aggregated data, we lost individuals as the real bearers of psychological motivational mechanisms and moved away from any opportunity to access these mechanisms. Even if our sample were a probability and representative one, the enumerative generalization of obtained regularities to a population of Canadian students would be only a weak form of scientific generalization. Its weakness is that it does not generate new knowledge over and above the empirical regularities discovered in the samples. As far back as the 17th century, Francis Bacon (1620/2000) called such a form of generalization "naïve" and "childish," and he deemed it unsuitable for scientific research.

The strong appeal of complex multivariate statistical calculations (such as multivariate regression analysis, path analysis, and SEM) is in their capacity to simultaneously enter and analyze several variables and their associations. These multi-variable analyses create an illusion that, by doing them, researchers uncover the systemic nature of a phenomenon under investigation. A quick look at all the positive and negative arrows presented in Figure 2 alleviates any doubts for many researchers that this model represents the complex nature of human motivation.

However, this is another illusion. In our statistical model, we have a formal compilation of the variables based on their covariances with each other. Why do they covary? Do these covariances reflect causal and/or interdependent/reciprocal relations among the variables? Or, are they reflections of accidental associations because of peculiarities of the samples and measures used? We do not know the answers to these questions; by using only statistics, we will never know. We have forgotten that motivational mechanisms are not formal compilations of variables through statistical associations. In fact, they are systems of interdependent socio-cultural, psychological, and psycho-physiological components that function through constant interactions with each other. These interactions are extended in time and embedded in a particular context that influences the dynamic and quality of these interactions (Juarrero, 1999); thus, they should be studied using a systemic approach (Capra & Luisi, 2014). As Znaniecki (1934) commented,

> The characters (aspects or facts) of any particular system, object, process, are not detached entities: they belong together and are mutually interdependent; knowledge of the system, object or process does not mean knowledge of each of the characters separately, but of all of them together as interdependent. Any progress in knowledge involves not only the discovery of new characters, but also a different and better understanding of the way all the characters, new and old, are combined in the given system. (p. 231)

The statistical analysis of motivational variables destroys the systemic nature of real motivational mechanisms. Such an analysis completely precludes the understanding of these systemic, temporal, and contextual aspects of motivational functioning.

By testing our model, we investigated the constructs and relations that we, researchers following other researchers, developed and imposed upon the real motivational mechanisms instead of inquisitively investigating and extracting these relations from the realm of students' actual functioning. Both the operationalization of the constructs and the directions of relations among them were suggested by researchers with little to no evidence about how students actually experience these constructs and how these constructs are related in their lives. We imposed our own conceptualizations and understandings on the psychological reality of academic functioning in order to convince others that this is how motivation really works within individual students. In actuality, the psychological reality of academic motivation is left unrepresented by this statistical model and not a single participant can identify with its claims. William James coined the term "psychologist's fallacy," which "referred to the psychologist's tendency to substitute his own categories for those of the person being studied" (Danziger, 1985, p. 12). For, example, as mentioned above, researchers have no consensus about how the SDT and AGT constructs interact among each other in producing various outcomes. Is SDT a mechanism of AGT or is AGT a mechanism of the SDT regulations? Researchers have claimed both of these opposing ideas. Statistics may equally justify either of these options, depending on the order in which researchers enter these constructs into the equations. Whatever order they use, it still remains unclear how these different motivational constructs really interact among themselves in regulating students' academic activities. This means that the motivational reality of university students remains untouched by our "play with the statistical symbols."

Because of the directional arrows among the variables and based on the idea that one variable accounts for a portion of the variance in another variable, many motivation researchers interpret these statistical associations as causal, following the Humean–Pearsonian interpretation of causality. As we explained, this outdated and long-time rejected conception of causality continues to guide positivist researchers' theorizing about the "effects" and "influences" in their statistical models. For example, Ntoumanis (2001) provided the following interpretation of the interaction of ego and task orientations in statistically predicting self-determined motivation scores: "Ego orientation was not *influential* in *determining* [emphasis added] the self-determination of high-task oriented individuals" (p. 407). Standage et al. (2003) used path analysis

> to provide greater insight into the motivational processes that account for varying levels of student motivation and also to examine the degree to which this motivation in turn predicts [it is not clear if this is a statistical or causal prediction] students' intention to partake in physical activity in their leisure time. (pp. 104–105)

It is also important to note that statistical associations do not differentiate the theoretically meaningful associations that can reflect the real causal relations and the accidental or spurious correlations that may happen because of various artifacts.

