Applying an Interactive Quantitative-Qualitative Framework: How Identifying Common Intent Can Enhance Inquiry

John H. Hitchcock¹ and Isadore Newman²

Abstract
This article offers a conceptual discussion of a series of qualitative and quantitative research endeavors that share common underpinnings and purpose. The article demonstrates several commonalities between approaches with the hope of encouraging readers to allow research questions to guide methodological choices and avoid overemphasizing division between research paradigms. The current work offers an extension of ideas previously published in Human Resource Development Review (Newman & Hitchcock, 2011) to address a call to explore aspects of theory-building in the context of research methods (Reio, 2010). Specific ideas discussed here focus on commonalities between quantitative and qualitative work when dealing with the broad notion of generalization, and include:

- connections between transferability, probabilistic generalization, naturalistic generalization, and external validity;
- commonalities between multiple (or collective) case studies and meta-analyses; and
- phenomenological perspectives and probability.

Keywords
mixed methods research, research methodologies, quantitative research, qualitative research

¹Ohio University, Athens, OH, USA
²Florida International University, Miami, FL, USA

Corresponding Author:
John H. Hitchcock, Ohio University, 305B McCracken Hall, Athens, OH 45701, USA
Email: hitchcoc@ohio.edu
Underlying Agreements Between Qualitative and Quantitative Research: Part II

*Human Resource Development Review* recently issued a call (Reio, 2010) to promote theory building that can in turn inform methodological choice. There has been at least one response to this call (Newman & Hitchcock, 2011). That piece demonstrated commonalities between quantitative and qualitative work to encourage readers to allow research questions to guide their methodological decisions; furthermore, the authors stated that there can still be a tendency to overemphasize division between research paradigms despite the considerable growth in mixed-method publication outlets that has occurred in recent years. The basic concept is that good research is good research, regardless of paradigm, and there are underlying methodological constructs that need to be considered regardless of orientation toward quantitative, qualitative, or mixed methods work. The purpose of this article is to further articulate the idea that there are common techniques and goals across quantitative and qualitative research approaches and that recognition of such commonalities can enhance empirical inquiry in research settings. In addition, the article presents an interactive continuum adopted from Ridenour and Newman (2008) that offers examples of how this can be broadly applied to questions that are of interest to human resource professionals.

Rather than focusing on a given paradigm, and perhaps reaffirming what we believe to be a false dichotomy, researchers can conceptualize an interactive framework between so-called qualitative and quantitative approaches. The overall philosophical stance we adopt is articulated in Newman and Hitchcock (2011) and Newman and Benz (1998). We offer our perspective and rationale that research strategies typically aligned with qualitative or quantitative camps are often more similar than different. This perspective has implications for training researchers. Note that we do not mean the training of quantitative or qualitative researchers, but individuals who can utilize multiple perspectives to address complex questions about social phenomena. The operating theoretical framework for the current work is the interactive continuum presented by Newman and Benz (1998) and Ridenour and Newman (2008). This framework can be summarized by Ridenour and Newman's (2008) definition of scientific research, where they state that

- There should be consistency between the research question, purpose, and methodological choices.
- Research needs to be systematic, verifiable, and transparent and it must demonstrate trustworthiness.
- Research must be replicable.
- Research should provide opportunity for self-correction.
- Research should be able to explain the phenomena under investigation.

