Error Propensities Amongst Finance and Accounting Personnel: Can We Quantitatively Measure Illusion of Control or Chaos?

Keith I. Taylor[1] & Halil Kiymaz[2]

[1]Rollins College of Liberal Arts, Winter Park, FL, USA

[2]Crummer Graduate School of Business, Rollins College, Winter Park, FL, USA

Correspondence: Keith Taylor, Department of Business, Rollins College of Liberal Arts, Winter Park, FL, 32789, USA. Tel: 1-407-748-1193. E-mail: kitaylor@rollins.edu

Received: May 25, 2022 Accepted: August 7, 2022 Online Published: August 22, 2022

doi:10.5430/afr.v11n3p34 URL: https://doi.org/10.5430/afr.v11n3p34

Abstract

Most financial and accounting tasks and analytics, whether associated with the past or future, assume knowledge of process, variation, and statistics. Yet, finance and accounting personnel averaging 13 years of experience could not distinguish non-random from random time-series strings in an assessment using statistical process control charts. Respondents scored no better than guessing compared to a series of true-false questions. Latent class analysis methods within partial least square structural equation modeling successfully uncovered segments of respondents with large explained variance and significant paths to explicate tendencies toward type I or type II error rates, i.e., an illusion of control or illusion of chaos. Relationships between the desirability of control, personal fear of invalidity, and error rates were more varied than expected.

Keywords: type I and type II errors, time series, statistical process control charts, partial least squares structural equation modeling, illusion of control, illusion of chaos

1. Introduction

1.1 Importance

Does anything happen by chance? If an individual thinks of an event or outcome as random when it is not, one type of error is made. Similarly, if another individual thinks of an event or outcome as pre-determined when it is not, another type of error is made. This study offers a path to measure these constructs for a group of accounting and finance professionals.

The research of error types traverses many diverse domains: in clinical trials as in Davidson (1986); in the judiciary processes as in Rizzolli (2016); with climate change as in Anderegg, Callaway, Boykoff, Yohe, and Root (2014); with innovation as in Tellis (2013), among others. The ideas of determinism and free will have been reported throughout human history. For example, Bennett (1998) provides a lengthy history of the development of randomizers and divination in antiquity, describing early games of chance and outlining the advancement of probabilistic thinking. Today, randomizers from antiquity have been replaced by random number generators. Studies of games of chance have led to probability theory. Randomness has yielded to self-similar fractals of Mandelbrot, chaos theory of Lorenz, or Bose-Einstein dynamics in the film industry by De Vany. Zbilut (2004) suggested randomness was “an active process which informs order and vice versa” instead of any process of scientific separation (p. 6).

1.2 Contributions

This study offers the following research contributions to explore randomness and error types for accounting and finance professionals. First, by design, statistical process control (SPC) charts employ probabilistic rules to separate random and non-random sequences but have not used to assess respondent error tendencies. Second, this instrument equally assesses type I and type II error rates, addressing research imbalances favoring type I errors, i.e., illusory perceptions and maladaptive behaviors. Third, to address the shortage of actual accounting and finance practitioners in academic studies, industry professionals were selected using vetted respondents provided online by Qualtrics. Fourth, partial least squares structural equation modeling (PLS-SEM) was employed as an exploratory tool to link this randomness assessment tool to two previously validated psychological scales.

Our findings show that accounting and finance professionals, when viewing time series did not distinguish between a random and non-random series. Second, depending on the type of assessment, different error types appeared. Third, multiple SEM exploratory tools uncovered many unexpected and varied relationships between control, fear, and error rates. Fourth, the overall model revealed minimal explained variance, so an objective instrument to identify the illusion of control or illusion of chaos must continue. Fifth, using Qualtrics' panels of current accounting and finance practitioners would continue to provide depth within an engaged scholarship approach. And finally, an enhanced understanding of variation that business processes generate and that the resulting financial statements reflect could lead to real organizational improvement.

Our study has the following structure. The literature review section examined the consequences and costs of error types, differences in modes of variation, and select studies describing various finance-related behavior contradictions. Next, the operational tools section demonstrated how an SPC chart can separate random/non-random sequences as valid and reliable measures for any time series. The methodology, analysis plan, and proposed model using PLS-SEM were justified in the analysis section. These results were illustrated and included reliability, validity, explained variance, and path coefficients. And lastly, a discussion ensued with suggested implications and possible future trajectories.

1.3 Scholarship Review

Accounting and finance personnel should understand error tendencies as the consequences and costs of both types I and II errors are large. Within auditing literature, Kinney and Salamon (1982) investigate the use of the STAR approach and report an overall type II error risk of 36.5%. Deshmukh, Karim, and Siegel (1998) employ signal detection theory to assess management fraud and estimate type II errors which are estimated to be 10 to 99 times more costly than type I. They also cite a statement by Arthur Anderson, which estimated that 12% of all Big Six revenue was related to type II errors. As organizational longevity continually declines, as evidenced by the age of firms in the SP500, understanding error types could ameliorate business strategies and reduce stagnation, financial distress, and bankruptcy, as exampled by recent retail collapses and similar industry upheavals.

Henderson, Mead, van Dijke, Ramsay, McDowall, and Dennis (2008) describe two forms of variation, common cause or systemic variation and special cause or non-random variation. An error occurs when mistaking a common cause for a special cause, and in the reverse case, when mistaking a special cause for a common cause. Similarly, a type I error occurs if an observer considers a time series as non-random when actually random. Oppositely, if an observer considers a time series as random when actually non-random, a type II error occurs. From a Bayesian view, Harris and Osman (2012) proposed a type I error as an illusion of control (IoCON) and a type II error as an illusion of chaos (IoCHA) when perceiving the world as controllable or uncontrollable versus its actual state.

