7-17-2017

HRDQ Submissions of Quantitative Research Reports: Three Common Comments in Decision Letters and a Checklist

Kim Nimon
University of Texas at Tyler, knimon@uttyler.edu

HRDQ Submissions of Quantitative Research Reports: Three Common Comments in Decision Letters and a Checklist

Kim Nimon

I have been on the *Human Resource Development Quarterly* (*HRDQ*) editorial team since 2010, when I was asked to fill the position of Assistant Editor of Quantitative Methods vacated by Dr. Greg Wang when he became the editor of the *Journal of Chinese Human Resource Management*. Since that time, I have served as Associate Editor under the leadership of Dr. Andrea Ellinger, and I now serve as co-editor along with Drs. Valerie Anderson and Jon Werner. Over my tenure, I have reviewed hundreds of submissions to *HRDQ* and have attempted to address limitations I observed by contributing method-related editorials.

In 2011, I wrote an editorial that considered the quality of quantitative research reports (Nimon, 2011). The 2011 editorial considered common issues related to reports of quantitative research including statistical assumptions (i.e., independence of observations, reliability of data), data analysis (i.e., measurement level, ecological validity, informed interpretation), and results (i.e., statistical and practical significance). Four years later, in Nimon and Astakhova (2015), we found that of the 63 quantitative articles reviewed, 100% included the reporting of an effect size or statistics that could be used to compute effect sizes. Only 5% of the applicable articles reviewed made no mention of reliability, which was in "stark contrast to Vacha-Haasee and Thompson (2011) who found that across 47 reliability generalization meta-analysis studies which represented 12,994 primary reports, 54.6% did not mention reliability" (Nimon & Astakhova, 2015, p. 234). While the editorial from 2015 shows promising results for *HRDQ* authors' ability to report quantitative research in a rigorous manner that allow for subsequent meta-analyses, I often find myself commenting on the same issues when I review manuscripts submitted to *HRDQ*, many of which are not identified by reviewers.

In this editorial, I build on my prior editorials and elaborate on three issues that I frequently note in decision letters to authors that report on

Human Resource Development Quarterly, vol. 28, no. 3, Fall 2017
Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/hrdq.21290

quantitative research. The purpose of this work is to assist authors in preparing manuscripts for potential publication in *HRDQ*, as well as to provide a resource to authors who may receive a decision letter noting such an issue. The three primary topics to be addressed in this editorial are:

1. Discrepancies between stated hypotheses and analyses.
2. Issues with mediated designs.
3. Harman's single-factor test.

Because this editorial is necessarily not comprehensive in considering the breadth and depth of quantitative method–related comments that could be included in a decision letter, I also provide a general checklist that authors may want to consider when preparing submissions to *HRDQ*. This will be presented after the three issues just mentioned are addressed.

Discrepancies Between Stated Hypotheses and Analyses

Authors frequently state hypotheses that consider a relationship between variables (e.g., positive affect is positively related to employee engagement). While such a hypothesis may be valid and can be tested with either a zero-order or implied correlation (e.g., see Zigarmi, Nimon, Houson, Witt, & Diehl, 2011), the problem occurs when the hypothesis is tested by a standardized regression weight from ordinary least squares regression or a standardized path coefficient from a path or structural equation model (SEM) where the regression, path, or SEM model contains additional paths to the same variable and those paths stem from correlated variables.

Consider for example the correlations among positive affect, work cognition, and employee engagement reported in Zigarmi et al. (2011, Table 2). As depicted in Panel A of Figure 1, the standardized path coefficient between positive affect and employee engagement is 0.77. Panel A indicates: (a) the correlation between positive affect and employee engagement is 0.77, and (b) for each standard deviation (SD) change in positive affect, employee engagement increases 0.77 of an SD. In Panel B of Figure 1, the standardized path between positive affect and employee engagement is 0.70. Panel B of Figure 1 *does not* indicate that the correlation between positive affect and employee engagement is 0.70. The panel indicates that for each *SD* change in positive affect, employee engagement increases 0.70 of an SD, holding work cognition constant. In this example, the weight that positive affect had on employee engagement decreased with the inclusion of work cognition in the model.

Now imagine that a measure of social desirability was included in Zigarmi et al. (2011) and the correlation to positive affect was 0.0 and the correlation to employee engagement was 0.1. As depicted in Panels A and B of Figure 2, the standardized path between positive affect and employee

Figure 1. Example Models Where a Standardized Path Coefficient Is Not a Measure of Relationship

