Using Fully Integrated Bayesian Thinking to Address the $1 + 1 = 1$ Integration Challenge

Anthony J. Onwuegbuzie, John Hitchcock, Prathiba Natesan, and Isadore Newman

*Department of Educational Leadership, Sam Houston State University, TX, USA and Department of Educational Leadership and Management/Department of Educational Psychology, University of Johannesburg, Johannesburg, South Africa; *Department of Instructional Systems Technology, Center for Evaluation and Education Policy, Indiana University Bloomington, IN, USA; †University of North Texas, TX, USA; ‡Florida International University, FL, USA

**ABSTRACT**

A fundamental argument for the integration of qualitative and quantitative methods is that doing so can create a form of synergy that allows for an understanding of social phenomena that is difficult to achieve using a mono-method approach. Such synergy has been expressed in a recent editorial that appeared in the *Journal of Mixed Methods Research* as the $1 + 1 = 3$ formula, which is meant to promote exploration of quantitative, qualitative and mixed disciplinary boundaries. There are clear advantages to expressing mixed methods synergy in this way but one potential drawback is doing so arguably reifies a quantitative-qualitative dichotomy that can undermine a fuller and more seamless kind of integration. An example of such fuller integration is seen via Bayesian mixed analyses that allow for using qualitative information to inform subsequent modeling and interpretation. Procedures that we conceptualize entail transforming data forms (e.g., quantitizing and qualitizing) and crossover mixed analyses in ways that treat data as information without necessarily adhering to a quantitative and qualitative dichotomy. This article presents Bayesian analyses in a way that is meant to pursue such integration and presents two applied examples that demonstrate what we refer to as Fully Integrated Bayesian Thinking (FIBT). We argue that FIBT is emblematic of an alternative formulaic option to express full integration via the following formula: $1 + 1 = 1$.

An important aspect of mixed methods work is *integration*, which has been defined as “the interaction or conversation between the qualitative and quantitative components of a study” (O’Cathain, Murphy, & Nicholl, 2010, p. 341). There is great potential to advance how research is conducted by further developing integration ideas but this potential also presents a fundamental challenge, as described in an editorial by the editors of the *Journal of Mixed Methods Research (JMMR)*:

Moving forward, we are posing to the mixed methods community to focus even greater attention to the “integration challenge.” We describe the integration challenge qualitatively as the imperative to produce a whole through integration that is greater than the sum of the individual qualitative and quantitative parts ... Now, with more experience under the field’s belt, we hope to get all mixed methods researchers to consider the mixed methods challenge. Quantitatively, we express this as $1 + 1 = 3$. That is, qualitative + quantitative = more than the individual components. We believe this framework should push all mixed methodologists to think about how integration has and can push research methods to higher levels that would not have been achieved by simply adding together the results of separate qualitative and quantitative studies conducted without full attention to integration ... The $1 + 1 = 3$ integration formula also gives permission to question the assumptions of qualitative and quantitative disciplinary borders and blinders, to test the waters, and to create and discover

**KEYWORDS**

Bayesian; Bayesian mixed analysis; crossover mixed analyses; Fully Integrated Bayesian Thinking; integrated analyses; integration; $1 + 1 = 1$
new ways of thinking and producing mixed methods results. Through these “brain stretches” mixed methods researchers can best illustrate how the additional work of doing mixed methods research results in a “value added” outcome from being in the “radical middle” (Onwuegbuzie, 2012). (Fetters & Freshwater, 2015, pp. 115-116)

In this article, we attempt to address part of this challenge by raising two interrelated points: (a) the synergy expressed by the formula is of central importance to the mixed methods community, so much so that we (the authors) push this idea even further; and (b) our own perspectives on integration has evolved as a function of our thinking about Bayesian estimation, presenting an avenue by which to question the qualitative and quantitative boundaries mentioned by Drs. Fetters and Freshwater. Others who are familiar with Bayesian work also might concur. For example, Tonelli and Shirts (2017) state: “While the importance of mechanistic and Bayesian reasoning are emphasized in guidelines for interpretation of sequence variants, the underlying framework for integrating knowledge derived from various quantitative and qualitative methodologies remains controversial and underexamined” (p. 1650).