Behavioral accounting and finance anomalies are reported in many studies. For example, Uecker and Kinney (1977) assess the representative and protectiveness heuristics among practicing CPAs; Kahneman and Riepe (1998) offer advice concerning investor false beliefs and biases; Fenton-O'Creevy, Nicholson, Soane, and Willman (2003) examine the IoCON and trading performance among actual traders; Taleb (2004) observes trader foibles based on his career in the industry; McSweeney (2006) suggests that use of net present value created an illusion of certainty; and Whitson and Galinsky (2008) examine the relationship between illusory perceptions, control, and conspiratorial beliefs by priming subjects with either a stable or volatile stock market scenario. More recently, Ackert, Church, and Qi (2015) report various factors causing investors to hold inferior portfolios, and Heuer, Merkle, and Weber (2016) study actual investor misattributed behaviors, among many others. These few examples of a body of inquiry conducted over 40 years suggest the need to continually scrutinize bias and error tendencies with novel approaches as this study presents.

1.4 Operational Tools and Hypotheses

1.4.1 SPC Charts as an Endogenous Variable

Statistical process control chart methodology provides eight rules for identifying non-random observations (for a detailed discussion of SPC chart development and use with diverse time series, see Taylor & Kiymaz, n.d.). Initial graphical representations of the 18-point SPC charts in Figure 1 were produced from a random number generator with a mean of 0.0, a standard deviation of 0.05, and a sample size of 100,000. After creating the SPC chart using SigmaXL software, random and non-random sequences were extracted. For ten non-random strings, six of eight Nelson's rules (1984) were selected once, while rules one and two were selected twice. Each survey question was randomized. A scenario was constructed which utilized stock returns with the above-mentioned parameters. Appendix Table A1 shows two typical questions from the answer key. Each individual was scored based on his/her overall percent correct and rates of type I or type II error.

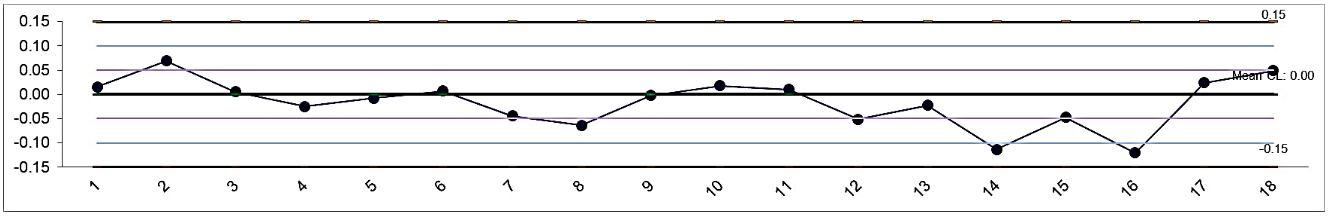

Figure 1. SPC Chart Randomness Diagram

Eighteen observations are from the average or centerline of 0.0. 0.15/-0.15 lines show control limits of three standard deviations.

1.4.2 Desirability of Control as an Exogenous Variable

Burger and Cooper (1979) developed the desirability of control (DC) scale. Individuals with high DC want to have control in life, prefer to influence others, often seek leadership, and appear decisive and assertive. Gebhardt and Brosschot (2002) created a Dutch version. They showed relationships between DC and LOC, coping style, repression, achievement motivation, personality characteristics, trait anxiety and depression, trait worry, burnout, and somatic complaints. Burger and Cooper (1979) and Gebhardt and Brosschot (2002) furnished coefficients of reliability, test-retest, and discriminant validity (mixed).

The DC scale is composed of 20 questions on a 6-point Likert scale of “this statement doesn’t apply to me” to “this statement always applies to me.” Two sample statements include “I prefer a job where I have a lot of control over what I do and when I do it” and “I try to avoid situations where someone else tells me what to do.” And of 20 items, seven are reverse coded. For this study, DC is a reflective, ordinal latent variable.

DC was anticipated to have a significant positive effect on type I error rate as individuals attempt to influence others and exert control. Therefore,

Hypothesis 1: High DC provides a significant positive path to type I error rate.

1.4.3 Personal Fear of Invalidity as an Exogenous Variable

Thompson, Naccarato, Parker, and Moskowitz (2001) developed the personal need for structure (PNS) and the personal fear of invalidity (PFI) scales from a basic human desire for structure and management of uncertainty. The authors noted instead of being concerned about structure; individuals might be more concerned with damage from committing errors and a high PFI. These persons were concerned with the risks of an undertaking, vacillating between courses of action, seeking alternatives, and exhibiting some distress when personal errors were enumerated. Past studies had reviewed both constructs together or with other variables (Clow and Esses (2005) for PNS-PFI-need for closure; Rietzschel, De Dreu, and Nijstad (2007) for PNS-PFI and for PNS-creativity; and Blais, Thompson, and Baranski (2005) for PNS-PFI-need for cognition). Thompson et al. (2001) furnished reliability coefficients and convergent and discriminant validity.

PFI is composed of 14 questions on a 6-point Likert scale of “this statement doesn’t apply to me” to “this statement always applies to me.” Two sample statements include “I may struggle with a few decisions but not very often” and “I never put off making important decisions.” And of 14 items, five are reverse coded. In this study, PFI is a reflective, ordinal latent variable.

PFI would have a significant positive effect on type II error rate to avoid making any errors, add qualifiers and subtleties, and might prefer difficult complexifying questions. Therefore,

Hypothesis 2: High PFI provides a significant positive path to type II error rate.

1.4.4 True-False Guessing as an Exogenous Variable

Fritz (1921) documented true-false guessing strategies and observed students chose more true answers when responding incorrectly. Krueger (1933), using frequency to assess correct answers in a true-false guessing strategy, found true rates of 60%. Morrison (1978) suggested an iterative statistical process to distinguish guessers from discriminators. Burton (2005) described myths related to true-false questions, and Poundstone (2014) suggested examiners were more likely to create true questions (56%). Cabeza, Rao, Wagner, Mayer, and Schacter (2001) examined brain images of subjects while employing recognition tests of true-false statements.