The interactive continuum facilitates research by providing a framework to promote these principles. This idea is summarized in Figure 1.
To explain the framework, we offer an example that could occur in a human resource development (HRD) setting where empirical evidence might be used to inform practice. Consider an investigation of the connection between employee motivation and compensation practices within a given industry. An operating idea may be that different bonus structures, and associated communication around decisions to provide bonuses, are thought to have differential influences on motivation. This might be empirically understood at an aggregate (nomothetic) level; experimental studies may show that, on average, different bonus schemes yield different levels of motivation. Of course, variance estimates and sample sizes will be necessary components for conducting a null hypothesis test of statistical significance. Recommended practice would be to estimate the magnitude of difference between groups and to think in terms of confidence.
intervals in lieu of focusing solely on p values (Thompson 2002a, 2002b; Wilkinson and the Task Force for Statistical Inference, 1999). The interactive continuum framework would encourage researchers to go further with investigating the connection between bonus structure and motivation by using multiple tools. The interconnections between different methods and how the bonus structure might be understood at different levels can almost start to seem endless, even in the context of a simple two-group study. The framework can, however, promote focus by the ever-present push to not lose sight of the purpose of any given question (see the top of the figure) and a culminating goal of continuously revising and confirming theory (bottom of the figure).

To elaborate, assume a study used a single “treatment” (the treatment is a new bonus structure) and control group (prior compensation practices). Further assume the intervention yielded an effect size showing that one can expect about a one standard deviation improvement on a measure of motivation if it is in place (i.e., effect size = 1.0). Although this would be an important finding and might start to inform compensation policy, more work can be pursued. Examination of variance will often show that some members of the treated group still, despite exposure to a powerful new approach, have less motivation than study participants in the control condition. The interactive continuum encourages examination as to why. In doing so, one could use a variant of discriminant function analyses to identify employees who responded well to the treatment from ones who did not (Schumacker & Newman, 2012). But it would help to have some qualitative understanding of participant characteristics to proceed with such an analysis. The framework implies that making choices around what variables to consider when developing a model requires an understanding of context. Factors one might consider could be employee cultural characteristics, satisfaction, perhaps evolving expectations of job performance, and so on. One could imagine a complex relationship between such characteristics and perceptions of a new bonus scheme. This is likely to be heavily dependent on context, cultural concerns, and individual employee’s history in an organization, perceptions of the message tied to a bonus, and so on. It is the case that qualitative methods would offer necessary techniques to understand the experience of study participants under the two different bonus structures. Both typical employees and outliers may offer interesting grounds for conducting case studies (which might also be done with multiple methods such as surveys, interviews, reviews of documents like employee performance history). Phenomenological techniques could be used to understand individual perceptions of what it is like to work in the company and focus on the merit of a bonus structure. Or as another example, ethnographic techniques might offer another lens focusing on the study of the culture in participating organizations in an attempt to learn how shared values, perceptions, and so on, might have impacted the relationship between bonuses and motivation. In sum, those who wish to understand (a) aggregate impacts of the bonus system on motivation, (b) impacts of the system on different subgroups, and (c) perceptions of the system across subgroups will likely want to use multiple research methods. There can be real value in further inquiry. Getting back to the example of a bonus system that yields an effect size of 1.0
on employee motivation, we may learn that some workers perceive the new structure to be challenging (they think other systems make it easier to achieve a bonus), or perhaps think a new approach favors certain subgroups (i.e., the system is not fair). or perhaps perceive a new approach to be demeaning (teachers may react poorly to the idea that they would teach better if paid more) and so on. The continuum, along with an appreciation of how disparate methods often address different facets of the same underlying questions, encourages use of multiple techniques.

The connections provided by the framework are not unique. There is a large body of work that connects different paradigms, types of methods, and philosophy (e.g., Creswell & Plano Clark, 2011; Johnson & Onwuegbuzie, 2004; Onwuegbuzie & Johnson, 2006; Tashakkori & Creswell, 2007; Tashakkori & Teddlie, 2003; 2011). Furthermore, others have listed criteria that need to be considered when using multiple methods in the context of scientific research (e.g., Johnson, 2009; see also Nastasi, Hitchcock & Brown, 2010, for an overview of several different frameworks). Our purpose is, however, not to rehash discussions of mixed methods work, but to apply the interactive continuum to continue to point out that (a) the walk from qualitative to quantitative perspectives is shorter than what is often portrayed in the literature, and (b) it can be used as a basis to further understand conceptual commonalities between paradigms. There are a number of similarities between the different paradigms when attending to basic principles of research and these similarities are captured by the framework. To our knowledge there has been a lack of specific discussion of how different topics in design can be informed by the different perspectives, or combination of perspectives, and we believe awareness of these connections can further enhance design. Newman and Hitchcock (2011) attempted to bridge this gap by discussing