In addition, researchers may utilize different operational definitions and different measures to assess the same theoretical constructs, and they may use diverse convenience samples from their universities; as such, one may say that their statistical models have the potential to be strongly distorted and wrongfully represent the theoretical models that guide their research.[6] In our study, several associations were not predicted while others even contradicted the theoretical assumptions that guided our investigation. Such inconsistencies within the results preclude their replicability (see above on the crisis of replicability) and, thus, they prevent any accumulation of knowledge and the scientific progress that can be associated with such an accumulation.

Conclusion

We produced a statistical simulacrum that is fallaciously believed to represent the real world of academic motivation. During this production, we substituted theoretical thinking about psychological reality with a formal statistics-based decision-making grounded in the null-hypothesis significance testing of the model. This substitution created an illusion of a logical and objective judgment about the value of the obtained results. We suggested that we discovered law-like relations that represent the systemic nature of motivational mechanisms, and that these results represent nearly universal regularities that can be used to explain students' behavior and can be generalized to a broader population. Instead, we obtained a de-personalized, a-contextual, and a-historical abstract conglomerate of variables connected by statistical associations that reflect nothing more than these variables' covariances. Moreover, there is nothing causal or determining about these relations because no matter what statistical manipulations we perform, correlation (covariance) does not mean real causation. Therefore, despite all the efforts and resources we put into conducting this study, the extracted model has no value and provides little to no contribution to scientific knowledge. Danziger (1990) echoes this conclusion by saying that

> not only did these methodological aberrations [intensive use of inferential statistics] legitimate a large-scale waste of time, efforts, and resources, they also confined psychological theorizing in an increasingly narrow mold, thus closing the door on alternative conceptualizations and practices that might have reversed the process of intellectual decline that the discipline was now beginning to suffer. (p. 155)

Byrne (2009) noted that "certainly, nomothetic scientist approaches based on transferring the language of variables to the social [and psychological] world has—in brutal summary—been largely useless" (p. 520). Such statistical models provide no opportunity to psychologists to move further in understanding and explaining human motivation, except by adding new variables and/or more people to samples. Neither of these additions has any potential in advancing motivational psychology further in its investigation of human motivation. Psychologists need to study motivation where it is actually happening—in the embodied human beings that are embedded into historical and sociocultural contexts—and apply their theoretical thinking and intellectual sophistication, rather than statistical programs, to discover real motivational mechanisms. One version of such an approach is reported in Chirkov and Anderson (2018).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Danziger (1987) called this approach "the Galtonian Model" (p. 35) and Gigerenzer (1987) called it "the Galton-Pearson tradition/program" (p. 7). This conveys that such thinking was initiated by Galton and implemented by Pearson in their anthropometric and biometric research.

2. The empirical studies are based on Anderson (2015).
3. This section of the article presents only the part of our investigation (Anderson, 2015) that is relevant for the purpose of the current presentation.
4. Because we wanted to compare the two paradigms on the same theoretical model, Study 1 includes only the constructs that we used in Study 2.
5. For example, there is a well-known expression that "intelligence is what the intelligence test measures."
6. For example, in the studies of Ntoumanis (2001) and Standage et al. (2003), the authors repeatedly announce that many discovered statistical associations do not correspond to their theoretical predictions: "The only exception was that ego orientation positively predicted intrinsic motivation to experience stimulation. This was not expected and it may be attributed to the possibility that athletes with high ego orientation will report positive sensations and excitement when they are able to meet their criteria for success (i.e. winning and outperforming others)" (Ntoumanis, 2001, p. 406). This explanation has never been tested. The researcher continues, "These results are contrary to the predictions of achievement goal theory and the hypotheses of this study and can be partially attributed to the high perceived competence of the participants (85% scored 4 or above on a 7-point scale)" (p. 406). Standage et al. (2003) assert, "In contrast to Ntoumanis (2001), who … found a moderate to strong path between cooperative learning and relatedness, the path between mastery climate and relatedness in the present study was dropped because it was nonsignificant (.07) … [Because of this,] the present findings, … depart from theoretical postulations and are not consonant with our hypotheses" (p. 105). Naturally, this list may be continued. The authors explained these inconsistencies by constitutions of their samples, poor measurement scales, and other artifacts that have to be taken care of during the design of the study. However, ultimately, these inconsistencies make the discovered regularities controversial, unconvincing, and non-replicable.