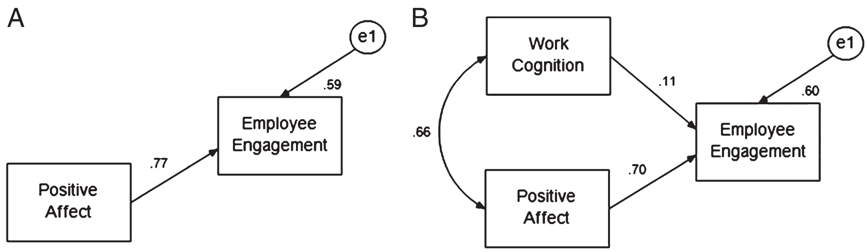
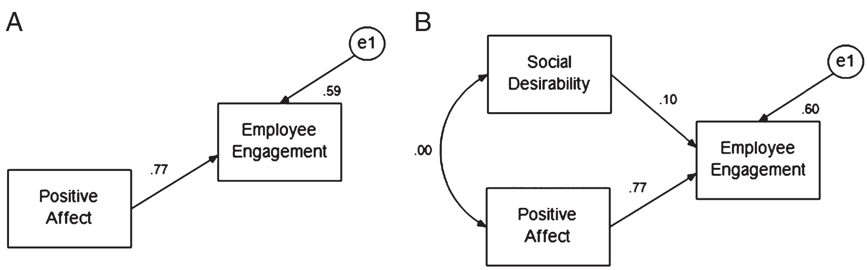


Figure 2. Example Models Where a Standardized Path Coefficient *Is* a Measure of Relationship



engagement remains at 0.77. Panels A and B indicate that (a) the correlation between positive affect and employee engagement is 0.77 and (b) for each SD change in positive affect, employee engagement increases 0.77 of a SD. In this hypothetical example, the weight that positive affect had on employee engagement was not changed with the inclusion of social desirability because social desirability had no relationship with positive affect.

It should seem clear that the level of correlation among predictors impacts whether a standardized regression weight or path coefficient can be interpreted as a measure of relationship (cf. Courville & Thompson, 2001). While some authors may calculate a variance inflation factor (VIF) as a technique to assess the "degree of multi-collinearity of the *i*th independent variable with the other independent variables in a regression model" (O'Brien, 2007, p. 673) and consider it appropriate to interpret standardized regression weights or path coefficients as measures of relationship if VIFs are low, I argue below that this not appropriate.

Consider for example measures from the High School and Beyond data set (Holzinger & Swineford, 1939), where 26 tests were administered to 301 students. Among those measures were scores from *numerical puzzles*, *add*, and *deduction* tests. When considering a regression model where *deduction*

Table 1. Comparison of Baron and Kenny's (1986) Conditions for Mediation

Condition	*Baron and Kenny (1986, p. 1176) Conditions for Mediation*	*Baron and Kenny (1986, p. 1177) Regression Equations and Conditions for Mediation*	*Saks (2006, p. 612) Conditions for Mediation Referencing Baron and Kenny (1986)*
First	Variations in levels of the independent variable significantly account for variations in the presumed mediator (i.e., Path *a*).	When regressing the mediator on the independent variable [first equation], the independent variable must affect the mediator.	The independent variables(s) (the antecedents of engagement) must be related to the mediator (employee engagement).
Second	Variations in the mediator significantly account for variations in the dependent variable (i.e., Path *b*).	When regressing the dependent variable on the independent variable [second equation], the independent variable must be shown to affect the dependent variable.	The mediator (employee engagement) must be related to the dependent variables(s) (the consequence of engagement).
Third	When paths *a* and *b* are controlled, a previously significant relationship between the independent and dependent variables is no longer significant, with the strongest demonstration of mediation occurring when Path *c* is zero.	When regressing the dependent variable on both the independent variable and on the mediator [third equation], the mediation must affect the dependent variable in the third equation. If these conditions all hold in the predicted direction, then the effect of the independent variable on the dependent variable must be less in the third equation than in the second. Perfect mediation holds if the independent variable has no effect when the mediator is controlled.	A significant relationship between the independent variable(s) (antecedents of engagement) and a dependent variable(s) (consequences of engagement) will be reduced (partial mediation) or no longer be significant (full mediator) when controlling for the mediator (employee engagement).

test scores were regressed on *add* and *numerical puzzles* test scores, the VIF for both predictors is 1.141, which is considered low and not indicative of a multicollinearity problem (cf. von Eye & Schuster, 1998).

As depicted in Panel A of Figure 3, the correlation between *numerical puzzles* test and *deduction* test scores is 0.04. While the relationship between *numerical puzzles* tests scores and *deduction* tests scores is positive (albeit small), the standardized path coefficient flips signs when *numerical puzzles* test scores is entered in the regression equation (see Panel B of Figure 3). Also note that the standardized path coefficient for *numerical puzzles* (i.e., 0.44) is higher than the correlation between *numerical puzzles* and *add* test scores (i.e., 0.35). What this suggests is the variable *add* test scores is suppressing variance in the variable *numerical puzzles* test scores that is irrelevant to predicting *deduction* test scores, thus making *numerical puzzles* tests scores a stronger predictor than it would be on its own. This finding indeed makes sense when you realize that the *add* test is a timed test and therefore the scores reflect something in addition to the ability to add. One can only imagine the confusion that would have entailed had the standardized path coefficient been interpreted as a negative measure of relationship. In fact, recognizing the standardized path coefficient has a different sign than the correlation coefficient contributed important knowledge to the understanding of the relationship among the study variables. For another such example, see Siebold and McPhee (1979).