To the first point, consider the $1 + 1 = 3$ integration formula. Although it has much potential for enhancing integration within the field of mixed methods research, it is arguable that this formula actually reflects all types of mixing because it reflects an assumption that a dichotomy of methodological approaches (i.e., qualitative approaches vs. quantitative approaches) exists, which we contend either represents a false dichotomy (Nastasi, Hitchcock, & Brown, 2010; Ridenour & Newman, 2008) or is immaterial (Onwuegbuzie & Leech, 2005). Yet, not subscribing to this dichotomy of methodological approaches allows mixed methods researchers to believe in the importance of taking multiple perspectives, to develop research questions with multiple components, to employ multiple methods of data collection and analysis, and to arrive at inferences that incorporate all outcomes that emerge from such a study (Onwuegbuzie & Tashakkori, 2015). Therefore, although the $1 + 1 = 3$ integration formula provides added value for addressing complex research questions relative to monomethod approaches, especially when the integration of the qualitative and quantitative components occurs only at the data interpretation (i.e., inference-making) phase, the mixing can involve partial integration. That is, this integration formula likely yields a mixed methods research study that can be placed at the low(er) end of the integration continuum (i.e., partial integration). Contrastingly, by integrating qualitative and quantitative elements at the data collection and data analysis phases, a mixed methods research study can be placed at the high(er) end of the integration continuum—potentially leading to full integration (cf. Creamer, 2018).

As demonstrated later, one can argue that a Bayesian perspective fits with such full integration if one agrees with Onwuegbuzie (2017), who referred the full integration of qualitative and quantitative elements at the data collection and data analysis phases, alongside the data interpretation phase, as yielding the following formula:

$$1 + 1 = 1$$

This $1 + 1 = 1$ formula entails integrating different forms of data, which includes transforming qualitative data into a quantitative form (i.e., quantitizing; e.g., using numeric counts of themes; Miles & Huberman, 1994; Onwuegbuzie & Teddlie, 2003; Sandelowski, Voils, & Knafl, 2009; Tashakkori & Teddlie, 1998), transforming quantitative data into a qualitative form (i.e., qualitizing; e.g., obtaining narratives to explore the meaning of numerical data; Onwuegbuzie & Teddlie, 2003; Sandelowski et al., 2009; Tashakkori & Teddlie, 1998), and using joint displays, which Onwuegbuzie and Dickinson (2008) call “crossover (mixed research) graphical displays” or “‘crossover’ visual extensions to summarize and integrate both qualitative and quantitative results within the same framework” in order to “enhance[e] researchers’ understanding (i.e., increased Verstehen) of social and behavioral phenomena in general and the meaning that underlie[e] these phenomena in particular” (p. 204). Moreover, this $1 + 1 = 1$ full integration formula promotes the use of crossover analyses, which involve using one or more analysis types associated with one tradition (e.g., quantitative analysis) to analyze data representing a different tradition (e.g., qualitative data); Onwuegbuzie & Combs, 2010). Such analyses hold great promise for a different and more seamless kind of integration across quantitative and qualitative thinking, facilitating mixed methods researchers in addressing increasingly complex research questions. Figure 1 shows a visual representation comparing $1 + 1 = 3$ partial integration and the alternative $1 + 1 = 1$ full integration as a function of degree of integration. Consistent with our assertion, Teddlie and Tashakkori (2009) declared that “We believe that this [use of crossover mixed analyses] is one of the more fruitful areas for the further development of MM [mixed methods] analytical strategies” (p. 281).
To this end, Onwuegbuzie and Hitchcock (2015) have advanced a mixed analysis framework for conducting analyses that increases the level of integration at the data collection and data analysis phases, and Newman, Onwuegbuzie, and Hitchcock (2015) describe a conceptual framework for pursuing fully integrated mixed analyses via the general linear model. Building on these efforts, we believe that another advancement is to rely on the logic of crossover analyses as a basis for including Bayesian estimation within mixed methods research. The way that Bayesian estimation works, relative to frequentist Null Hypothesis Significance Testing (NHST), might appeal to mixed methods researchers interested in maximizing integration in general and conducting crossover analyses in particular. Therefore, in this article, we review Bayesian approaches and connect these back to how a qualitative understanding of context can be used potentially to enhance $1 + 1 = 1$ full integration. In short, we introduce the concept of **Fully Integrated Bayesian Thinking** (FIBT). In the remainder of this article, we first provide an overview of Bayesian approaches to statistical analyses. Then, to promote the use of FIBT and demonstrate what we mean by *full integration*, we provide two examples wherein FIBT is applied to making predictions about dichotomous decisions and meta-analyses. This yields new mixed methods research tools (with the understanding that these represent only two examples of how FIBT might be used within mixed methods inquiry).

Our overall goals for this article are (a) to begin to address Fetters and Freshwater’s (2015) challenge by demonstrating a fuller and seamless way of thinking about mixed methods integration and (b) to promote the use of Bayesian estimation within the mixed methods research community.

### An Overview of Bayesian Statistical Analyses

Bayesian methods are gaining popularity within the social sciences (e.g., Gill, 2014; Kruschke, 2015). This approach has been contrasted with classic frequentist approaches (that are familiar to most readers of this journal) based on both theory (e.g., Gelman et al., 2013; Gill, 2014; Kruschke, 2015) and performance (e.g., De la Torre, Stark, & Chernyshenko, 2006; Kieftenbeld & Natesan, 2012; Natesan, 2015). In Bayesian estimation, parameters are estimated as distributions and not as point estimates (that have a standard error of the estimate), as is the case with frequentist approaches. In Bayesian estimation, each value in the estimated parameter distribution is a probable value that the parameter can take, and each such value is associated with a probability that the parameter would take that value. This distribution is called the posterior distribution.