- qualitative perspectives when interpreting statistical significance and replicability of findings;
- practical significance and the need to understand context;
- how data saturation is connected to generalization and sample-size concerns;
- quantitative perspectives of triangulation and negative case analyses;
- the similarity between reliability and external auditors;
- thematic analyses and factor development; and
- use of qualitative techniques to better understand participant selection in quasi-experiments (p. 384).

The work relied on logical analyses of different techniques; the conclusion was that so-called qualitative and quantitative approaches have more in common than what is often suggested in the literature. When considering exploratory factor (quantitative) and thematic analyses (qualitative) for example, both share the underlying goal of reducing data into manageable chunks of information (factors or themes). For that matter, once thematic analyses are completed some researchers prefer to check for whether external auditors will independently identify the same themes drawn from raw data and this is
highly consistent with notions of reliability and interobserver agreement. Yet another commonality here is that themes are often subjected to external review (consider peer-debriefing) to understand subjective interpretation, and factors to be understood in light of theory and could be subject to experts to examine content validity. Our point here is that these two fundamental techniques hold a great deal in common.

Although the above list provides a brief introduction of how quantitative, qualitative, and mixed methods research are perhaps not as different as often depicted in the literature, this represents only a beginning. The current work extends the discussion by identifying and describing additional research topics that can be informed by quantitative and qualitative perspectives. Here we focus on issues of generalization because of apparent commonalities that were not covered in Newman and Hitchcock (2011). Specific points discussed here are

- connections between transferability, probabilistic generalization, naturalistic generalization, and external validity;
- commonalities between multiple (or collective) case studies and meta-analyses;
- phenomenological perspectives and probability.

Connections Between Transferability, Generalization, and External Validity

In the book, *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century* (2001), David Salsburg discusses the poor fit between logic and statistically based decisions (i.e., decisions based on probability). In terms of making decisions on the basis of statistical models and probability, he states,

... decisions based on probabilistic arguments are not logical decisions. Logic and probabilistic arguments are not compatible ... In logic, there is a clear difference between a proposition that is true and one that is false. But probability introduces the idea that some propositions are probably or almost true. That little bit of resulting unsureness blocks our ability to apply the cold exactness of material implication in dealing with cause and effect. One of the solutions proposed for this problem in clinical research is to look upon each clinical study as providing some information about the effect of a given treatment. The value of that information can be determined by a statistical analysis of the study but also by the quality of the study. This extra measure, the quality of the study, is used to determine which studies will dominate the conclusions. The concept of the quality of a study is a vague one and not easily calculated. The paradox remains, eating at the heart of statistical methods. Will this worm of inconsistency require a new revolution in the twenty-first century? (pp. 300-301)
Salsburg offers this writing after demonstrating that continued use of null hypotheses can yield a patently wrong conclusion. He offers an example where 10,000 lottery tickets for a prize drawing are sold, one of which is a winner. The use of hypothesis testing would yield a conclusion that no ticket will win the lottery (each has a 0.0001 chance), but we know one ticket must be a winner. So basing decisions on probability alone can yield incorrect decisions. Salsburg wonders if this problem might be figured out via a revolution within statistics, and this field must yield a solution if we are to rely only on hard calculation. But some qualitative concepts might be useful here. It may be a stretch to say that the "worm" Salsburg writes about can be addressed with these concepts, but perhaps connecting ideas about generalization, transferability, naturalistic generalization and external validity, can help researchers remember to use probability as a tool, while not solely relying on it, when making inferences about their observations.