References

Anderson, J. (2015). *Exploring the mechanisms of academic motivation: An integration of self-determination and achievement goal theories from a critical realist perspective* (Unpublished Master's thesis). University of Saskatchewan, Saskatoon, Canada.

Bacon, F. (2000). *The new organon*. Cambridge, UK: Cambridge University Press. (Original work published 1620)

Barlow, D. H., & Nock, M. K. (2009). Why can't we be more idiographic in our research? *Perspectives on Psychological Science*, *4*(1), 19–21. doi: 10.1111/j.1745–6924.2009.01088.x

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*(4), 396–403. doi: 10.1111/j.1745–6916.2007.00051.x

Bhaskar, R. (2008). *A realist theory of science*. London, UK: Verso. (Original work published 1975)

Bickhard, M. H. (1992). Myth of science: Misconceptions of science in contemporary psychology. *Theory & Psychology*, *2*, 321–337.

Bickhard, M. H. (2001). The tragedy of operationalism. *Theory & Psychology*, *11*, 35–44.

Blumer, H. (1956). Sociological analysis and the "variable". *American Sociological Review*, *21*(6), 683–690.

Bourdieu, P. (2004). *Science of science and reflexivity*. Chicago, IL: The University of Chicago Press.

Bunge, M. (2004). Does it work?: The search for explanatory mechanisms. *Philosophy of the Social Sciences*, *34*(2), 182–210.

Byrne, D. (2009). Causality and interpretation in qualitative policy-related research. In D. Byrne & C. C. Ragin (Eds.), *The SAGE handbook of case-based methods* (pp. 511–521). Los Angeles, CA: Sage.
Capra, F., & Luisi, P. L. (2014). *The systems view of life: A unified vision*. Cambridge, UK: Cambridge University Press.
Carlson, R. (1971). Where is the person in personality research? *Psychological Bulletin*, *75*(3), 203–219.
Carnap, R. (1995). *An introduction to the philosophy of science*. New York, NY: Dover Publications. (Original work published 1966)
Chang, H. (2009, Fall). Operationalism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/fall2009/entries/operationalism/
Chirkov, V. (2016). *Fundamentals of research on culture and psychology: Theory and methods*. New York, NY: Routledge.
Chirkov, V., & Anderson, J. (2018). Statistical positivism versus critical scientific realism. A comparison of two paradigms for motivation research: Part 2. A philosophical and empirical analysis of critical scientific realism. *Theory & Psychology*, *28*, 737–756.
Chirkov, V., Ryan, R. M., & Willness, C. (2005). Cultural context and psychological needs in Canada and Brazil: Testing a self-determination approach to internalization of cultural practices, identity and well-being. *Journal of Cross-Cultural Psychology*, *36*(4), 425–443.
Ciani, K. D., Sheldon, K. M., Hilpert, J. C., & Easter, M. A. (2011). Antecedents and trajectories of achievement goals: A self-determination theory perspective. *British Journal of Educational Psychology*, *81*, 223–243.
Cohen, J. (1994). Earth is round ($p < .05$). *American Psychologist*, *49*(12), 997–1003.
Crandall, C. S., & Sherman, J. M. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology*, *66*, 93–99.
Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*(11), 671–684.
Danziger, K. (1985). The methodological imperative in psychology. *Philosophy of Social Sciences*, *15*(1), 1–13.
Danziger, K. (1987). Statistical method and the historical development of research practice in American psychology. In L. Kruger, G. Gigerenzer, & M. S. Morgan (Eds.), *The probabilistic revolution* (Vol. 2, pp. 35–47). Cambridge MA: MIT Press.
Danziger, K. (1990). *Constructing the subject: Historical origin of psychological research*. Cambridge, UK: Cambridge University Press.
Danziger, K. (1992). The project of experimental social psychology: Historical perspectives. *Science in Context*, *5*(2), 309–328.
Danziger, K. (2000). Making social psychology experimental: A conceptual history, 1920–1970. *Journal of the History of the Behavioral Sciences*, *36*(4), 329–347.
Danziger, K., & Dzinas, K. (1997). How psychology got its variables. *Canadian Psychology*, *38*(1), 43–48.
Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*, 71–75.
Duda, J. L., Chi, L.-K., & Newton, M. (1995). Task and ego orientation and intrinsic motivation in sport. *International Journal of Sport Psychology*, *26*, 40–63.
Dyrlund, A. K. (2009). *An examination of the integrative relationship among the factors of achievement goal theory and self-determination theory: Addressing existing problems and missing links* (Unpublished PhD thesis). Tallahassee, Florida State University.