Bayesian work can potentially entail $1 + 1 = 1$ seamless integration by accounting for contextual knowledge, which can be qualitatively understood, with quantitative analyses. The basis for this is borne out of the idea that...
qualitative inquiry can help lay bare and justify how researchers think about prior information. This is important because, in Bayesian estimation, prior information about a given parameter is supplied by the researcher and its accuracy can vary depending on many factors, including context, previous research findings, the researcher’s belief about the parameters, and history. Furthermore, there is a natural alignment between the progression of evidence that is used in Bayesian approaches wherein the results of probability estimation (called posterior distributions) can be used to inform a prior in subsequent work, and how sequential designs can be used in singular studies or across multiple studies in a program of research.

Consider Bayes’s theorem:

$$p(\theta|\text{Data}) = \frac{p(\text{Data}|\theta)p(\theta)}{p(\text{Data})}$$  \hspace{1cm} (1)

where $p$ refers to probability, $\theta$ is the parameter (or parameters) of interest, and $p(\theta|\text{Data})$ is the posterior distribution of the parameters. The posterior distribution is the probability distribution of the parameters of interest given (i.e., conditional on) the observed data. In Equation 1, $p(\text{Data}|\theta)$ is the likelihood of the data, that is, the information contained in the data; $p(\theta)$ is the prior distribution—that is, the researcher’s belief about the parameters; and $p(\text{Data})$ is the marginal probability of the data. The theorem expresses the key idea that a researcher’s beliefs about a parameter and accumulation of evidence can together yield improved posteriors.

Bayesian statistics has not yet gained the popularity that it arguably deserves in the social sciences (Boedeker, Natesan, & Onwuegbuzie, 2017), and there are two basic reasons for why this is so. First is the matter of $p(\theta)$, or the prior distribution, which is defined by the researcher; because the researcher provides this definition, many frequentists question the objectivity of Bayesian methodology (Gelman, 2008). $P(\text{Data})$ provides the second reason. Specifically, the denominator in Equation 1 takes the form of a sum for discrete parameters and an integral for continuous parameters. The integral quickly gets complicated with increase in model complexity. Such integrals are intractable or unsolvable. Indeed, because there was no solution to this problem for several models, Bayesian statistics lay dormant for centuries. However, the Gibbs sampler, introduced by Geman and Geman (1984) and other Markov chain Monte Carlo (MCMC) methods, solved this issue by approximating the posterior distribution without solving for the denominator. That is, the parameter estimate is now sampled from the product of prior and the likelihood. This process is repeated iteratively to form a distribution of parameter estimates called the posterior distribution. One could think of the posterior, which is the subjective probability distribution of the parameter, as an analog to the sampling distribution of a statistic in the frequentist framework obtained by hypothetical resampling.

The Intuitive Appeal of Bayesian Estimation

The theorem can be restated as the following proportionality:

$$p(\theta|\text{Data}) \propto p(\text{Data}|\theta)p(\theta)$$  \hspace{1cm} (2)

or:

$$\text{Posterior} \propto \text{Likelihood} \times \text{prior}$$  \hspace{1cm} (3)

The left-hand side of the posterior, warrants special attention for readers who are new to Bayesian estimation. This is because in introductory statistics courses, students typically are taught about the $p_{\text{calc}}$, which is defined as the probability of obtaining a sample that is at least as extreme as the observed data, assuming that the null hypothesis is true [$p(D|H_0 = \text{true})$]. This is often misinterpreted in NHST by assuming that the null is true in the population and calculating the probability of obtaining the observed data from a population where this is so [i.e. $p(H_0 = \text{true}|D)$]. Cohen (1994) however discusses, in detail, how the two conditional probabilities—that is, $p(H_0 = \text{true}|D)$ and $p(D|H_0 = \text{true})$—are not the same. To summarize one of Cohen’s points, when a null hypothesis is rejected (e.g., $p_{\text{calc}} < .05$), what is learned is that the observed data are unlikely if the null is true in the population. However, what most researchers actually want to know is the probability that the null hypothesis is true given the observed data. For this question, the Bayesian approach is needed and it is Proportionality 2 that yields what researchers typically want, namely, $p(H_0 = \text{true}|D)$.