A true-false assessment in which questions were designed to be unanswerable was created by extracting various sentences from books in a personal library. Ten true questions were based on random word selection, and ten false questions were created by selecting either the word before or after the correct one within the same sentence. Appendix Table A2 shows two questions from the answer key. Each individual was scored based on his/her overall percent correct and rates of type I or type II error identically as the SPC chart assessment.

Following Fritz (1921), Krueger (1933), Burton (2005), and Poundstone (2014), respondents would answer a higher percentage of true responses to false responses. And when answering true, the probability of making a type I error would also increase. Therefore,

Hypothesis 3: The distribution of the actual percentage of True responses in the True-False assessment group will be significantly higher than the percentage from a random distribution.

Lastly, if respondents can distinguish between an actual deterministic series and a stochastic one, the percentage correct in the SPC error rate evaluation should be superior to respondents' guessing strategy. Therefore,

Hypothesis 4: The percentage correct of the SPC Error Rate assessment is significantly higher than the percentage correct in the True-False assessment.

2. Method

The proposed exploratory model for this study which delineates types of positive and negative relationships between latent constructs is shown in Figure 2 using a partial least squares structural equation model (PLS-SEM). Hypotheses H1 and H2 are related to the multivariate model, while hypotheses H3 and H4 are evaluations of sample attributes that should mirror historical studies.

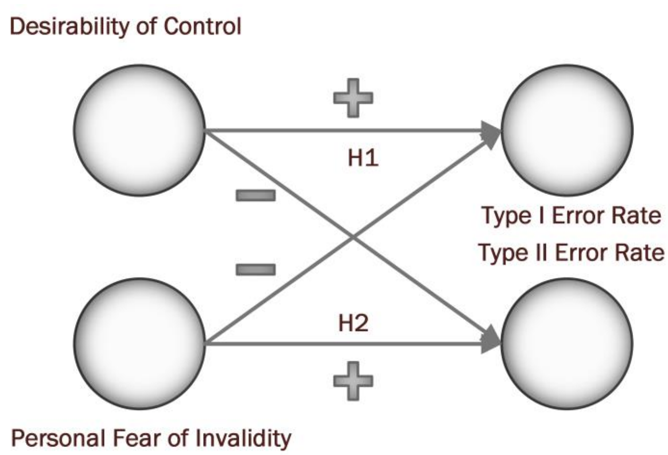

Figure 2. Proposed Model and Associated Relationships

Latent constructs are circles with arrows indicating the direction of relationships. Hypotheses are numbered on the path of the appropriate relationship.

2.1 Justification of Methodology

Following a pragmatic view, a multivariate design was employed based on Buhl, Goodson, and Neilands (2007) and Astrachan, Patel, and Wanzenried (2014). Following Christensen, Johnson, and Turner (2014), we undertook several pilot surveys that revealed issues related to the scenario, instructions, interpretations, and chart visibility, including HTML and image enhancements and timing. Following Hair, Hult, Ringle, and Sarstedt (2017), an eight-stage PLS-SEM procedure was employed.

2.2 Sampling

The survey instrument was based using 84 observed variables, 11 latent constructs, and the screening variables. Qualtrics also provided a panel selection process that included confidential respondent attributes. There was an initial soft launch to provide a small set of responses for examination and analysis before releasing to the entire panel. Data verification was executed, including calculations to check work experience, inconsistent answers, and ballot-box stuffing. Based on information obtained and submitted to Qualtrics, they performed an additional reexamination of all responses, which included the application of in-platform quality measures (speeding, duplication checks) and external data cleaning (back-end proprietary analyses).

2.2.1 Power and Precision

The following parameter estimates were used to determine the minimum sample size for this study: a minimum effect size of 0.20 based on previously unpublished research (using a 0.30 would not change in minimum respondent count); a standard power level of 0.80; the latent and variable counts previously mentioned; and a standard probability level of 0.05. These combined factors yielded a minimum sample size of 588 based on the model structure and 488 to detect an effect.

2.3 Subjects and Final Sample

The final sample of 1,050 responses was composed of 64% male and 36% female. Using generational categories as Dimock (2019) proposed, respondents were 54% Millennials, 35% Generation X, and 11% Baby Boomers. Fifty-two percent of individuals had some college or a college degree, while 47% had some graduate work or a graduate degree. Respondents had an average experience of 13 years and were based in the U.S.

3. Results

We report the relationships between the SPC chart and the true-false assessment for the final dataset in Table 1. Power was maintained when examining actual effect sizes with a sample size of 850. An evaluation of non-response bias was completed between the first and last halves following Blair and Zinkhan (2006) to assess for a device effect that was not significant.

Table 1. SPC versus True-False Assessments: Comparison of Attributes

This table compares attributes for SPC versus True-False Assessments for 1,050 accounting and finance professionals.

Relationship	**SPC**	**TF**	**Count**	***P-value***	**Effect Size**	**Odds ratio (OR)**
Percentage Correct	0.5022	0.4997	1,050	0.9062	0.0051	1.0051
Percentage Type1 Error	0.2119	0.2947	1,050	**0.0000**	**0.1911**	**1.3912**
Percentage Type2 Error	0.2859	0.2118	1,050	**0.0001**	**0.1717**	**1.3496**
First Answer Bias	0.5740	0.5397	1,050	0.1126	0.0692	1.0637
Percentage Correct (non-mobile device)	0.5019	0.5001	899	0.9380	0.0037	1.0037
Percentage Type1 Error (non-mobile device)	0.2046	0.2756	899	**0.0004**	**0.1667**	**1.3472**
Percentage Type2 Error (non-mobile device)	0.2935	0.2244	899	**0.0008**	**0.1583**	**1.3084**
Percentage Correct (mobile device)	0.5043	0.4974	151	0.9038	0.0139	1.0140
Percentage Type1 Error (mobile device)	0.2553	0.2368	151	0.7083	0.0431	1.0783
Percentage Type2 Error (mobile device)	0.2404	0.2368	151	0.9408	0.0085	1.0154

Note. Odds Ratio from Sullivan and Feinn, (2012). Significant relationships are bolded.