To clarify, correlation coefficients are bounded by −1.0 and 1.0. However, standardized regression weights and path coefficients do not have such bounds. As such, it is confusing to reference paths as measures of relationship except in the case where there is only one predictor or where predictors are perfectly uncorrelated (cf. Courville & Thompson, 2001). I expect that what authors may intend to hypothesize are direct effects, rather than relationships. Note that a standardized direct effect estimates the amount of change in a dependent variable *Y* as the proportion of an SD, given a change in an independent variable *X* of a full SD, controlling for other parents of *Y* (Kline, 2016, p. 232). It is, of course, *controlling for other parents of Y* that may distinguish a standardized direct effect from a bivariate correlation or relationship. As in the example previously presented, authors should interpret standardized weights in conjunction with structure coefficients or bivariate correlations in the presence of correlated predictors (cf. Courville & Thompson, 2001).

Figure 3. Example Models Where a Standardized Path Coefficient and Correlation Coefficients Are Different Signs

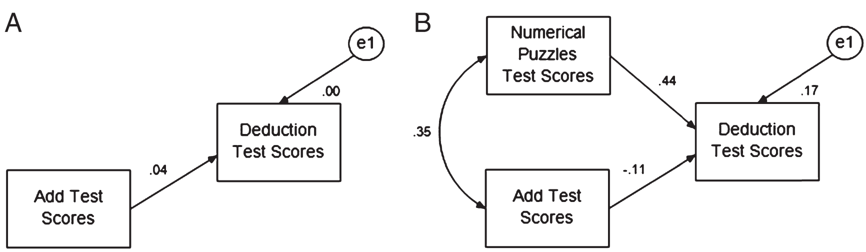
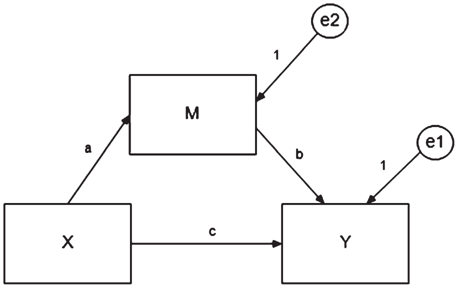


Figure 4. Three-Variable Mediation Model



Mediated Designs

I expect that one of the many reasons authors may hypothesize bivariate relationships is a reliance on the Baron and Kenny (1986) model of mediation that in some cases has been misreported. Consider, for example, the conditions for mediation proposed by Baron and Kenny referencing paths, as depicted in Figure 4, and how Saks (2006) described the conditions to establish mediation. As seen in Table 1, Saks referenced relationships for the first, second, and third conditions for mediation. Also note that there is a substantive difference between Baron and Kenny's second condition and Saks' interpretation. In Saks, the second condition only considered the relationship between the mediator and the dependent variable. However, Baron and Kenny (1986, p. 1176) indicated a second condition for mediation as "variations in the mediator significantly account for variations in the dependent variable (i.e., Path *b*)," which requires regressing the dependent variable on both the independent and on the mediator (see Figure 4 and Baron and Kenny's third regression equation in Table 1). This is further clarified by Kenny (2016).

> Use *Y* as the criterion variable in a regression equation and *X* and *M* as predictors (estimate and test path *b*). It is not sufficient just to correlate the mediator with the outcome because the mediator and the outcome may be correlated because they are both caused by the causal variable *X*. Thus, the causal variable must be controlled in establishing the effect of the mediator on the outcome.

Whether Baron and Kenny's (1986) model of mediation or related articles contribute to authors stating hypotheses that are inconsistent with their analyses is somewhat irrelevant to a couple more issues I want to discuss related to studies that report on "mediation." First, I agree with Kline (2016), who suggested that "the term mediation should be reserved for designs that feature time precedence" (p. 135). This indicates that for authors who report on cross-sectional data, the term *mediation* should not be used. Rather than referencing a mediating variable, authors can reference an intervening variable. In the model depicted in Figure 4, rather than indicating that variable *M* mediates the relationship between *X* and *Y*, authors could indicate that *X* has an indirect effect or association on *Y,* controlling for the direct effect or association of *X* on *Y* as appropriate (cf. Kline, 2016, p. 232). Note that the total effect of *X* on *Y* is estimated "controlling for other variables that sever all back-door (noncausal) paths between *X* and *Y*, leaving only direct or indirect causal paths between them" (Kline, 2016, p. 232). As such, in designs that are more complex than the three-variable mediation design depicted in Figure 4, indirect effects must be interpreted considering appropriate controls including variables that sever noncausal paths between *X* and *Y* and other mediators. Note that for the remainder of the editorial, the term *mediated* will be used broadly to include those designs that may or may not qualify as mediated to be consistent with statistical literature that does not make a distinction based on research design.

The second point related to mediated designs is the need for authors to indicate the model informing their research. While Baron and Kenny's (1986) model has historically informed mediated designs published in *HRDQ*, more recent literature suggests a more relaxed set of conditions to inform mediation. For example, Zhao, Lynch, and Chen (2010), referencing a three-variable mediation model, as depicted in Figure 4, argued "there should be only one requirement to establish mediation, that the indirect effect a x b be significant" (p. 198). Published corollaries of only considering the indirect effect include: "The strength of the mediation should be measured by the size of the indirect effect, not by the lack of the direct effect" (Zhao et al., 2010, p. 198). "The *X-Y* test is never relevant to establishing mediation" (Zhao et al., p. 200). An indirect effect may be claimed even if the direct path from the mediator to the outcome (e.g., Path *b* in Figure 4) is not statistically significant (Preacher & Hayes, 2008, p. 31).