This raises another important aspect of Bayesian statistics, which is the generation of credible intervals. In frequentist statistics, confidence intervals (CI) are obtained and reported because we know that the point estimates do not express the uncertainty associated with the statistical estimate. However, the frequentist 95% CI does not express the probability that the true value can be found in that interval (a frequentist approach attempts to deal with this using replicability estimates). Rather, a 95% CI is associated with infinite replications of the data analysis, of which 95% of them will contain the true population value. This 95% CI does not imply that
there is a 95% probability of this CI containing the population statistic (Gorard, 2015). In essence then, the frequentist CI obtained by the researcher gives us very little confidence as to whether the true value is contained in the interval at all (again, see Cohen, 1994). With respect to the Bayesian counterpart, the credible interval is more straightforward to interpret. That is, the probability that the true value is contained in a 95% credible interval is 0.95. This is because the Bayesian posterior distribution contains parameter estimates obtained from thousands of data analysis iterations of the estimates. Thus, the distributions of the parameter estimates have a shape and probabilities associated with each probable parameter value. Examining these distributions provides the researcher with a better understanding of how best to trust a point estimate (which is usually an expected a posteriori estimate in frequentist methods).

Subjective Priors

In Proportionalities 2 and 3, the prior information that is given by the researcher has an influence on the calculation of the posterior. Hence, the question of what prior to use and its subjectivity has been discussed extensively (e.g., Gelman et al., 2013). Importantly, the effect of this prior decreases with increase in sample size because the impact of the likelihood will increase for larger samples or when the data have higher quality (e.g., low measurement error). This is because high quality data and generally large samples give more information in the form of likelihood $p(\text{Data}|\theta)$. Therefore, priors need to be chosen more carefully when small samples are analyzed. Depending on the amount and the type of information that they provide, priors can, in general, be at least of three types: (a) an informative prior, (b) an uninformative prior, and (c) a hyperprior. An informative prior provides very specific information or information with less uncertainty, whereas an uninformative prior yields less information or information with more uncertainty. A hierarchical or hyperprior can be a good compromise between these two ends of informativeness, where a prior can be placed on priors and the researcher can remain noncommittal. For instance, consider an item response model where the person abilities ($\theta$) are measured on a standard unit normal scale. Equations 4-6 give examples of informative, uninformative, and hyperpriors, respectively.

\[
\begin{align*}
\text{Informative prior: } & \theta \sim N(0,1) \\
\text{Uninformative prior: } & \theta \sim N(0, 10000) \\
\text{Hierarchical/hyperprior: } & \theta \sim N(\mu, \sigma), \text{ where } \\
& \mu \sim N(0, 10) \\
& \sigma \sim \text{gamma}(1, 1)
\end{align*}
\]

Equations 4 and 5 represent priors with variances 1 and 10,000, respectively. This means that the information provided by the prior in Equation 4 is packed closely around 0, with 95% of the values lying between -2 and 2, and so on (because this is a normal distribution). The information around the prior in Equation 5 is not as closely packed around 0, but rather 95% of the values lie between -20,000 and 20,000, and so on. Even though they are both normally distributed around 0, the probability density (i.e., the height of the bell curve) around 0 will be much smaller for the uninformative prior in Equation 5 than for the informative prior in Equation 4 (an example of visualization is demonstrated in Figure 2). In Equation 6, we do not know the mean and standard deviation of the distribution from which $\theta$ needs to be sampled. Therefore, we specify the mean and standard deviation for this distribution to be unknown parameters $\mu$ and $\sigma$, respectively. In turn, $\mu$ and $\sigma$ are sampled from normal and gamma distributions, respectively. This specification is called a hierarchical or hyperprior, where prior distributions are specified for prior parameters ($\mu$ and $\sigma$). What happens when one uses an uninformative prior such as the one in Equation 5? This would most probably take longer time and more iterations to converge and, depending on the sample size, the estimates might yield large posterior standard deviations. Alternatively, using a hyperprior such as the one in Equation 6 will tend to yield more accurate expected posteriori estimates with larger standard errors than do more informative priors but smaller standard errors than do uninformative priors (Natesan, Nandakumar, Minka, & Rubright, 2016). However, for large samples, the accuracy of the estimates usually is as good as are those from informative priors. In sum, the choice of prior requires careful consideration, especially for smaller samples. At its least optimum point, if the researcher has no information about the prior distribution and has no means of arriving at a subjective probability, then he/she may use uninformative or hierarchical priors. A key point here is that the use of qualitative data can play an important role here in obtaining an informative prior, which could be seen as a kind of crossover analysis. This represents a form of $1 + 1 = 1$ seamless integration of mixed methods thinking. Researchers can modify the weight given to the qualitative prior by altering its informativeness or its variability similar to the examples shown in Equations 4-6. But before getting into the use of qualitative inquiry for this purpose, it might help to review the idea of subjective probability.
Subjective probability defined. In the research context, subjective probability is the researcher’s personal judgment about how likely a particular outcome is; however, rather than being based on any precise computation, a subjective probability represents a rational—and hopefully a reasonable—assessment by the researcher. A subjective probability may contain no formal calculations and reflects the researcher’s perceptions, opinions, historical data or results, and/or past experience. Therefore, subjective probabilities often differ from one researcher to the next and because the probability is subjective, it reflects at least some personal bias. The higher the degree of personal bias (to the degree that a bias yields inaccurate thinking about the phenomenon of interest), the less informative or accurate the probability estimate likely will be for making inferences, thereby adversely affecting the integrity of the quantitative findings. Of course, personal bias is likely to be less of a concern if it is: (a) transparent and (b) any decisions that inform subjective probability are well justified with logic, prior data, and extant literature. Add to this, bias can be well understood and made clearer by the reflective researcher who is accustomed to describing motivation for carrying out research and a priori expectations of findings. Indeed, to promote confidence in the subjective prior, any number of credibility techniques (e.g., member checking, triangulation; for a review of 24 techniques, see Benge, Onwuegbuzie, & Robbins, 2012; Onwuegbuzie & Leech, 2007) could be built into a design so that a subjective decision can be well-justified.
Often, frequentists are uncomfortable with the subjectivity allowed by the use of priors (we think this discomfort also contributes to the quantitative–qualitative paradigm argument). Yet, all statisticians and students of statistics are taught to make subjective judgments. For instance, Cohen (1994) argued about the subjectivity in the choice of $P_{crit} = 0.05$ for NHSTs. Similarly, he discouraged the use of guidelines such as Cohen's (1988) $d$ values of 0.2, 0.5, and 0.8 as small, medium, and large effect size values associated with the two-sample case, respectively. Further, the very idea of converting a probability continuum into a strict dichotomous reject/do-not-reject decision is subjective in itself because it denies the possibility of other decisions (Thompson, 1996, 1997). For yet another example of subjective decision-making, consider structural equation modeling (SEM). All SEM researchers acknowledge that there are many possible models to explain the phenomenon under study and yet choose to test only a selection of models that support their hypotheses (e.g., Schumaker & Lomax, 1996), which is a function of their judgement—representing a subjective sampling decision (Onwuegbuzie & Collins, 2017). The decision to retain all SEMs based on fit indices is also not entirely objective but guided by recommended fit indices that are based on simulation studies of some models.