3.1 SEM Measurement Model

Following Klarner, Sarstedt, Hoeck, and Ringle (2013) and Hair et al. (2017), outer (measurement) models with reflective constructs were reviewed. DC and PFI, two reflective measures, were analyzed by examining convergent validity through their outer loadings, which are visible in Appendix Table B1. Values greater than 0.7 were acceptable, while those between 0.4 and 0.7 needed additional evaluation. DC7, DC10, DC16, DC19, and DC20 were retained, as were PFI3, PFI4, PFI5, PFI6, PFI13, and PFI14.

Cronbach's alpha and composite reliability measures are shown in Appendix Table B2. The SPC error rate reliability measures are 1.00 by definition of the errors themselves. Variance inflation factors (VIF) were examined for collinearity among the items, and DC and PFI passed. Fornell Larker criterion was used for reflective measures, DC and PFI, which are visible in Appendix Table B3 to measure discriminant validity. For single measures, Hetero-Trait Mono-Trait (HTMT) values were evaluated and are listed in Appendix Table B4. A complete bootstrap procedure was executed to verify that 95% confidence intervals contained 1.0. All measures passed discriminant validity substantiation.

3.2 SEM Structural Model and Hypotheses Evaluation

We followed a standardized procedure by Hair et al. (2017) to analyze the inner (structural) model and proposed hypotheses. We report our findings in Figure 3, in which the structural model, path coefficients, and R^2 are shown.

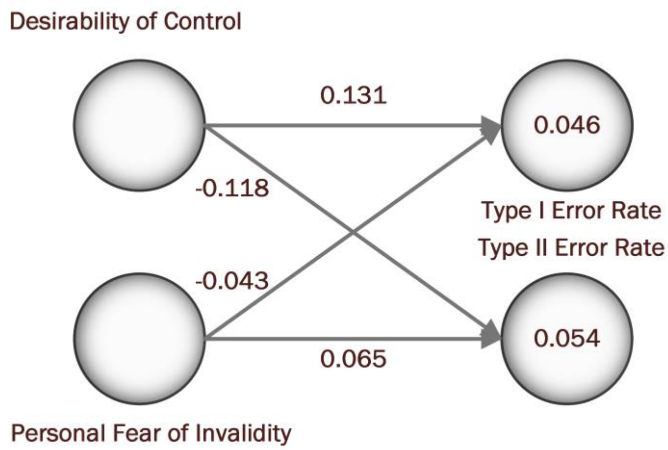

Figure 3. Structural Model with Path Coefficients and R^2

This figure reports the structural model with path coefficients to the respective error rates and corresponding R^2, 0.046, for type I error rate, 0.054 for type II error rate.

R^2 for type I and type II errors rates were 0.046 and 0.054, respectively, for each endogenous variable shown on the right. To evaluate effect size significance, F^2 was computed for both error types. The values were 0.007 and 0.002, respectively; however, those results indicated extremely weak relationships (typically a value of 0.02).

Appendix Table B5 lists the structural models' path coefficients and significance levels. Hypothesis H1, stating DC provides a positive path to type I error rate, was supported (mean = 0.131, standard deviation = 0.048, *p-value* = 0.007). Hypothesis H2, stating PFI provides a positive path to type II error rates, was not supported (mean = 0.065, standard deviation = 0.051, *p-value* = 0.198), was not supported.

To assess the significance of the predictive value of the model, Q^2 was calculated using a blindfolding procedure. As explicated by Hair et al. (2017), Q^2 "examines whether a model accurately predicts data not used in the estimation of model parameters" (p. 325). Type I and type II error rate predictive values were 0.039 and 0.050, respectively. The calculated Q^2 effect sizes were -0.011 and 0.012, respectively, and weak.

Model fit measures included goodness-of-fit (GoF), standardized root mean square residual (SRMR), and root mean square residual covariance (RMS_{theta}). Hair et al. (2017) did not recommend interpreting any of these measures as current research did not have solid results. However, studies pointed to RMS_{theta} as a promising metric, so, for completeness, its value was 0.20.

Hypothesis H3 predicted the percentage of true responses in the true-false assessment group would be significantly higher than the percentage true in a random distribution. A t-test of the percentage true rate of 0.5397 against 0.500 produced a *p-value* = 0.069 and effect size = 0.070. Running 10 simulations with sample sizes of 1,050 using an inverse binomial function yielded results of mean = 0.5010, standard deviation = 0.0037, skewness = 0.0071, and kurtosis = -1.4675. A t-test calculation resulted in minuscule changes from using 0.5000. Examining the confidence intervals with the upper limit yielded a non-significant *p-value* = 0.205 and effect size = 0.055. Using the lower limit produced a significant *p-value* = 0.023 and effect size = 0.099; however, the literature-based expectation was closer to 0.56 true. Therefore, Hypothesis H3 was not supported.

Hypothesis H4 predicted the percentages of correct responses would be significantly higher in the SPC assessment than in the true-false assessment because finance and accounting personnel would know and recognize non-randomness. From Table 1, it is clear respondents had incorrect strategies as there was no significant difference between the two assessments (*p-value* = 0.9062, effect size = 0.0051, and odds ratio = 1.0051). Therefore, H4 was not supported. This lack of random sequence knowledge in corporate finance and accounting personnel was disappointing but not surprising.

3.3 Modeling unobserved heterogeneity

Hair, Sarstedt, Ringleand Gudergan (2018) indicated traditional clustering methods such as k-means clustering did not meet expectations for PLS models. Instead, the authors recommended methods such as latent class techniques. This study employed finite mixture partial least squares (FIMIX-PLS) and prediction-oriented segmentation partial least squares (POS-PLS) to maximize explained variance.