Elliot, A. J., & Hulleman, C. S. (2017). Achievement goals. In A. J. Elliot, D. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation* (2nd ed., pp. 43–60). New York, NY: The Guilford Press.

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100*(3), 613–628.

Elms, A. C. (1975). The crisis of confidence in social psychology. *American Psychologist*, *30*, 967–976.

Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences*, *42*(2), 131–149.

Ferguson, C. J. (2015). "Everybody knows psychology is not a real science": Public perception of psychology and how we can improve our relationship with policymakers, the scientific community, and the general public. *American Psychologist*, *70*(6), 527–542.

Fisher, R. A., Sir. (1970). *Statistical methods for research workers*. Edinburgh, UK: Oliver and Boyd.

Funder, D. C. (2009). Naive and obvious questions. *Perspectives on Psychological Science*, *4*(4), 340–344.

Gigerenzer, G. (1987). The probabilistic revolution in psychology—An overview. In L. Kruger, G. Gigerenzer, & M. S. Morgan (Eds.), *The probabilistic revolution* (Vol. 2, pp. 7–9). Cambridge, MA: MIT Press.

Gigerenzer, G. (1993). The superego, the ego, and the id of statistical reasoning. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 311–339). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gilbert, D. T., Fiske, S. T., & Lindzey, G. (Eds.). (1996). *The handbook of social psychology* (4th ed., Vol. 1). Boston, MA: McGraw Hill.

Gilbert, D. T., King, G., Pettigrew, S., & Wilson, T. D. (2016). Comment on "Estimating the reproducibility of psychological science". *Science*, *351*(6277), 1037-b.

Grace, R. C. (2001). On the failure of operationalism. *Theory & Psychology*, *11*, 5–33.

Green, C. D. (1992). Of immortal mythological beasts: Operationism in psychology. *Theory & Psychology*, *2*, 291–320.

Halfpenny, P. (2001). Positivism in the twentieth century. In G. Ritzer & B. Smart (Eds.), *Handbook of social theory* (pp. 371–385). London, UK: Sage.

Halfpenny, P. (2015). *Positivism and sociology: Explaining social life*. London, UK: Routledge. (Original work published 1982)

Halpin, P. F., & Stam, H. J. (2006). Inductive inference or inductive behavior: Fisher and Neyman-Pearson approaches to statistical testing in psychological research (1940–1960). *The American Journal of Psychology*, *119*(4), 625–653.

Hammersley, M. (2012). What's wrong with quantitative research? In B. Cooper, J. Glaesser, R. Gomm, & M. Hammersley (Eds.), *Challenging the qualitative-quantitative divide: Explorations in case-focused causal analysis* (pp. 27–56). London, UK: Continuum.

Hempel, C. G. (1942). The function of general laws in history. *Journal of Philosophy*, *39*(2), 35–48.

Hempel, C. G. (1966). *Philosophy of natural science*. Englewood Cliffs, NJ: Prentice-Hall.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*(2), 135–175. doi:10.1086/286983

John, I. D. (1992). Statistics as rhetoric in psychology. *Australian Psychologist*, *20*(3), 144–149.