With that background, consider the connections between statistical generalization (e.g., Heeringa, West, & Berglund, 2010), transferability (Denzin & Lincoln, 2005; Hellström, 2008), naturalistic generalization (Stake, 1997), and external validity (Shadish, Cook, & Campbell, 2002). Statistical generalization deals with classic inferential statistics and probability (i.e., inferring knowledge of a population parameter based on what was observed in a sample), and thus the degree to which a sample is representative of some underlying population. By contrast, transferability deals with the degree to which findings from a study can be applied to new settings. When invoking statistical generalization, the burden of proof is very much on the researcher and several steps must be described. These include particulars of the target population, how a sample was drawn from the population, and estimates of the degree of confidence the researcher has that whatever has been observed in the sample reflects the population of interest. When invoking transferability (a qualitative notion), the burden is shifted more toward the reader of a study. Researchers must describe the particulars of the study in sufficient detail (often referred to as thick description) so as to allow readers to judge the degree to which findings are relevant to their own needs (Brantlinger, Jimenez, Klingner, Pugach, & Richardson, 2005; Denzin & Lincoln, 2005; Hellström, 2008). Another qualitative notion, naturalistic generalization, deals more with making generalizations on the basis of logic, where one considers similarities across contexts, people, and so forth. Finally, external validity is drawn from experimental design literature and considers the degree to which one can infer a causal relationship holds across different people, settings, context, and measures and so on. Shadish et al. (2002) describe this aspect of their experimental validity typology in terms of factors that can threat external validity.

Getting back to probability statements, consider a census, which attempts to measure an entire population without relying on sampling. This of course tends to require massive resources (think of the federal census) and so most of the time researchers work with a sample, which forms the basic rationale for inferential statistics. To be able to make reasonable inferences, it is important that the sample is similar to the population of interest. If the sample is randomly drawn from the target population and
there is no reason to worry much about issues like coverage error (i.e., when not all members of the target have a chance of being in a sample), then this "similarity" condition is met (Dillman, Smyth, & Christian, 2009). From there, estimation of sampling error provides a sense of how confident we can be that whatever is observed in the collected sample accurately reflects a population parameter. Of course, obtaining a random sample in many studies is easier said than done and other sources of error routinely enter into the endeavor (e.g., nonresponse bias). So it behooves researchers to have a fallback plan and consider the degree to which characteristics of an observed sample likely reflect population characteristics; that is, we should at least rely on logic when it is hard to invoke probabilistic generalization. This strikes us as similar to Stake's (1997) ideas around naturalistic generalization, which suggests that generalization is dependent on the degree of similarity from one setting to another. Such logical generalization appears to be common. Researchers often conduct so-called quantitative studies (surveys, experiments, etc.) without the benefit of random sampling, yet probability statements are used anyway. In a strict sense, this makes a strong assumption that the sample is representative of a given population and it may thus be a real problem to rely on null hypotheses to rule out chance as the reason for any observed relationships in a given sample. This does not necessarily mean that hypothesis tests should not be used, but why not also apply ideas like naturalistic generalization at the same time? If naturalistic generalization relies on elements of logic, as it must because the idea relies on similarity (in this case, similarity between the sample and population of interest), the two approaches can be used in a complementary fashion. To clarify, mechanistic applications of null hypothesis tests when randomization was not used will likely introduce biased estimates (some readers may prefer to use "error" in lieu of "bias"), so we argue that theory and logic must be more clearly applied to assess the degree to which such biases are likely to undermine the capacity to draw solid conclusions. In short, we advocate for explicit use and discussion of naturalistic generalization when probabilistic generalization may be a concern. We also recommend that, when feasible, the application of inferential statistics be paired with ideas like thick description so as to promote transferability.