Researcher expectation anticipated six groups based on respondent answer patterns and research of personality types (see Gerlach, Farb, Revelle, and Amaral, 2018). In Appendix Table B6, 11 separate fit indexes are listed across eight segments calculated in eight separate trials using FIMIX-PLS. The minimum values in AIC_4 and BIC suggested six segments. Also, the entropy statistic normed (EN) of 0.643, which was larger than the hurdle rate, provided additional confirmatory evidence.

Based on six segments, POS-PLS was executed. Appendix Table B7 lists path coefficients for each of the POS groups, and Table B8 indicates R^2 values. Segments two through six exhibited strong R^2. Once each respondent had been assigned to a group, a determination was made as to the characteristics of these groups. A review of quantitative data, including screening criteria and demographic variables, did not yield insights. However, several segments revealed a type of finance-accounting behavior based on relationships to error types which was much richer than anticipated.

For example, using Appendix Tables B7 and B8, segment six indicates either an individual with high DC or one with high PFI with strong positive tendencies to type I error rate and a strong negative tendency to type II error rate. These individuals could be described as “control freaks” or perhaps micromanagers or as fearful individuals who check everything multiple times. These activities would cause many type I errors through the effort to verify nonexistent effects. These persons could be individuals most subject to maladaptive behaviors described by Fenton-O’Creevy et al. (2003).

In opposition to what was expected, segment five individuals had stronger tendencies of DC to type II error rates and PFI to type I error rates. Again, controlling behavior results from the assumption of subject matter expertise, which may be exhibiting hubris or other IoCON behavior, leading them away from seeing any effects (type II). Conversely, semi-fearful finance individuals will be checking for effects constantly and continue making more type I errors, as in segment six.

Segment two individuals are similar to those anticipated in this study. These individuals have a weaker tendency of DC to type I error, but it remains significant. They control and make type I errors, much less than segment six. Moreover, this segment contains the most fearful individuals in the sample, the highest path of PFI to type II error rate, leading to indecision and IoCHA. Their remedy would be to follow Harris and Osman’s (2012) advice and “pretend” they had control despite their perception of chaos.

As an example of the strength of these results, POS-PLS groups were added to the original model. Appendix Table B10 shows path coefficients and *p-values* for segment six. All paths are significant. For the type I error rate, R^2 was 0.983, while the type II error rate was 0.986. The effect sizes of DC and PFI, f^2, were 39.0 and 0.032, respectively.

3.4 True-False PLS-SEM Model

Because of respondent guessing, another PLS structural model using the true-false assessment was possible. In an identical procedure, the true-false model was constructed. Variables retained included DC2, DC4, DC5, DC11, DC12, and DC15; PFI3, PFI4, PFI5, PFI6, PFI7, PFI11, PFI12, and PFI14. The equivalent FIMIX-PLS and POS-PLS processes were followed, which produced five segments. R^2 remained weak, yet some segments had high explained variance, though not as elevated as the SPC model. A crosstab count of respondents in both SPC and true-false assessments is shown in Appendix Table B10. While the initial reaction to this difference might appear problematic, Table 1 showed individuals reacted differently to each assessment, i.e., missing effects in SPC graphs and, perhaps, overthinking in a true-false guessing game. Moreover, this may provide evidence for Brunswik’s representative design (see Gigerenzer, 2000), discussed in the next section.

4. Discussion

The most interesting finding is accounting and finance personnel, when reviewing time series, do not distinguish between a random and non-random series. Using a time-series graph without slope, individual assessments were no better than chance. Finance personnel should have general knowledge about randomness, variation, and error implications along with their behavioral tendencies. Providing average point estimates should be superseded by simulations and confidence intervals incorporating extreme events. Moreover, in an organization with limited resources, allowing many type I errors will not necessarily decrease the probability of type II errors.

Second, these assessments revealed human flexibility in problem-solving. Table 1 showed a significant type II error rate in the SPC assessment, a professional task, and a significant type I error rate in the true-false assessment, a gaming task. The latter was designed to give typical random responses but, instead, was used by some to assign hidden intentions. Gigerenzer (2000) describes the last precept of Brunswik's probabilistic functionalism as representative design, the necessity to generalize beyond just one assessment of stimulus-response. These two inclinations reinforce caution when developing proxies and should be understood by analysts and decision-makers. The movement of respondents between groups, dependent on the measurement tool, is clearly exposed in the results.

Third, SEM-PLS, FIMIX-PLS, POS-PLS, and other modules offer new exploratory tools for complex relationships which can facilitate the representative design of endogenous variables. Their use uncovered a variety of groups and relationships between control, fear, and error rates. This combination of new statistical methods with older constructs coincides with the "tools to theories" heuristic advocated by Gigerenzer (1991, 2000).

Fourth, the overall model revealed minimal explained variance, so an objective instrument that identifies the universal tendencies toward IoCON or IoCHA amongst all individuals can continue.

Fifth, using Qualtrics' panels or research using current accounting and finance practitioners would continue to provide additional depth within an engaged scholarship approach.

Sixth, since much of the heuristics and bias research focuses on illusory behavior, i.e., observing non-existent effects. The SPC and true-false assessments allow identical measurement of both error types.

And finally, understanding variation in business processes that generate a variance in the financial statements creates knowledge towards an actual path of improvement. When executing their craft, finance and accounting practitioners should spearhead knowledge of error types and associated costs. Evoking Talab (2014) and Stimmler (2013) warned that both randomly-generated forecasts and enhanced complexity without new information increased financial risk-taking behaviors. This knowledge could provide a nuanced direction for finance professionals and may prevent the extinction of our craft as had been documented in other professions, notably in *The Specialist*, the saga of Lem Put, privy builder (Sale, 1929).

5. Conclusion

The importance of recognizing cognition bias cannot be understated. Definitive rules govern the SPC chart, which is based on the probability of occurrence. Individuals, who assume their positional authority, intuition, or experience could identify trends and shifts without experiments or science do so at their own peril. Moreover, realizing that financial assessments can be accurate or generate two types of errors, i.e., identifying something non-existent (type I) and not identifying something that exists (type II), are vital for the profession's future.