Of course, from a qualitative perspective, subjectivity is a well-understood idea. When it comes to subjective probabilities, Onwuegbuzie and Collins (2014, 2017) outlined how subjective probabilities are used by virtually all qualitative researchers when interpreting their findings. Furthermore, connecting the notion of subjective probability to mixed methods inquiry, debriefing techniques developed for qualitative research (e.g., Chenail, 2011; Onwuegbuzie, Leech, & Collins, 2008) and mixed methods research (Collins, Onwuegbuzie, Johnson, & Frels, 2013) are likely to help researchers be cognizant of the biases that they bring to the study and how these biases influence the subjective decisions made. In sum, thinking about subjective probability can help researchers make their thinking transparent to those who might consume and critique their work, and we encourage mixed methods researchers to embrace the idea when applying FIBT.

Starting to Address the Integration Challenge Via FIBT

To summarize the key issues raised so far in this article, Bayesian estimation has seen a relatively recent resurgence because of the aforementioned technical advances, the relative ease of interpreting credible intervals, and, perhaps most interesting to the mixed methods research community, related analyses can rely on priors that utilize the subjective judgement of an analyst. Central to this article is the idea that prior qualitative information can be used better to inform priors. Now, we offer our first (hypothetical) example of FIBT based on Say, Thomson, Robson, and Exley's (2013) qualitative interview study that explored pregnant women's attitudes to external cephalic version (ECV). ECV is a process by which a breech baby can be turned into a head first position. Women who have breech presentation have to decide whether they want to attempt ECV; therefore, there is merit in understanding the attitudes that they carry about the process. Four themes of women's attitudes emerged from the Say et al. interviews.

1) ECV as a means of enabling natural birth: Some women perceived ECV as a way to avoid caesarean section (CS), to experience vaginal delivery.

2) Concerns about ECV: Some women considered ECV as unnatural and “manipulating” what the baby had “chosen.” On a related point, ECV was presented in interviews as having only a 50% success rate. Some women thought that this might be because there is not always enough space for the baby to turn, whereas some perceived the procedure as being invasive. Furthermore, some women feared potential complications of ECV and others feared that the baby might revert to breech position after ECV.

3) Lay versus professional accounts of ECV: Based on hearsay, some women believed that it was a painful, dangerous, and unsuccessful procedure.

4) Breech presentation as a means of choosing planned CS: Some women chose CS irrespective of the breech as a means to avoid pain and because it was more convenient.