My purpose in presenting this alternative model of mediation is not to advocate for one over the other.[1] Rather, it is provided as means to illustrate the importance of identifying the underling model used when testing mediation. As illustrated by Zhao et al. (2010), interpretation of the same data may yield different results depending on the conditions for mediation employed. For other models of mediation, see, for example, Preacher and Hayes (2008).

The last point I will address regarding mediated designs relates to the statistical significance of indirect effects. When testing the statistical significance of indirect effects, current literature suggests that bootstrapping is a preferred technique over the Sobel test. The Sobel test may be inaccurate because it makes assumptions that are usually untenable. Kline (2016) also cautioned against making "hair-splitting distinctions among *p* values from significance tests for indirect effects"—even from bootstrapped tests—"especially if the sample size is not large." He recommended that researchers "rely more on whether the magnitudes of indirect effects are substantively meaningful, given the research context" (p. 465).

Harman's Single-Factor Test

In *HRDQ* submissions, the most common technique that authors appear to consider as a statistical means to assess common method variance (CMV) is Harman's single-factor test. Podsakoff, MacKenzie, Lee, and Podsakoff (2003) indicated that despite its apparent appeal, Harman's single-factor test is insensitive to detecting CMV. Podsakoff et al. (2003) identified more robust statistical techniques to assess and control for method effects that I will not repeat here. However, a technique that I often recommend authors to consider is the CFA marker variable technique described by Williams, Hartman, and Cavazotte (2010).

In the CFA marker technique, a marker variable is chosen that "(a) is influenced by the same causes of CMV (e.g., affectivity, acquiescence) as a set of substantive variables, and (b) is not theoretically related to those substantive variables" (Simmering, Fuller, Richardson, Ocal, & Atinc, 2015, p. 474). Five CFA models with the marker and substantive variables are tested. The CFA model is a traditional measurement model including the marker variable and all substantive variables. In the Baseline model, the correlation from each substantive variable to the marker variable is set to 0 and the measurement parameters of the marker variable are set to the values from the Baseline model. The Method-C model builds on the Baseline model by fixing method factor loadings to be equal. The Method-U model builds on the Baseline model by allowing method factor loadings to be freely estimated. As indicated by Williams et al. (2010),

> A comparison of the Method-C to the Baseline Model provides a test of the presence of equal effects associated with the marker latent variable. A comparison of the Method-C and Method-U models allows for a comparison of the CMV and UMV models discussed by Lindell and Whitney. … Method-R Model is identical to the Method-C and Method-U model, only the substantive factor correlations are constrained to their values from the Baseline Model. The comparison of the Method-R Model with either the Method-C or Method U Models (depending on which is retained in their direct comparison) provides a test of bias in the substantive factor correlations due to the marker-based variance that may be present (p. 494).

In my own research (e.g., Shuck, Nimon, & Zigarmi, in press), I have found the CFA marker variable technique to be informative. Therefore, I often recommend that authors review Richardson, Simmering, and Sturman

(2009), as it reviews the CFA marker technique along with other techniques that assess CMV and bias.

A General Checklist for Reports of Quantitative Research

Table 2 contains a checklist for quantitative research reports that I use when writing decision letters to authors who submit manuscripts to *HRDQ*. As I expect the checklist to continue to evolve over my tenure as co-editor, the most current version of the checklist can be accessed at profnimon.com/HRDQxList.pdf. Note that Item 1 has already been elaborated in this editorial, so I will say no more, other than I think it is incumbent on authors to be experts on the methods that they report on and to be sure that their hypotheses match their analytic strategy.

Method

Authors should fully describe their samples (Item 2). According to the American Psychological Association (APA, 2009), "human samples should be fully described with respect to gender, age, and, when relevant to the study, race or ethnicity. Where appropriate, additional information should be presented (generation, linguistic background, socioeconomic status, national origin, sexual orientation, special interest group membership, etc.)" (p. 4). Although much research reported in *HRDQ* is conducted on convenience samples, it would be helpful to know how the sample demographics compare to the intended population (cf. Kline, 2008, p. 68).

I have discussed the importance of testing and reporting how data meet the statistical assumptions associated with the data analysis reported (Issue 3) in a prior editorial (i.e., Nimon, 2011). In addition to that editorial, authors may find helpful the special issue that Osborne (2013) edited, as well as texts that are specific to their data analytic strategy (e.g., Kline, 2016, for SEM). The big picture is that failure to meet statistical assumptions may impact the reliability and validity of the statistics reported. I expect authors to identify the statistical assumptions for their analyses, report on how the data did or did not meet them, and address the subsequent data analytic strategy accordingly. For example, multivariate normality is considered by some (e.g., Kline, 2016) as a statistical assumption for confirmatory factor analysis (CFA) and structural equation modeling (SEM). However, if the data are not multivariate normal, authors may need to report bootstrapped estimates.