What if a study can reasonably argue that the full population of interest was available to the researcher? That is, there is no sample because data is collected from everyone in the target population. This might happen in a census, but could also occur in a case study where a small but unusual group represents a bounded phenomenon of interest. If a researcher is dealing with a small number of people and it makes sense to conceptualize them as representing the totality of some specialized population, then the mathematics of null hypothesis tests of statistical significance may not be particularly useful. There is not a probability statement to make. But in this scenario, some form of inference to other times and settings will still often be of interest. To demonstrate the point about time as a consideration in external validity, consider for example a study that found unambiguous evidence that a video-game that encourages children to engage in arithmetic work can promote related academic achievement; but the study was conducted in the early 1990s and claimed that part of the reason
why it works is due to a novelty effect. That is, a key element of the treatment is children perceive they are playing a game, and in essence are entertained, while working on arithmetic problems. Given advancement in technology and the wider use of such games, it is reasonable to question the external validity of the study to contemporary circumstances. Because researchers may wish to generalize findings across time, or for that matter different outcome measures or even populations, we see that naturalistic generalization and transferability should be invoked, even when a researcher might argue that the full population of interest was included in a study. To accommodate these scenarios, researchers should be aware how each lens can be used in a complimentary fashion.

Another angle here is that purposive sampling, maximum variation sampling, and so on, which tend to align with qualitative work (Collins, 2010; Onwuegbuzie & Leech, 2007; Patton, 2002) connects to Shadish et al.'s (2002) discussion of external validity when designing a study (not just when making inferences based on results). Their focus is on the degree to which findings from an experimental design might be valid in other contexts, times, to different populations, subpopulations, outcome measures, and so on. This represents a much broader concern than a focus of whether observations from a sample were likely to be reflective of some population effect. From a design point of view, a priori recognition of interest in subpopulations can be dealt with by recruiting a sample that captures the characteristics that are necessary to support generalization to these groups. In a simple example, if one is interested in examining treatment effects for both males and females, then it behooves the researcher to have a sampling scheme that will allow for inclusion of both sexes to test for related interaction effects.

In sum, we argue that commonality across these ideas allows for applying all of these concepts at once. Most research endeavors entail a basic desire to question if, and to what extent, observations made during the course of a study could be true in some other context. Those who espouse a quantitative paradigm recognize the idea of external validity and qualitative researchers recognize the notions of transferability and naturalistic generalization. Consistent with a focus on the commonalities between qualitative and quantitative work, we see an advantage to remaining open to all of these ideas during the design and conduct of a study and when communicating findings. It is laudable, if not necessary, to pursue some assurance that whatever is observed in a study is reflective of a target population and to assess the degree to which probabilistic generalization is problematic. But even in those cases where one can have a high degree of confidence that a sample is well reflective of a population, one still must contend with the difference between probabilistic and logical arguments (Salsburg's point described above). It is also the case that researchers should want to think through broader issues of external validity; if, after all, a study is well justified in saying that it obtained a well-reflective sample of a given population, what concerns arise when thinking about findings with similar but different populations? Or what about questioning the degree to which findings transfer across time, new settings, or other measures? Our point is that even when dealing with those rare cases when one can argue that probabilistic
generalization is reasonable, researchers might consider naturalistic generalization, transferability, and other concerns in external validity. Models like the interactive continuum framework are designed to promote such thinking.

Now for a moment consider these ideas in the context of systematic inquiry in a HRD setting. Revisiting the above example regarding bonus systems, it is logical to assume that the same system simply cannot be expected to have a universal impact when considering variation in industry, culture, language, and so forth. Indeed, it is not hard to imagine how a bonus structure deemed to be useful in one setting may be problematic in other settings. As a simple example, just imagine how an academic professor might react to a bonus structure that encourages quick graduation of PhD students. So return to the basic desire to question if, and to what extent, observations made during the course of a study could be true in some other context. To do so one must think beyond probabilistic generalization (i.e., the degree to which a sample is representative of a target population) and instead deal with transferability and external validity. This is because program impacts are being pondered across what are essentially different populations. Therefore, we advocate using both quantitative and qualitative approaches. Readers who agree with this point may start to wonder if research reports should routinely discuss transferability, external validity, and other aspects of generalization. This is a point which is revisited below, but first we offer a different example of how commonalities between disparate research techniques can promote a more comprehensive understanding of findings.