With this background, we offer a scenario of how qualitative themes might be used within the Bayesian estimation process, representing a fuller form of integration that perhaps starts to address Fetters and Freshwater's (2015) challenge. It stands to reason that women whose responses fell under Theme 1 might be expected to choose ECV when the option is presented, whereas the others might be expected not to choose ECV. This qualitative information can seamlessly inform a prior and thus facilitate Bayesian analyses to develop more precise probability estimates for future planning. This would be useful because the capacity to predict ECV choice might inform future efforts to guide women's choices and to allocate resources within, say, a given hospital. To demonstrate the process, assume that in the initial interview study that followed a probability sampling scheme, three responses fell under Theme 1 and five fell under the other themes. From there, assume a new study within the hospital was conducted where 11 women out of 13 chose ECV. The probability that women would choose ECV
in the future can be computed via a beta-binomial problem where the prior and posterior are both beta distributions, \( BE(p|3, 5) \) and \( BE(11 + 3, 2 + 5) = BE(14, 7) \), respectively (for technical details refer Gill [2014], Chapter 2). A beta distribution is a continuous probability distribution and is commonly used in Bayesian analysis to describe the probability of success. The two parameters used in the beta distribution are shape parameters. Understanding the basis for the prior distribution selection is the point at which quantification of qualitative information (i.e., quantitizing) occurs. The parameters of the prior distribution are based on a simple frequency count from the initial interview study of how many women provided responses that fell within each theme. Three women provided interview responses that fell under a theme showing a positive attitude towards ECV, and five women fell under themes suggesting a negative attitude towards ECV. This is an example of a qualitatively informed Bayesian prior.

To draw out the value of the qualitatively informed prior, suppose a researcher instead uses a non-informative prior. This would likely entail assuming a 50-50 chance of women selecting the procedure, that is, \( BE(p|1,1) \). This choice will result in a posterior \( BE(11 + 1, 2 + 1) = BE(12, 3) \). These two sets of solutions are represented graphically in Figure 2. Critically, the highest posterior density interval (HPD, which is a specific credible interval) is narrower and shows a more conservative estimate of how likely women are to choose ECV when the qualitatively informed prior is utilized. This suggests that women are less prone to see ECV as a means for enabling natural childbirth; as a result, it becomes easier for analysts to see that there are consequences for women’s attitudes around the procedure. A key take away here is that this example demonstrates that it is possible to integrate qualitative information into Bayesian estimation, representing a mixed methods framework that can yield findings that are not as easily obtained using monomethod approaches. Further, this Bayesian estimation is predicated on qualitative work that was needed to recognize and to differentiate Theme 1 from other options, representing FIBT in particular and, more generally, the 1 + 1 = 1 framework. This is because there is use of subjective priors when estimating models, and estimating models was contingent upon the interview study. To demonstrate FIBT-crossover analysis further, the following section describes its application to advance meta-analyses.

A FIBT application to meta-analysis. Since Glass coined the phrase meta-analysis in 1976, this technique has become popularized among quantitative researchers. In a classical meta-analysis (i.e., synthesis of quantitative studies using frequentist approaches), the analyst combines findings from as many available individual research studies as possible that address one or more related research hypotheses in order to synthesize the results (Glass, 1976). There are two major goals in such meta-analyses: (a) to estimate the mean effect size across the selected studies; and (b) to examine the variability of effect sizes across studies as a function of study design effects (e.g., whether the effect size was different for intervention studies that utilized randomization vs. studies that did not)—a technique called homogeneity analyses. Several strengths of meta-analyses have been documented, including the synthesis of findings so that policy decisions can be informed by several studies, and related techniques allow for the incorporation of moderating variables when summarizing findings and facilitate the identification of characteristics of studies (e.g., study quality indices such as quantitative research design used and sample size) as potential explanations for consistent/inconsistent findings across studies (see, e.g., Onwuegbuzie & Frels, 2016). However, meta-analysis also has a key limitation: “the quantitative components do not allow the reviewer to identify the more qualitative distinctions among studies—yielding a ‘lack of context’ criticism” (Onwuegbuzie & Frels, 2016, p. 252). As an example, conducting what they termed a best-evidence synthesis, Slavin, Cheung, Groff, and Lake (2008) systematically reviewed research (n = 33) on the achievement outcomes of four approaches for improving the reading of middle and high school English language learners in an attempt to assess how much of a scientific premise there is for competing claims about the effects of these programs. Yet, as noted by Slavin and Cheung (2003), their synthesis—which we call a monomethod synthesis—excluded qualitative research studies, even though they acknowledged that these are “valuable in suggesting programs or practices that might be effective” (p. 17). Thus, best-evidence syntheses completely omit information from qualitative research studies, even after acknowledging the value of such studies. Slavin and Cheung (2003) assume that information stemming from qualitative research studies cannot be incorporated into a meta-analysis. In contrast, by adopting \( 1 + 1 = 1 \) full integration thinking, FIBT can be used to address this problem.