As mentioned previously, *HRDQ* submissions often report on cross-sectional data (Issue 4). As such, there is a concern that the data may be subject to CMV that could bias results. While there is disagreement in the field as to whether common method bias inflates common method correlations (Conway

Table 2. A Checklist for Quantitative Research Reports

Section of Manuscript / *Issue*	*Relevant References*
Introduction	
1. Hypotheses consistent with analyses	Statistical textbooks (e.g., Kline, 2016)
Method	
2. Sample description	APA (2009)
3. Statistical assumptions	Osborne (2013) Zientek, Nimon, & Brown (2016)
4. Procedural remedies for common method bias	Podsakoff et al. (2003)
Results	
5. Descriptive statistics	APA (2006)
6. Construct validity	Graham et al. (2003) Henson & Roberts (2006) Schreiber et al. (2006)
7. Statistical assessment of common method variance and bias	Richardson et al. (2009) Simmering et al. (2015) Williams et al. (2010)
8. Test statistics, *df*s, *p* values, effect sizes, and indications of uncertainty (e.g., *SE*s or CIs) as well as sufficient statistics to verify *df*s and *p* values and to support replication studies	APA (2006) Callahan & Reio (2006) Epskamp & Nuijten (2015) Henson (2006)
9. Tests of regression models	Nimon & Oswald (2013)
10. Tests of canonical models	Nimon et al. (2011)
11. Tests of SEM models	Cortina et al. (2016) Kline (2016) Schreiber et al. (2006)
12. Tests of indirect effects	Wen & Fan (2015) Zhao et al. (2010)
13. Nested models	van Mierlo et al. (2009) West et al. (2007)
14. Tests of measurement invariance	Vandenberg & Lance (2000)
15. Instrument development	Hinkin (1998) Worthington & Whittaker (2006)
Final Checks	
16. Reliability not attributed to instruments	Thompson & Vacha-Haase (2000)
17. No claims of causality without appropriate design	Gubbins & Rousseau (2015)
18. Errors in writing	Onwuegbuzie et al. (2010).

Note. The checklist is an initial version. The most current version can be found at profnimon.com/HRDQxList.pdf.

& Lance, 2010), most researchers would agree that the procedural remedies offered by Podsakoff, MacKenzie, Lee, & Podsakoff (2003), including "temporal, proximal, psychological, or methodological separating of measurement," "protecting respondents anonymity and reducing evaluation apprehension," and "counterbalancing question order," should be considered a priori techniques to reduce the likelihood of CMV (pp. 887–888).

Results

In quantitative research articles, correlations between study variables may be reported in a correlation matrix (Item 5). The inclusion of a correlation matrix is consistent with American Educational Research Association (AERA, 2006) standards, which call for matrix summaries to be included in research reports. Combined with sample size, Ms, SDs, and measures of reliability, a correlation matrix allows researchers to conduct analyses as if they had access to the original dataset. Alternatively, a covariance matrix may be reported because it provides the same information as a correlation matrix and a set of SDs.

The American Statistical Association (2007) recommended that "for every measure in every research process it is essential to provide appropriately defensible evidence for the validity, reliability, and fairness of [scores on] the measure" (p. 11). Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) are common techniques for examining factor validity. Results from CFA can also be used to assess reliability, convergent validity, and discriminant validity. Henson and Roberts (2006) provided excellent guidance for how to report EFA analyses, and Schreiber, Nora, Stage, Barlow, and King (2006) described how to report CFA analyses. Also note that Graham, Guthrie, and Thompson (2003) demonstrated the importance of reporting and interpreting structure coefficients in addition to path coefficients when reporting on CFA analyses. Note, however, EFA and CFA results may not provide a robust set of information that describes how well scores from a measure relate to scores from other established measures in a predictable pattern (i.e., nomological validity). Therefore, validity studies are often conducted before an instrument is considered sufficiently vetted to be used in substantive research. There is also the concern of common method variance and bias when data have been selected using the same technique. As discussed previously, the Harman single-factor test is known to be highly conservative in detecting CMV. Therefore, I recommend that authors consider more robust techniques. Richardson, Simmering, and Sturman (2009) provided an excellent review of statistical techniques for the detection of and correction for CMV.

The AERA (2006) recommended that "for each of the statistical results that is critical to the logic of the design and analysis, there should be included": "the test statistic," "the associated significance level," "an effect size of some kind," and "an indication of the uncertainly of that index of effect (such as a standard error or confidence interval)" (p. 36). As it relates to significance levels, the APA (2010) recommended that "exact probabilities to two or three decimal places" be reported (p. 139). Degrees of freedom should also be reported so that significance levels can be independently confirmed (cf. Epskamp & Nuijten, 2015). When reporting regression and path coefficients, authors often report only standardized weights as the magnitude of weights can be compared. However, unstandardized weights are actually better for replication studies and "when the scales of all variables are meaningful rather than arbitrary" (Kline, 2016, p. 29).

Most of the analyses reported in quantitative research reports submitted to *HRDQ* are part of the general linear model (GLM). The GLM encompasses a set of analyses that "(a) are correlational, (b) yield effect sizes analogous to r^2, and (c) apply weights to a measured variable to yield scores on latent variables" (Thompson, 2006, p. 360). Because these analyses are part of a single analytic family, these procedures are hierarchical, in that some procedures are special cases of others. Regression encompasses most univariate GLM analyses, canonical correlation analyses encompasses univariate and multivarite GLM analyses, and SEM subsumes all GLM analyses. As such, the checklist in Table 2 considers substantive models using the GLM hierarchy.