Juarrero, A. (1999). *Dynamics in action: Intentional behavior as a complex system*. Cambridge, MA: A Bradford Books.

Koch, S. (1992). Psychology's Bridgman vs Bridgman's Bridgman: An essay in reconstruction. *Theory & Psychology*, *2*, 261–290.

Kruger, L., Gigerenzer, G., & Morgan, M. S. (Eds.). (1987). *The probabilistic revolution*. Cambridge, MA: MIT Press.

Lambdin, C. (2012). Significance tests as sorcery: Science is empirical—significance tests are not. *Theory & Psychology*, *22*, 67–90.

Lamiell, J. T. (2013). Statisticism in personality psychologists' use of trait constructs: What is it? How was it contracted? Is there a cure? *New Ideas in Psychology*, *31*, 65–71.

Langfeld, H. S. (1945). Symposium on operationism. *The Psychological Review*, *52*(5), 241–294.

Leahey, T. H. (1980). The myth of operationalism. *The Journal of Mind and Behavior*, *1*(2), 127–143.

Lilienfeld, S. O. (2012). Public skepticism of psychology: Why many people perceive the study of human behavior as unscientific. *American Psychologist*, *67*(2), 111–129.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, *70*(3), 151–159.

Mandler, G., & Kessen, W. (1964). *The language of psychology*. New York, NY: Wiley.

Manicas, P. T., & Secord, P. F. (1983). Implications for psychology of the new philosophy of science. *American Psychologist*, *38*(4), 399–413.

Meehl, P. E. (2006). When shall we use our heads instead of the formula. In N. G. Waller, L. J. Yonce, W. M. Grove, D. Faust, & M. F. Lenzenweger (Eds.), *A Paul Meehl reader: Essays on the practice of scientific psychology* (pp. 263–270). Mahwah, NJ: Lawrence Erlbaum.

Michell, J. (2003). The quantitative imperative: Positivism, naive realism and the place of qualitative methods in psychology. *Theory & Psychology*, *13*, 5–31.

Mill, J. S. (1965). *On the logic of moral sciences*. Indianapolis, IN: The Bobbs-Merril Company. (Original work published 1843)

Molenaar, P. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives*, *2*(4), 210–218.

Morris, W. E., & Brown, C. R. (2016). David Hume. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://plato.stanford.edu/archives/spr2016/entries/hume/

Mulaik, S. A. (1987). Toward a conception of causality applicable to experimentation and causal modeling. *Child Development*, *58*(1), 18–32.

Neckerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, *5*(2), 241–301.

Ntoumanis, N. (2001). Empirical links between achievement goal theory and self-determination theory in sport. *Journal of Sports Sciences*, *19*, 397–409.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716-aac4716. doi:10.1126/science.aac4716

Pancer, S. M. (1997). Social psychology: The crisis continues. In D. Fox & I. Prilleltensky (Eds.), *Critical psychology: An introduction*. London, UK: Sage.

Patnoe, S. (Ed.). (1988). *A narrative history of experimental social psychology: The Lewin tradition*. New York, NY: Springer-Verlag.

Pearson, K. (1957). *The grammar of science*. New York, NY: Meridian Books. (Original work published 1892)

Petrie, H. G. (1971). A dogma of operationalism in the social sciences. *Philosophy of Social Science*, *1*(2), 145–160.

Polkinghorne, D. (1983). *Methodology for the human sciences: Systems of inquiry*. Albany: State University of New York Press.

Ray, C. (2000). Logical positivism. In W. H. Newton-Smith (Ed.), *A companion to the philosophy of science* (pp. 243–251). Malden, MA: Blackwell.

Reichenbach, H. (1958). *The rise of scientific philosophy*. Berkeley: University of California Press.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*(3), 351–357.

Rodrigues, A., & Levine, R. V. (Eds.). (1999). *Reflections on 100 years of experimental social psychology*. New York, NY: Basic Books.

Rorer, L. G. (1991). Some myths of science in psychology. In P. E. Meehl, D. Cicchetti, & W. M. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of Paul E. Meehl*. Minneapolis: University of Minnesota Press.

Rosnow, R. L. (1981). *Paradigms in transition: The methodology of social inquiry*. New York, NY: Oxford University Press.