Commonalities Between Multiple (or Collective) Case Studies and Meta-Analyses

Case studies entail the in-depth examination of bounded phenomena (Stake 1995, 2006; Yin, 2009). We struggle with whether it is best to think of the case study approach as a distinct methodology or as a research goal. Case studies can after all be examined with qualitative, quantitative, or mixed methods approaches (Stake, 2005) and rely on several combinations of methods to understand the case (Yin, 2009). Regardless of how one conceptualizes the individual case study approach, there are some who argue that collective, or multiple, case studies do represent a distinct methodology (having said that, Yin has suggested that case study skill sets should transfer to collective case study work). Stake (2006) outlines a series of recommendations when conducting these designs along with a justification for their use. The key idea is the multicase study provides an opportunity to compare and contrast among cases, which provides both an opportunity for understanding a phenomenon in different contexts, and this can provide a better sense of specific cases as well. In a sense, understanding of the individual is enhanced by seeing the whole (i.e., several cases) via contrasting and comparing findings. At the multicase level, findings are to be synthesized into themes that are better understood by triangulating findings across cases. This promotes a sense of the complexities of each case as well as any generalizations that appear to be reasonable within the set of cases.
To maintain both a sense of the individual cases as well as aggregate findings, Stake recommends the use of the cross-case dialectic, where localized (i.e., within case) findings are compared to overall findings. Stake suggests findings can be adversarial, or rather, in direct conflict; but looking for these differences should provide a better sense of the phenomena of interest. In this sense, singular generalization is often not so much the goal as it is a desire to see if a phenomenon behaves in particular ways in particular contexts. This notion is, by the way, consistent with the qualitative notion of transferability (Branlinger et al., 2005; Nastasi & Schensul, 2005). Nevertheless, cross-case thematic analyses promote an examination of that which is common across available cases.

This type of movement from the individual to the aggregate is also seen in classic meta-analytic work. Meta-analysis is a procedure where multiple studies of a given phenomenon are systematically collected and analyzed to see what aggregated findings indicate (Hedges, & Olkin, 1985; Lipsey & Wilson, 2001; McNeil & Newman, 1994). The Greek word root “meta” (above) provides a sense of purpose. Classic meta-analytic steps consider whether there is reason to believe observed effects might generalize beyond the available corpus of studies, and this is somewhat akin to considering the degree to which multiple-case study findings can be expected to transfer to other scenarios. In addition, meta-analysts will often examine if a set of effect sizes appear to vary considerably, and if so, to see if such variation can be explained by particulars of individual studies. This is conceptually similar to the same tension Stake describes when contrasting cross-case with individual case study findings. Overall, what the two types of studies indicate is that there are conceptual similarities between multicase studies and meta-analyses, in that both types of approaches seek to synthesize (when they can), compare and contrast findings across several studies to develop a better understanding of some phenomenon of interest.

Getting back to HRD concerns, consider an effort to document examples of effective managerial job performance. Cases of different managerial successes (and failures) might be blended with meta-analytic thinking that promotes an understanding of general principles managers might follow. That is to say, meta-analytic work might collate and summarize a series of empirical studies about practices of effective leaders, and this could be supplemented with detailed descriptions of managerial behavior and impacts on staff development, motivation, and so forth. Indeed, combining both sets of techniques can yield iterative information on this particular topic.