Dependent on the type of substantive model tested (Issues 9–13), additional information needs to be reported. Nimon and Oswald (2013) presented a comprehensive set of statistics to report for regression models, along with R syntax to compute the statistics presented. Nimon, Henson, and Gates (2010) presented guidelines for presenting canonical models along with SPSS and R syntax. Schreiber et al. (2006) reviewed statistics to report for SEM analyses. In addition, authors should consider the work of Cortina, Green, Keeler, and Vandenberg (2016) and report sufficient information so that the *df*s for SEM models can be independently confirmed. Also note that I concur with the best practices in SEM presented by Kline (2016, pp. 452–468), which identifies the need for researchers to test models informed by theory, test alternative models, and "never retain a model based solely on global fit testing" (p. 461). For SEM models with indirect effects, authors should consider Zhao et al. (2010) as well as Wen and Fan (2015) for guidance on reporting indirect effects.

The GLM analyses also encompass data that violate the assumption of independent observations and are clustered in some way (e.g., employees in teams, employees cross-classified into departments and divisions). Although authors who have clustered data may report a low intraclass correlation coefficient (ICC) to indicate that the data analytic strategy does not need to take into account the clustered nature of the data, Roberts (2002) argued that the absence of a significant ICC does not indicate that the assumption of independence has not been violated. More appropriately, such data should be analyzed with multilevel models (Issue 13).

When authors report on multilevel models, oftentimes I find they consider a random intercept, but do not test for random slopes. I find this problematic as efficiently modeling random slopes is a key feature of multilevel models and without such analyses authors "run the risk of reporting findings that are opposite what they would be if the data were analyzed with the appropriate technique" (Nimon, 2011, p. 389). Although there are many textbooks to choose from when seeking guidance on reporting multilevel models, I find West, Welch, and Galecki (2007) very informative. The journal also occasionally receives submissions where multilevel data have been aggregated. Although Osborne (2000) demonstrated problems with the aggregation strategy for a set of data, I recognize there may be times when such a data analytic strategy is consistent with the aims of the study. In such cases, I advise authors to follow the recommendations of van Mierlo, Vermunt, and Rutte (2009) when presenting their analyses.

Occasionally, *HRDQ* receives submissions that are validity studies including those that report on a new instrument or test for measurement invariance (Issues 14–15). As with substantive studies, authors need to present the framework informing their validity studies. Just as there are many ways to assess measurement invariance (see Vandenberg & Lance, 2000), there are many models to follow when presenting findings from a new instrument (e.g., Hinkin, 1998; Worthington & Whittaker, 2006).

Final Checks

Before submitting a manuscript to *HRDQ,* authors should conduct some final checks. I present three checks that are frequent recommendations in decision letters.

First, be careful not to attribute reliability to an instrument. As discussed in a prior editorial (Nimon, 2011), I agree with Thompon and Vacha-Haase (2000) that reliability is a property that can be attributed to data, but not the instrument (Issue 16). For example, employees in an individualistic society may not respond to items that are designed to measure perceptions of autonomy the same as those in a collectivistic society. If the reliability of data is a problem, the related variable may have to be omitted from subsequent substantive analyses. In no case does it make sense to indicate that even though a reliability estimate is low in a given study, the data is considered sufficiently reliable because another author reported high reliability when using the same scale.

Second, if reporting on cross-sectional data (Issue 17), verify that no claims of causality or prediction have been made. Prediction generally requires a longitudinal design and claims of causality require an experimental or quasi-experimental design (Gubbins & Rousseau, 2015).

Finally, check that the manuscript contains no errors. Confirm the statistics reported in tables match what are described in the narrative (Issue 18). I understand how such discrepancies occur, but it is confusing to reviewers and the editorial team when such errors are present, as it is often difficult to determine which statistic is correct. I also recommend that authors have a professional editor review their manuscripts prior to submissions. Often, reviewers have a challenging time providing a meaningful review when the manuscript has grammatical and APA errors. For guidelines for avoiding the most common APA errors, see Onwuegbuzie, Combs, Slate, and Frels (2010).

Concluding Thoughts

The field of statistics is an area of research and, as such, what is considered good practice changes over time. For example, while most current doctoral students learn that multivariate analysis of variance (MANOVA) should not be followed up with univariate analyses, more mature researchers may not have learned the benefits of performing multivariate group comparisons following a statistically significant MANOVA (cf. Enders, 2003). The field is also informed by simulation studies that give guidance as to the reliability of statistics under certain data conditions (e.g., Richardson et al., 2009). Finally, advances in statistical software are constantly changing what is considered good practice for reporting on quantitative research (cf. Zientek & Thompson, 2009).