Rozeboom, W. W. (1960). The fallacy of the null-hypothesis significance test. *Psychological Bulletin*, *57*(5), 416–428.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. New York, NY: The Guilford Press.

Salmon, W. C. (2000). Logical empiricism. In W. H. Newton-Smith (Ed.), *A companion to the philosophy of science* (pp. 233–242). Malden, MA: Blackwell.

Sayer, A. (1992). *Methods in social science: A realist approach* (2nd ed.). London, UK: Routledge.

Schaller, M. (2016). The empirical benefits of conceptual rigor: Systematic articulation of conceptual hypotheses can reduce the risk of non-replicable results (and facilitate novel discoveries too). *Journal of Experimental Social Psychology*, *66*, 107–115.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, *46*(1), 26–47.

Smith, L. D. (1986). *Behaviorism and logical positivism: A reassessment of the alliance*. Stanford, CA: Stanford University Press.

Spadish, W. R., Cook, T. D., & Campbell, D. T. (2001). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin Company.

Stam, H. J. (2006). Pythagoreanism, meaning and the appeal to number. *New Ideas in Psychology*, *24*, 240–251.

Standage, M., Duda, J. L., & Ntoumanis, N. (2003). A model of contextual motivation in physical education: Using constructs from self-determination and achievement goal theories to predict physical activity intentions. *Journal of Educational Psychology*, *95*(1), 97–110.

Thompson, S. K. (2012). *Sampling* (3rd ed.). New York, NY: John Wiley & Sons.

Tolman, C. W. (Ed.). (1992). *Positivism in psychology: Historical and contemporary problems*. New York, NY: Springer-Verlag.

Toomela, A. (2008). Variables in psychology: A critique of quantitative psychology. *Integrative Psychological and Behavioral Science*, *42*, 245–265.

Toomela, A. (2010, July). Quantitative methods in psychology: Inevitable and useless. *Frontiers in psychology, 1*. doi: 10.3389/fpsyg.2010.00029

Tsang, E. W. K., & Kwan, K.-M. (1999). Replication and theory development in organizational science: A critical realist perspective. *Academy of Management Review*, *24*(4), 759–780.

Vallerand, R. J., Fortier, M. S., & Guay, F. (1997). Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. *Journal of Personality and Social Psychology*, *72*, 1161–1176.

Valsiner, J. (1986a). Where is the individual subject in scientific psychology? In J. Valsiner (Ed.), *The individual subject and scientific psychology* (pp. 1–14). New York, NY: Plenum Press.

Valsiner, J. (1986b). Between groups and individuals: Psychologists' and laypersons' interpretations of correlational findings. In J. Valsiner (Ed.), *The individual subject and scientific psychology* (pp. 113–151). New York, NY: Plenum Press.

Vansteenkiste, M., Duriez, B., Simons, J., & Soenens, B. (2006). Materialistic values and well-being among business students: Further evidence for their detrimental effect. *Journal of Applied Social Psychology*, *36*(12), 2892–2908.
Watson, D., Tellegen, A., & Bartol, K. M. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070.
Znaniecki, F. (1934). *The method of sociology*. New York, NY: Rinehart & Company.

Author biographies

Valery Chirkov is a Professor and a member of the Applied Social Psychology program in the Department of Psychology at the University of Saskatchewan, Saskatoon, Canada. Dr. Chirkov's research interests concern the philosophy and methodology of psychological and cultural research, theories of culture for psychologists, the application of self-determination theory of human motivation in cross-cultural research, and psychology of immigration and acculturation. In addition to numerous articles and book chapters on these topics, in 2011, he coedited the book *Human Autonomy in Cross-cultural Contexts: Perspectives on the Psychology of Agency, Freedom, and Well-being* and in 2016 published the textbook *Fundamentals of Research on Culture and Psychology: Theory and Methods*.

Jade Anderson is a PhD student at the University of Regina Johnson Shoyama Graduate School of Public Policy, Regina, Canada. Her research interests include methodology of applied research, integration of health and education systems, and disparities in access to healthcare in rural populations. Recent publications include *Realist interviewing: Componential and structural analysis applied to motivation research*, with Valery Chirkov.