Additional finance and accounting research could focus on the eight individual non-random rules or the number of observations in time series strings. While this paper focused on finance and accounting personnel, other individuals might have different tendencies. Additional research could be undertaken with actual practitioners in other business functional areas and from many different industries.

Acknowledgments

We thank Christopher Posey, Lucy F. Ackert, Ph.D., Deborah I. Prosser, Ph. D., for providing invaluable input for the earlier version of the article.

References

Ackert, L. F., Church, B. K., & Qi, L. (2015). An Experimental Examination of Portfolio Choice. *Review of Finance, 20, July 2016*, 1427–1447. https://doi.org/10.1093/rof/rfv036.

Astrachan, C. A., Patel, V. K., & Wanzenried, G. (2014). A Comparative Study of CE-SEM and PLS-SEM for Theory Development in Family Firm Research. *Journal of Family Business Strategy, 5*, 116-128. https://doi.org/10.1016/j.jfbs.2013.12.002

Bennett, D. J. (1998). *Randomness.* Cambridge, MA: Harvard University Press. https://doi.org/10.4159/9780674020771

Blair, E. & Zinkhan, G. M. (2006). Nonresponse and Generalizability in Academic Research. *Journal of the Academy of Marketing Science, 34(1),* 4-7. https://doi.org/10.1177/0092070305283778

Blais, A-R., Thompson, M. M., & Baranski, J. V. (2005). Individual Differences in Decision Processing and

Confidence Judgments in Comparative Judgment Tasks: The Role of Cognitive Styles. *Personality and Individual Differences, 38*(7), 1701-1713. https://doi.org/10.1016/j.paid.2004.11.004

Buhl, E. R., Goodson, P., & Neilands, T. B. (2007). Structural Equation Modeling: A Primer for Health Behavior Researchers. *American Journal of Health Behavior, 31*(1), 74-85. https://doi.org/10.5993/AJHB.31.1.8

Burger, J. M., & Cooper, H. M. (1979). The Desirability of Control. *Motivation and Emotion, 3*(4), 381-393. https://doi.org/10.1007/BF00994052

Burton, R. (2005). Multiple-Choice and True-False Tests: Myths and Misapprehensions. *Assessment and Evaluation in Higher Education, 30,1,* 65-72. https://doi.org/10.1080/0260293042003243904

Cabeza, R., Rao, S. M., Wagner, A. D., Mayer, A. R., & Schacter, D. L. (2001). Can Medial Temporal Lobe Regions Distinguish True from False? An event-related functional MRI Study of Veridical and Illusory Recognition Memory. *Proceedings of the National Academy of Sciences, 98*(88), 4805-4810. https://doi.org/10.1073/pnas.081082698

Christensen, L. B., Johnson, R. B., & Turner, L. A. (2014). *Research Methods, Design, and Analysis.* Boston, MA: Pearson Education.

Clow, K. A., & Esses, V. M. (2005). The Development of Group Stereotypes from Descriptions of Group Members: An Individual Difference Approach. *Group Processes and Intergroup Relations, 8*(4), 429-445. https://doi.org/10.1177/1368430205056469

Deshmukh, A., Karim, K. E., Siegal, P. H. (1998). An Analysis of Efficiency and Effectiveness of Auditing to Detect Management Fraud: A Signal Detection Theory Approach. *International Journal of Auditing, 2*, 127-138. https://doi.org/10.1111/1099-1123.00035

Dimock, M. (2019). Defining generations: Where Millennials end and Generation Z begins. *Pew Research Center*. Retrieved from https://pewrsr.ch/2szqtJz

Fenton-O’Creevy, M., Nicholson, N., Soane, E., & Willman, P. (2003). Trading on Illusions: Unrealistic perceptions of Control and Trading Performance. *Journal of Occupational and Organizational Psychology, 76*, 53-68. https://doi.org/10.1348/096317903321208880

Fritz, M. (1921). Guessing in a True-False Test. *Journal of Educational Psychology, 18*(8), 558-561. https://doi.org/10.1037/h0074440

Gebhardt, W. A., & Brosschot, J. F. (2002). Desirability of control: psychometric properties and relationships with locus of control, personality, coping, and mental and somatic complaints in three Dutch samples. *European Journal of Personality, 16*, 423-438. https://doi.org/10.1002/per.463

Gerlach, M., Farb, B., Revelle, W., & Amaral, L. A. N. (2018). A Robust Data-driven Approach Identifies Four Personality Types Across Four Large Data Sets. *Nature of Human Behavior, 2*, 735–742. https://doi.org/10.1038/s41562-018-0419-z

Gigerenzer, G. (1991). From Tools to Theories: A Heuristic of Discover in Cognitive Psychology. *Psychological Review, 98*(2), 265-267. https://doi.org/10.1037/0033-295X.98.2.254

Gigerenzer, G. (2000). *Adaptive Thinking: Rationality in the Real World.* Oxford University Press. New York, New York.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM).* Thousand Oaks, CA: Sage Publications.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2018). *Advanced Issues in Partial Least Squares Structural Equation Modeling*. Thousand Oaks, CA: Sage Publishing, Inc. https://doi.org/10.1007/978-3-319-05542-8_15-1

Harris A. J. L., & Osman, M. (2012). The Illusion of Control: A Bayesian Perspective. *Synthesis, 189*, 29-38. https://doi.org/10.1007/s11229-012-0090-2

Henderson, G. R., Mead, G. E., van Dijke, M. L., Ramsay, R., McDowall, M. A., & Dennis, M. (2008). Use of Statistical Process Control Charts in Stroke Medicine to Determine if Clinical Evidence and Changes in Service Delivery Were Associate with Improvements in the Quality of Care. *Quality Safety Health Care, 17,* 301-306. https://doi.org/10.1136/qshc.2006.020784

Heuer, J., Merkel, C., & Weber, M. (2016). Fooled by Randomness: Investor Perception of Fund Manager Skill.