This suggests at least two final recommendations for authors contributing quantitative research to *HRDQ*. First, know your analyses. I often tell my doctoral students that a fool with a tool is still a fool. With current statistical software, very advanced analyses can be accomplished with just a click of a button. However, much more knowledge is required to interpret those results in an accurate and meaningful manner. There are several ways to keep up with advances in quantitative methods. International conferences such as the *Academy of Management (AOM)* and the *Society for Industrial and Organizational Psychology (SIOP)* often include method-related workshops. Statistical camps (e.g., Stats Camp) provide an opportunity for researchers to stay abridged of the latest quantitative development and work with leaders in the field. Journals such as *Organizational Research Methods*, *Behavior Research Methods*, *Psychological Methods*, and *Multivariate Behavioral Research* regularly publish research that considers advancements in statistical science.

Second, prospective authors should consider conducting secondary data analyses using published literature. As stated by Zientek and Thompson (2009), "the inclusion of correlation/covariance matrices, standard deviations, and means can enhance findings … by permitting secondary researchers to (a) conduct commonly utilized traditional univariate and multivariate analyses not initially performed in primary studies, (b) produce effect sizes and other statistics not included in prior published literature, and (c) conduct analyses once difficult to perform" (p. 343).

This editorial has emphasized three issues that frequently arise when quantitative research is reviewed for *HRDQ,* namely, concerns over discrepancies between stated hypotheses and the analyses conducted, issues with mediated designs, and the proper use of Harman's single-factor test. In addition, a general checklist has been provided which should assist authors as they prepare their manuscripts for submission. May we all make better use of the quantitative research tools available to us to advance the field of human resource development through research.

Acknowledgments

I want to thank the *Human Resource Development Quarterly* editorial team including Drs. Anderson, Brown, Gubbins, Reio, Sheehan, Werner, and Yoon, as well as Ms. Mandolen Mull for their review and input to this editorial.

Note

1. I find the decision tree to conceptualize types of mediations and the interpretation of the data pattern for conclusions about theory very intuitive (see Zhao et al., 2010, Figure 2b).

References

American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher, 35*(6), 33–40. https://doi.org/10.3102/0013189X035006033

American Psychological Association (APA). (2009). *Publication manual of the American Psychological Association: Supplemental material: Writing clearly and concisely*, Chap. 3. Retrieved from http://supp.apa.org/style/pubman-ch03.00.pdf

American Psychological Association (APA). (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.

American Statistical Association. (2007). *Using statistics effectively in mathematics education research*. Alexandria, VA: Author.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Callahan, J. L., & Reio, T. (2006). Making subjective judgments in quantitative studies: The importance of using effect sizes and confidence intervals. *Human Resource Development Quarterly, 17,* 159–174. https://doi.org/10.1002/hrdq.1167

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology, 25,* 325–334. https://doi.org/10.1007/s10869-010-9181-6

Cortina, J. M., Green, J. P., Keeler, K. R., & Vandenberg, R. J. (2017). Degrees of freedom in SEM: Are we testing the models that we claim to test? *Organizational Research Methods*, *20*, 350–378. https://doi.org/10.1177/1094428116676345

Courville, T., & Thompson, B. (2001). Use of structure coefficients in published multiple regression articles: Beta is not enough. *Educational and Psychological Measurement, 61,* 229–248. https://doi.org/10.1177/0013164401612006

Enders, C. K. (2003). Performing multivariate group comparisons following a statistically significant MANOVA. *Measurement and Evaluation in Counseling and Development, 36,* 40–57.

Epskamp, S., & Nuijten, M. B. (2015). statcheck: Extract statistics from articles and recompute *p* values (R package version 1.0.1.).

Graham, J. M., Guthrie, A. C., & Thompson, B. (2003). Consequences of not interpreting structure coefficients in published CFA research: A reminder. *Structural Equation Modeling, 10*(1), 142–153. https://doi.org/10.1207/S15328007SEM1001_7

Gubbins, C., & Rousseau, D. M. (2015). Embracing translational HRD research for evidence-based management: Let's talk about how to bridge the research–practice gap. *Human Resource Development Quarterly, 26,* 109–125. https://doi.org/10.1002/hrdq.21214

Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist, 34,* 601–629. https://doi.org/10.1177/0011000005283558

Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement, 66,* 393–416. https://doi.org/10.1177/0013164405282485

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1,* 104–121. https://doi.org/10.1177/109442819800100106

Holzinger, K. J., & Swineford, F. (1939). *A study in factor analysis: The stability of a bi-factor solution (No. 48)*. Chicago, IL: University of Chicago Press.

Kenny, D. A. (2016). The four steps. Retrieved from http://davidakenny.net/cm/mediation.htm.

Kline, R. B. (2008). *Becoming a behavioral science researcher: A guide to producing research that matters*. New York, NY: Guilford Press.