Phenomenological Perspectives, Context and Probability

On another topic, probability statements (i.e., $p$ values) represent a fundamental aspect of quantitative work. Several researchers have strongly criticized the rote application of $p$ values when making decisions about null hypotheses (see Cohen, 1994; Kline, 2004; Morgan, 2003; Schmidt & Hunter, 1997; Thompson, 2002a, 2002b; Wilkinson and Task Force for Statistical Inference, 1999). One general recommendation for those who criticize rote application null hypothesis tests of statistical
significance is the need to think through the meaning of a $p$ value. We suggest that one such approach could be to ponder the contextual meaning of a probability statement, and recommend this be considered both from the emic point of view of researchers and the etic perspective of researchers; that is, the respective points of view of research participants and outsiders (Nastasi & Schensul, 2005). In the latter case, there may be times when strict adherence to controlling for Type I error is less critical, such as when working within the context of an efficacy study where an intervention is being implemented under optimal circumstances and the primary goal of the study is to focus on implementation. Simply put, a researcher might be more interested in getting initial evidence of an exploratory program and less concerned with fully controlling for Type I error rates. When presenting probability statements to research participants by contrast, it is necessary to understand the value they may construe about the relative chances of a Type I error. To offer a blatant example, the required level of evidence when making decisions about interventions that deal with life and death outcomes (one might wish for a lower $p$ value cutoff) can be expected to be quite different from, say, alterations to marketing plans to improve sales (here users of research might tolerate higher levels of Type I error). So there can be value in determining how participants may react to and perceive different types of intervention scenarios since different research contexts require different levels of evidence. This contextual information should in turn influence decisions around $p$ value cutoffs, whether to apply a one- or two-tailed test, and so on.

A summation of the issue then is that it may be all well and good if some finding is statistically significant, but what does this finding mean to the individual? In many respects this connects to prior comments on practical significance and a call for examination of context to understand the importance of rejecting a null hypothesis (Newman & Hitchcock, 2011). But we wonder if the idea can be extended somewhat by drawing on some phenomenological frames by understanding how an individual experiences, say, a treatment. Adopting such a lens may start to open the door toward understanding idiographic and explanatory causation (Johnson & Christensen, 2011; see also Shadish et al., 2002) where one can examine individualized perspectives of causal mechanisms as well whether such mechanism hold up in unique/rare settings. To elaborate a little, we may learn that some treatment is almost universally effective, but we might still wish to know about those rare circumstances when it does not appear to work. In short, applying how an individual experiences and perceives a treatment can promote localized and rare applications. Although phenomenology as a discipline does not appear in the interactive continuum, it easily fits within section C (examples of qualitative research) and following this idea to its natural conclusion would be to suggest that intervention studies can benefit from targeted attempts to understand the experiences of a few participants. This way if one concludes that a treatment “works” on the basis of a $p$ value, there can be descriptive information wrapped around the finding to examine additional questions like how the treatment worked and, if it did not work in individualized cases, then why?
Again let us consider a human resources example. Consider the impact of an employee wellness program. One might design a study to see if introduction of such a program has, in aggregate, a positive impact on any number of employee outcomes in areas such as health, motivation, and so forth. Typical experimental and quasi-experimental designs will rely on some probabilistic statement that deals with whether we can expect that findings observed in the sample are likely replicated in some population of employees. This is all well and good but the continuum encourages efforts to explore treatment impacts at an individual level that will help to identify people it does not work for, therefore improving our understanding of exceptions to the general finding and, furthermore, to gain a deeper understanding of how the program works and not just if it worked. The implications of this statement is that research that is based on group analyses has some obvious advantages over research that only looks at individual differences, such as single subject designs. However single subject designs have certain advantages over group analyses in that they can identify how an individual subject is specifically responding to a treatment. We are suggesting that mixed methods allows one to combine the strengths of multiple methods, for a better understanding of both group and individual outcomes. It appears that medical research is also moving in this direction, where they are not only interested in if a drug generally works for the majority of the population but also in identifying who it works for and for whom it does not.

So What?

Peer review of manuscripts often helps to improve presentation. An earlier version of this work was of course reviewed and we were encouraged to clarify the “So what?” question behind these points. As an answer, we remind readers that Thomas Reio (2010) issued a call for discussing methodological developments in HRD. Specifically, he stated,

We cannot overemphasize the importance of the need for HRD scholars to contribute to the development of theory-building research methods in the social sciences. Without an understanding of how and why such methods support theory building and how to apply them productively, we risk stagnating as a field. (p. 223)

He later specifies mixed methods research as one area where such work is needed.