To begin to understand how, note that Bayesian meta-analyses (which still use only quantitative research studies) have several advantages over classical meta-analyses. The previous discussion of Bayesian work should help readers see that Bayesian meta-analyses allow analysts to use credible intervals that describe the probability that they contain a true value, and use of a coherent modelling framework that can use evidence from different sources (Sutton & Abrams, 2001). However, Bayesian meta-analyses still only incorporate quantitative findings and ignore all the qualitative findings. Yet, there is a form of Bayesian meta-analysis approach that incorporates (i.e., integrates) qualitative data and we frame these as FIBT meta-analyses.
There are at least two ways of conducting a FIBT meta-analysis. The first is by giving the quantitative findings more emphasis than the qualitative findings such that the qualitative evidence is used to update prior probabilities, and with only the quantitative findings contributing to the likelihood. The second way is to allow the qualitative and quantitative findings to affect the results in the same way, which yields a measure of uncertainty in the estimate of the probability of an effect. That is, the qualitative and quantitative findings are simultaneously incorporated into the likelihood function. Further, a FIBT meta-analysis can involve integration of qualitative and quantitative findings (a) at the \textit{participant level}, by linking findings to the numbers of participants yielding them; (b) at the \textit{study level}, by creating what Onwuegbuzie and Teddlie (2003) referred to as an inter-respondent matrix that contains a series of 1s and 0s depending on whether the themes are present or absent, respectively (see also, Onwuegbuzie, 2003); or (c) at both the participant and study levels. It should be noted that by manipulating the degree of informativeness in the prior, one can allow the prior to have equal or even more influence on the posterior than the data likelihood. However, Bayesian analysts are very much aware of this issue, and sensitivity analysis is one of the checks that represent a best practice for ensuring that undue influence of the prior does not occur. Indeed, much research has been conducted regarding the sensitivity of the estimates to prior specification that can be used to investigate the impact of priors on the estimates (Gelman, 2002).

To demonstrate FIBT meta-analysis, we need not construct an example—especially bearing in mind the space constraint for this article—because three meta-analyses that can be framed as aligning with FIBT appear to have been conducted to date: Crandell, Voils, Chang, and Sandelowski (2011); Roberts, Dixon-Woods, Fitzpatrick, Abrams, and Jones (2002); and Voils et al. (2009). Each of these studies is described briefly in the following sections.

\textit{Roberts et al.’s (2002) FIBT meta-analysis.} In what Voils et al. (2009) identified as representing “the first to demonstrate the use of Bayesian methods to synthesize qualitative and quantitative research findings” (p. 231), Roberts et al. (2002) used findings from the qualitative research studies to provide prior information, thereby allowing this evidence to inform the analysts as to which factors might be related to the uptake of immunization in children. Findings from the body of qualitative research studies were compiled and combined with the meta-analysts’ own expert judgement to provide an estimate of the prior probability that each factor was related to receiving immunization (representing what we judge to be emblematic of seamless integration). Next, they used evidence from the quantitative research studies on the same topic to update these prior probabilities, which yielded posterior probabilities that each factor was associated with the decision to immunize. Simply put, Roberts et al. gave more emphasis to the quantitative findings than the qualitative findings. This is because the qualitative findings were used to obtain prior probabilities of the relationships, which, subsequently, were updated via Bayesian estimation.

\textit{Voils et al.’s (2009) FIBT meta-analysis.} In contrast to Roberts et al. (2002), Voils et al. (2009) used qualitative and quantitative findings from the extant literature with equal emphasis to examine whether HIV-positive women were less likely to continue taking their antiretroviral medications if their medication schedule was more complex than if their medication schedule was less complex. They identified 11 qualitative research studies and six quantitative research studies that informed their meta-analysis. To make the qualitative and quantitative findings compatible, the findings from the qualitative studies were quantitized. These quantitized data were synthesized in the same way as were findings from the quantitative research studies to yield an estimate of the probability that, for a given research participant, the complexity of the regimen was associated with adherence.