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Nimon, K. (2011). Improving the quality of quantitative research reports. *Human Resource Development Quarterly, 22,* 387–394. https://doi.org/10.1002/hrdq.20091

Nimon, K. F., & Astakhova, M. (2015). Improving the rigor of quantitative HRD research: Four recommendations in support of the general hierarchy of evidence. *Human Resource Development Quarterly, 26,* 231–247. https://doi.org/10.1002/hrdq.21219

Nimon, K., Henson, R., & Gates, M. (2010). Revisiting interpretation of canonical correlation analysis: A tutorial and demonstration of canonical commonality analysis. *Multivariate Behavioral Research, 45,* 702–724. https://doi.org/10.1080/00273171.2010.498293

Nimon, K., Oswald, F. L. (2013). Understanding the results of multiple linear regression: Beyond standardized regression coefficients. *Organizational Research Methods, 16,* 650–674. https://doi.org/10.1177/1094428113493929

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41,* 673–690. https://doi.org/10.1007/s11135-006-9018-6

Onwuegbuzie, A. J., Combs, J. P., Slate, J. R., & Frels, R. K. (2010). Editorial: Evidence based guidelines for avoiding the most common APA errors in journal article submissions. *Research in the Schools*, *16*(2), ix–xxxvi.

Osborne, J. (2000). Advantages of hierarchical linear modeling. *Practical Assessment, Research & Evaluation*, *7*(1). Retrieved from http://pareonline.net/getvn.asp?v=7&n=1

Osborne, J. (Ed.). (2013). Sweating the small stuff: Does data cleaning and testing of assumptions matter in the 21st century? *Frontiers in Psychology,* 1–158. https://doi.org/10.3389/978-2-88919-155-0

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88,* 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Preacher, K. J., & Hayes, A. F. (2008). Contemporary approaches to assessing mediation in communication research. In A. F. Hayes, M. D. Slater, & L. B. Snyder (Eds.), *The Sage sourcebook of advanced data analysis methods for communication research* (pp. 13–54). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781452272054.n2

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods, 12,* 762–800. https://doi.org/10.1177/1094428109332834

Roberts, J. K. (2002). The importance of the intraclass correlation in multilevel and hierarchical linear modeling designs. *Multiple Linear Regression Viewpoints, 28,* 19–31.

Saks, A. M. (2006). Antecedents and consequences of employee engagement. *Journal of Managerial Psychology, 21,* 600–619. https://doi.org/10.1108/02683940610690169

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research, 99,* 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Shuck, B., Nimon, K., & Zigarmi, D. (in press). Untangling the predictive nomological validity of employee engagement: Decomposing variance in employee engagement using job attitude measures. *Group & Organization Management.*. https://doi.org/10.1177/1059601116642364

Siebold, D. R., & McPhee, R. D. (1979) Commonality analysis: A method for decomposing explained variance in multiple regression analysis. *Human Communication Research, 5,* 355–363. https://doi.org/10.1111/j.1468-2958.1979.tb00649.x

Simmering, M. J., Fuller, C. M., Richardson, H. A., Ocal, Y., & Atinc, G. M. (2015). Marker variable choice, reporting, and interpretation in the detection of common method variance: A review and demonstration. *Organizational Research Methods, 18*(3), 473–511. https://doi.org/10.1177/1094428114560023

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach.* New York, NY: Guilford Press.

Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement, 60,* 174–195. https://doi.org/10.1177/0013164400602002

Vacha-Haase, T., & Thompson, B. (2011). Score reliability: A retrospective look back at 12 years of reliability generalization studies. *Measurement and Evaluation in Counseling and Development, 44,* 159–168. https://doi.org/10.1177/0748175611409845

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4–70. https://doi.org/10.1177/109442810031002

van Mierlo, H., Vermunt, J. K., & Rutte, C. G. (2009). Composing group-level constructs from individual-level survey data. *Organizational Research Methods, 12,* 368–392. https://doi.org/10.1177/1094428107309322

von Eye, A., & Schuster, C. (1998). *Regression analysis for social sciences*. San Diego, CA: Academic Press.

Wen, Z., & Fan, X. (2015). Monotonicity of effect sizes: Questioning kappa-squared as mediation effect size measure. *Psychological Methods, 20,* 193. https://doi.org/10.1037/met0000029

West, B. T., Welch, K. B., Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software.* Boca Raton, FL: Chapman & Hall. https://doi.org/10.1201/b17198

Williams, L. J., Hartman, N., & Cavazotte, F. (2010). Method variance and marker variables: A review and comprehensive CFA marker technique. *Organizational Research Methods, 13,* 477–514. https://doi.org/10.1177/1094428110366036

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *Counseling Psychologist, 34,* 806–838. https://doi.org/10.1177/0011000006288127

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37,* 197–206. https://doi.org/10.1086/651257

Zientek, L. R., Nimon, K., & Brown, B. (2016). Analyzing data from a pretest-posttest control group design: The importance of statistical assumptions. *European Journal of Training and Development, 40,* 638–659. https://doi.org/10.1108/EJTD-08-2015-0066

Zientek, L. R., & Thompson, B. (2009). Matrix summaries improve research reports: Secondary analyses using published literature. *Educational Researcher, 38,* 343–352. https://doi.org/10.3102/0013189X09339056

Zigarmi, D., Nimon, K., Houson, D., Witt, D., & Diehl, J. (2011). A preliminary field test of an employee work passion model. *Human Resource Development Quarterly, 22,* 195–221. https://doi.org/10.1002/hrdq.20076

Kim Nimon is Associate Professor at the University of Texas at Tyler.

Corresponding author:
Kim Nimon can be contacted at knimon@uttyler.edu