Review of Finance, 605-635. https://doi.org/10.1093/rof/rfw011

Kahneman, D., & Riepe, M. W. (1998). Aspects of Investor Psychology. *Journal of Portfolio Management, Summer 1998*, 52-65. https://doi.org/10.3905/jpm.1998.409643

Kinney, W. R., & Salamon, G. L. (1982). Regression Analysis in Auditing: A Comparison of Alternative Investigation Rules. *Journal of Accounting Research, 20*(2), 350-366. https://doi.org/10.2307/2490745

Klarner, P., Sarstedt, M., Hoeck, M., & Ringle, C.M. (2013). Disentangling the Effects of Team Competences, Team Adaptability, and Client Communication on the Performance of Management Consulting Teams. *Long Range Planning, 46*(3), 258-286. https://doi.org/10.1016/j.lrp.2013.03.001

Krueger, W. C. F. (1933). Distributions of Scores Based on Correct Guessing for True-false Tests of Various Lengths. *Journal of Educational Psychology 24*(3), 185-188. https://doi.org/10.1037/h0072930

McSweeney, B. (2006). Net Present Value: the Illusion of Certainty. *Strategic Change, 15*, 47-51. https://doi.org/10.1002/jsc.746

Morrison D. G. (1978). A Probability Model for Forced Binary Choices. *The American Statistician, 32*(1), 23-25. https://doi.org/10.1080/00031305.1978.10479238

Nelson, L. S. (1984). The Shewhart Control Chart—Tests for Special Causes. *Journal of Quality Technology, 16*(4), 237-239. https://doi.org/10.1080/00224065.1984.11978921

Poundstone, W. (2014). *Rock Breaks Scissors: A Practical Guide to Outguessing and Outwitting Almost Anybody.* New York, NY: Little, Brown and Company, Hachette Book Group.

Sale, C. (1929). *The Specialist.* St. Louis, MO: Specialist Publishing Company, Wainwright Building.

Rietzschel, E. F., De Dreu, C. K. W., & Nijstad, B. A. (2007). Personal Need for Structure and Creative Performance: The Moderating Influence of Fear of Invalidity. *Personality and Social Psychology Bulletin, 33*(6), 855–866. https://doi.org/10.1177/0146167207301017

Stimmler, Mary. (2013). Essays on Financial Risk Taking and Embedded Heuristics. UC Berkeley: Business Administration, Ph.D. Program. Retrieved from: http://www.escholarship.org/uc/item/01z2b0hj

Sullivan, G.M., & Feinn, R. (2012). Using Effect Size – or Why the *P-value* I not Enough, *Journal of Graduate Medical Education, September 2012*, 279-282. https://doi.org/10.4300/JGME-D-12-00156.1

Taleb, N. N. (2004). *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets.* New York, NY: Random House.

Taleb, N. N. (2014). *AntiFragile: Things that Gain from Disorder.* New York, NY: Random House.

Tellis, G. J. (2013). *Unrelenting Innovation. How to Build a Culture of Market Dominance.* San Francisco, CA: John Wiley and Sons.

Taylor, Keith I., Kiymaz, H. (n.d.). *Deming and Juran Were Correct: Verifying Estimates of the 85/15 Rule Using Two Statistical Process Control Chart Methods with Diverse Times Series*, [Unpublished Manuscript]. Rollins College, Winter Park, FL.

Thompson, M. M., Naccarato, M. E., Parker, K. C. H., & Moskowitz, G. (2001). The Personal Need for Structure (PNS) and Personal Fear of Invalidity (PFI) scales: Historical Perspectives, Present Applications and Future Directions. *In G. Moskowitz (Ed.), Cognitive Social Psychology: The Princeton Symposium on the Legacy and Future of Social Cognition*, 19-39. Mahwah, NJ: Erlbaum.

Uecker, W. C., & Kinney, Jr., W. R. (1977). Judgmental Evaluation of Sample Results: A Study of the Type and Severity of Error Made by Practicing CPAs. *Accounting, Organizations, and Society, 2*(3), 269-275. https://doi.org/10.1016/0361-3682(77)90016-2

Whitson, J., & Galinsky, A. (2008). Lacking Control Increases Illusory Pattern Perception. *Science, 322*(115), https://doi.org/10.1126/science.1159845

Zbilut, J. P. (2004). *Unstable Singularities and Randomness: Their Importance in the Complexity of Physical, Biological, and Social Sciences*. Elsevier. Amsterdam, The Netherlands.

APPENDIX A

Sample Survey Questions

Table A1. Final Statistical Process Control Chart (SPC) Question Sample List

Survey	**Answer**	**Error**	**Rule Violation; 0 = no Rule Violation**
SPC1	Random	Type1	0
SPC2	Non-Random	Type2	Test 6: 4 out of 5 points more than 1 standard deviation (StDev) from center line (CL) (same side)

Table A2. Final True-False Question Sample List

Survey	**Answer**	**Error**	**Actual Question**
TF1	FALSE	Type1	The Picture of Dorian Grey, chapter 4, 32th word is "with".
TF2	TRUE	Type2	The Picture of Dorian Grey, chapter 2, 14th word is "with".

APPENDIX B

Analytics of Results

Table B1. Outer loadings from Initial Model

Indicator	Char Type1	Char Type2	DC	PFI	SPCType2 Error Rate	SPCType1 Error Rate
DC1			-0.335			
DC10			0.614			
DC11			-0.448			
DC12			-0.570			
DC13			-0.231			
DC14			-0.143			
DC15			-0.534			
DC16			0.757			
DC17			-0.156			
DC18			-0.542			
DC19			0.696			
DC2			-0.597			
DC20			0.721			
DC3			-0.604			
DC4			-0.425			
DC5			-0.475			
DC6			-0.362			
DC7			0.695			
DC8			-0.118			
DC9			-0.160			
PFI1				-0.468		
PFI10				-0.502		
PFI11				0.572		
PFI12				0.608		
PFI13				-0.633		
PFI14				0.720		
PFI2				-0.376		
PFI3				0.650		
PFI4				0.701		
PFI5				0.726		
PFI6				0.726		
PFI7				0.580		
PFI8				0.459		
PFI9				-0.668		
SPCType1						1.000
SPCType2					1.000	

Note. Outer loadings greater than 0.7 exhibit convergent validity.