\textit{Crandell et al.’s (2011) FIBT meta-analysis.} Crandell et al. (2011) were interested in examining 10 factors that were associated with adherence/nonadherence to HIV medication regimens. Crandell et al. conducted a different type of FIBT meta-analysis wherein, as was the case for Voils et al.’s (2009) study—findings from the 12 qualitative research studies and 15 quantitative research studies were given equal emphasis. Because of the lack of subject-level information in the qualitative research studies, Crandell et al. instead decided to examine all qualitative and quantitative research studies at the study level. They quantitized the quantitative findings (i.e., into themes), and then categorized each study according to the presence/absence of themes; specifically, using Cohen’s $d$ effect size (i.e., the standardized difference between means), they coded the quantitative research studies as follows: as “1” if $d \geq .20$, as $.50$ if $.20 < d < .20$, and as 0 if $d \leq .20$. For the qualitative research studies, they coded a variable as 1 if it related to adherence, as .5 if it related to both adherence and nonadherence, or neither; and as 0, if it related to nonadherence. For both the qualitative and quantitative research studies, if a study did not address a certain variable, the study was coded as missing because they did not want to assume a lack of effect of a variable that was not examined at all. After the themes had been extracted from the qualitative and quantitative research studies, the authors created a matrix that summarized the thematic coding of all the research studies, with each column corresponding to one of the selected factors and each row to an individual qualitative or quantitative research study. For each variable, the cell entries were as follows: 1 = factor was reported as promoting adherence, .5 = factor was reported as having no effect on adherence, or 0 = factor was
reported as promoting nonadherence. Because this process generated many missing cells, a Bayesian data augmentation method was used to impute data, under the assumption that each row in the table, if completely observed, represented a random vector drawn from some type of multivariate distribution. The Bayesian data augmentation method involved using the observed data to estimate the unknown parameters of the multivariate distribution, then drawing from this distribution to provide an estimate of the missing values themselves, with the parameters of the distribution being re-estimated. This step is repeated until a large number of estimates had been obtained. Then, these estimates were summarized to describe the parameters of the multivariate distribution from which each row was selected. Subsequently, Crandell et al. summarized, ranked, and compared the effects of each of the 10 factors on medication adherence while summarizing information from both quantitative and qualitative research studies. This is not trivial. Promoting adherence to medication regimens is obviously a critical aspect of care, and physicians would seemingly want to know that their research-based practice accounts for qualitative information.

Heuristic example. Returning to Slavin et al.’s (2008) systematic review of research (n = 33) on the achievement outcomes of four approaches for improving the reading of middle and high school English language learners, these researchers could have transformed their traditional meta-analysis into a FIBT meta-analysis by using the qualitative evidence to update the prior probabilities and the 33 sets of quantitative findings to contribute to the likelihood, or by allowing both the qualitative and quantitative findings to play an equal role (i.e., have equal weight). Similar to Crandell et al.’s (2011) FIBT meta-analysis, a matrix could be constructed that summarizes the coding of the themes extracted from the qualitative and quantitative research studies, wherein each theme is quantitized across all these studies.

Conclusion

We agree with Fetters and Freshwater (2015) that their 1 + 1 = 3 integration framework “should push all mixed methodologists to think about how integration has and can push research methods to higher levels” (p. 116). However, we contend that a 1 + 1 = 1 integration offers a complementary framework that should motivate all mixed methodologists to think not just about integration but about full integration. Indeed, by illustrating FIBT applications to dichotomous probability estimation and meta-analyses, we illustrate how use of FIBT approaches—representing crossover analyses—necessitate the integration of qualitative and quantitative data analyses. In our examples, the underlying Bayesian analysis was transformed from a traditional quantitative analysis to an integrated analysis. In particular, we outlined how qualitative data and analysis can be incorporated into Bayesian analysis as a means of obtaining improved posteriors—either in a subordinate manner to qualitative data and analysis, or in an equal status manner.

FIBT is a potentially exciting trend in mixed methods research because it involves full integration of methodological approaches—which includes full integration at the following stages of the mixed methods research process: research conceptualization (e.g., full integration at the philosophical levels that includes [a] integration of qualitative-based and quantitative-based mental models and [b] integration of qualitative-based and quantitative-based analytical assumptions; generating research questions that are best addressed via integration of qualitative and quantitative data and analyses), research planning (e.g., integration of samples such as when extracting both qualitative and quantitative studies to be used within the same meta-analytic framework), research implementation (e.g., integration of qualitative and quantitative analysis and data), and research dissemination (e.g., enhanced integration of inferences [i.e., meta-inferences] for stakeholders). Even more encouragingly for the field of mixed methods research, because FIBT applications necessitate both qualitative and quantitative analyses, they motivate the full integration of qualitative and quantitative analysts working together on the same set of data, rather than in the 1 + 1 = 3 case wherein the qualitative and quantitative analysts of the mixed methods research team analyze separately the qualitative data and quantitative data, respectively. Thus, the 1 + 1 = 1 integration framework offers exciting possibilities for full integration (Hitchcock & Onwuegbuzie, in press).

However, we only just started to scratch the surface of FIBT applications; Newman and Schumacker (2012), for example, describe mixing methods for Regression Discontinuity Designs and others have described mixing qualitative, randomized controlled trial and survey methodologies (e.g., Hitchcock & Nastasi, 2011). If these forms of mixing occur, then it seems likely that they yield opportunities for FIBT. As these ideas become developed, it will be important to bear in mind that FIBT requires knowledge of qualitative, Bayesian, and mixed methods research approaches, thereby likely necessitating the formation of interdisciplinary, transdisciplinary, and multidisciplinary teams to promote maximal integration—which, in itself, would be a very positive development!
References