In this article and our earlier work, we have described several commonalities across different types of designs that are normally thought of as adhering to a qualitative or quantitative paradigm. In this current work we go beyond discussion of commonalities and call for the use of an interactive framework when thinking about design and when judging findings from published studies. In doing so we hope to impose on HRD researchers, practitioners who rely on empirical studies, and journal editors to take an expansive view of research design when interpreting findings. We
encourage thinking about transferability, generalization, and external validity in tandem. We hope there is clear effort to not only understand aggregate effects but also systematic exploration of exceptions to the norm, and overall we call for expanded thinking of what constitutes research.

This is a challenge (one that the authors also struggle with) but we see value in encouraging a mixed methods lens; that is, a lens that considers the synergies between quantitative and qualitative work, when thinking about research. In doing so, we encourage researchers, practitioners, and journal editors to consider Ridenour and Newman's aforementioned definition of scientific research. Consider the design features used in both qualitative (e.g., techniques used to promote credibility) and quantitative (e.g., techniques that promote validity in its various forms) inquiry (e.g., Denzin & Lincoln, 2005; Shadish et al., 2002). Above all, we encourage our readers to consider how features of one type of design can be used to minimize weakness in another type of design (Newman, Newman, & Newman, 2010; Tashakkori & Newman, 2010; Tashakkori & Teddlie, 2010).

Conclusion

When two distinct chemicals interact they can yield a new product. When hydrogen and oxygen are combined in the right way, one has water. This is a metaphor for what this article hopes to articulate. Interacting continuums are useful, but the idea is to emphasize the blending of methods. Validity is a concept to strive for and it is nonsensical to assume one camp or tradition owns it. Researchers should use whatever tools are at their disposal to strive for making valid inferences and understanding conditions under which they are valid. The warrants of the arguments to be presented are based in logical analyses of the purpose and approach in seemingly disparate methods. Although the methodological paradigm wars have been rejected by many (Tashakkori & Teddlie, 2010), it is still traditional to present methods as quantitative, qualitative, or mixed. There may be value in such organization, but there can be a cost in terms of provincial ownership of method and limited creativity in design. This work expands on the idea that different research paradigms share much in common and we wish to give voice to a perspective that is sometimes best to think of research as research and not get caught up in paradigm discussion. It is our hope that establishing the above points will continue to promote methodological diversity among social science, and thus HRD, researchers as well as a stronger understanding of the purpose of different techniques. Conceptual organizers of different methodological approaches such as the Interactive Continuum in Figure 1 are developed to encourage such diversity. The Interactive Continuum can also inform and guide theory building efforts in the social sciences and HRD.
Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

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

Note
1. The paradoxes and logical difficulties associated with null hypothesis tests of statistical significance are of course recognized by others (e.g., Cohen, 1994; Kline, 2004; Morgan, 2003; Schmidt & Hunter, 2002).

References


Bios

**John H. Hitchcock** is an associate professor of education research and program evaluation in the Patton College of Education and Human Services at Ohio University. He has coauthored more than 20 scholarly publications and was a coprincipal investigator of two federally funded randomized controlled trials. His primary interests are in mixed-methods research, program evaluation, culturally relevant intervention and assessment, and special education research.

**Isadore Newman** is the visiting scholar for the College of Education at Florida International University and he is a distinguished professor emeritus at the University of Akron. He has authored or coauthored more than 125 referred articles, more than 300 referred presentations, and approximately 17 books, chapters, and monographs. He was also the principal evaluator on millions of dollars in federal and state grants and was one of the founding editors of the Midwestern Educational Research Journal, editor of the Ohio Journal of Science, and was editor of Multiple Linear Regression Viewpoints Journal for 19 years.