Table B2. Reliability Metrics for the Final Model

Measure	Cronbach's Alpha	Composite Reliability	Average Variance Extracted (AVE)
DC	0.852	0.894	0.629
PFI	0.861	0.896	0.589
SPCType2 Error Rate	1.000	1.000	1.000
SPCType1 Error Rate	1.000	1.000	1.000

Note. Cronbach's alpha should be a minimum value of 0.7.

Table B3. Fornell-Larker Criteria for Discriminant Validity for the Final Model

Measure	Char Type1	Char Type2	DC	PFI
DC	-0.378	0.412	**0.793**	
PFI	0.365	-0.371	-0.758	**0.768**

Note. DC and PFI have highest row values.

Table B4. Hetero-Trait Mono-Trait (HTMT) Criteria for Discriminant Validity for the Final Model

Measure	DC	PFI	SPCType2 Error Rate
PFI	0.874		
SPCType2 Error Rate	0.224	0.201	
SPCType1 Error Rate	0.211	0.183	0.690

Note. In a complete bootstrapping procedure, confidence intervals for single measures.

Table B5. Structural Model Assessment with relationships and P-values.

Relationship	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	*P-Values*
DC -> SPCType2 Error	**-0.118**	**-0.120**	**0.050**	**2.387**	**0.017**
DC -> SPCType1 Error	**0.131**	**0.130**	**0.048**	**2.711**	**0.007**
PFI -> SPCType2 Error	0.065	0.068	0.051	1.289	0.198
PFI -> SPCType1 Error	-0.043	-0.046	0.049	0.888	0.375

Note. Significant values are bolded.

Table B6. Fit Indexes for One- to Eight-Segment Solution (FIMIX-PLS)

Fit Index/Segment Number	1	2	3	4	5	6	7	8
AIC (Akaike's Information Criterion)	5,868	5,647	5,506	5,448	5,382	5,335	5,324	**5,305**
AIC[3] (Modified AIC with Factor 3)	5,876	5,664	5,532	5,483	5,426	5,388	5,386	**5,376**
AIC[4] (Modified AIC with Factor 4)	5,884	5,681	5,558	5,518	5,470	**5,441**	5,448	5,447
BIC (Bayesian Information Criteria)	5,907	5,731	5,635	5,621	5,600	**5,597**	5,631	5,657
CAIC (Consistent AIC)	5,915	5,748	5,661	5,656	**5,644**	5,650	5,693	5,728
HQ (Hannan Quinn Criterion)	5,883	5,679	5,555	5,514	5,465	**5,434**	5,441	5,439
MDL[5] (Minimum Description Length with Factor 5)	**6,130**	6,204	6,359	6,595	6,825	7,072	7,357	7,633
LnL (LogLikelihood)	(2,926)	(2,807)	(2,727)	(2,689)	(2,647)	(2,614)	(2,600)	(2,582)
EN (Entropy Statistic (Normed))		0.496	0.658	0.514	0.707	0.643	0.652	0.544
NFI (Non-Fuzzy Index)		0.533	0.643	0.461	0.649	0.555	0.547	0.406
NEC (Normalized Entropy Criterion)		529	359	510	308	375	365	479

Note. Row minimum is optimal and bolded. Preferred combinations: AIC_3/CAIC, AIC_3/BIC, AIC_4/BIC. EN should be larger than 0.50. This solution used AIC_4/BIC.

Table B7. Path Coefficients for Original Sample and Six POS Segments (POS-PLS)

Relationship	Original	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6
DC -> Type1	**0.131**	**0.217**	**0.202**	-0.015	-0.142	**-0.529**	**1.513**
DC -> Type2	**-0.118**	**-0.202**	**-0.107**	0.118	**0.218**	**0.503**	**-1.508**
PFI -> Type1	-0.043	**0.118**	**-0.833**	-0.313	-0.079	**0.454**	**0.716**
PFI -> Type2	0.065	-0.084	**0.909**	0.281	0.170	**-0.484**	**-0.686**
Group Size	1,050	749	80	38	61	54	68

Note. Significant values are bolded.

Table B8. R^2 Original, by POS Segment, and Weighted (POS-PLS)

R^2	Original	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Weighted
SPCType1 Error	0.046	0.049	0.946	0.894	0.962	0.948	0.983	0.308
SPCType2 Error	0.054	0.060	0.965	0.908	0.947	0.961	0.986	0.318
Group Size	1,050	749	80	38	61	54	68	1,050

Table B9. Path Relationships Original and POS Segments for Segment Six

Relationships	Original Sample (O)	Sample Mean (M)	Standard Deviation (STDEV)	T Statistics (|O/STDEV|)	*P-values*
DC -> SPCType1 Error	**1.514**	**1.512**	**0.082**	**18.480**	**0.000**
DC -> SPCType2 Error	**-1.508**	**-1.506**	**0.079**	**19.122**	**0.000**
PFI -> SPCType1 Error	**0.714**	**0.707**	**0.072**	**9.949**	**0.000**
PFI -> SPCType2 Error	**-0.686**	**-0.676**	**0.072**	**9.516**	**0.000**

Note. Significant values are bolded.

Table B10. Comparison of SPC and True-False Models Assignment of Respondents

Segments	TF1	TF2	TF3	TF4	TF5	Total
SPCGroup1	453	89	65	84	58	749
SPCGroup2	32	23	9	6	10	80
SPCGroup3	14	3	9	7	5	38
SPCGroup4	16	14	9	19	3	61
SPCGroup5	23	8	9	9	5	54
SPCGroup6	28	11	12	6	11	68
Total	566	148	113	131	92	1,050