Abstract: In evaluation and applied social research, focus groups may be used to gather different kinds of evidence (e.g., opinion, tacit knowledge). In this article, we argue that making focus group design choices explicitly in relation to the type of evidence required would enhance the empirical value and rigor associated with focus group utilization. We offer a descriptive framework to highlight contrasting design characteristics and the type of evidence they generate. We present examples of focus groups from education and healthcare evaluations to illustrate the relationship between focus group evidence, design, and how focus groups are conducted. To enhance the credibility of focus group evidence and maximize potential learning from this popular qualitative data collection method, we offer a set of questions to guide evaluators’ reflection and decision making about focus group design and implementation.
Focus Group Evidence: Implications for Design and Analysis

Focus groups generate evidence\(^1\) that is commonly used for evaluating diverse programs and policy (Balch & Mertens, 1999; Krueger & Casey, 2009; Poitras Duffy, 1993). They are employed in different types of evaluations: needs assessment, program theory development, and implementation and outcome evaluation. Flexible and efficient, focus groups add a social dimension to verbal data in evaluation. The variations in how to carry out a focus group are as many as the uses (Belzile & Oberg, 2012; Denzin & Ryan, 2007; Farnsworth & Boon, 2010; Kamberelis & Dimitriadis, 2013; Morgan, 2012) and there is a robust and rich evaluation and applied social research literature that elucidates procedural and practical issues of planning and implementing focus groups (e.g., Krueger & Casey, 2009; Morgan, 1997). With so much variety from which to choose, focus group design decisions may be more complex and nuanced than meets the eye.

We propose that focus groups should be designed with a focus on the type of evidence to be generated (e.g., opinion, tacit knowledge). Gathering different types of focus group evidence (i.e., the type of information gathered and inferences to be drawn) requires different kinds of research designs (Belzile & Oberg, 2012; Kitzinger, 1994). For instance, depending on the evaluation purpose and context, evaluators may be interested in gathering opinions which reflect people’s stable personal dispositions. Evaluators could also be interested in accessing tacit knowledge that is more dynamic and socially constructed. The specific design choices made by the evaluator in planning focus groups such as the role of participant interaction, focus group structure (e.g., semi-structured, non-standardized), the role of the moderator (e.g., neutral, ancillary), and data analysis approach (e.g., verbal content or verbal content and participant...

\(^1\) For the purposes of this paper, we define evidence as “information helpful in forming a conclusion or judgment” (Schwandt, 2009, p. 199).
interaction) are notably different depending on the type of evidence to be generated (e.g., basic information, tacit knowledge). In this paper, we argue that making focus group design choices explicitly in relation to the type of evidence required would enhance the empirical value and rigor associated with focus group utilization.

We begin by presenting an overview of focus groups that includes concept definitions, a brief history, and a short summary of recent theoretical developments (Belzile & Oberg, 2012; Kamberelis & Dimitriadis, 2013; Markova, Linell, Grossen, & Orvig, 2007; Morgan, 2012). Then we present a descriptive framework based on contrasting theoretical perspectives and explain the focus group design characteristics associated with each to highlight distinctions critical in guiding evaluators’ design decisions. Our review and analysis are unavoidably selective and we expect that advocates of particular approaches to focus groups may dispute some of the distinctions we delineate.

To illustrate differences between focus group perspectives, particularly in relation to design and implementation of focus groups in evaluation, we present three examples. These focus groups were planned and conducted as part of evaluation projects in the areas of education and healthcare. After critically analyzing each example in relation to the descriptive framework, we discuss the strengths and weaknesses of each approach including the kind of evidence derived. The paper concludes by suggesting a set of questions that evaluators might address when designing focus groups to make more explicit the logic behind their design decisions.

**The Nature of Focus Groups**

The focus group is a particular type of group interview where the moderator (or researcher/evaluator) asks a set of targeted questions designed to elicit collective views about a specific topic (Fontana & Frey, 2005; Merton & Kendall, 1946). The character of participants’
interactions as well as the type of data collected distinguish the focus group from other methods—specifically, participants interact with “each other as well as the moderator” (Wilkinson, 1998, p. 182). Focus groups may be characterized as a particular kind of group interview or as a collective conversation, reflecting substantial variation in the degree to which groups are managed by the researcher or are allowed to be more free-flowing (Kamberelis & Dimitriadis, 2011; Krueger & Casey, 2009).

Typically, a group of 6-8 participants, purposefully selected based on a significant homogeneous characteristic, engage in a face-to-face 1-2 hour discussion of a limited set of topics. Focus group research often utilizes some type of purposive sampling scheme, such as typical or maximum variation (Patton, 2002; see MacDougall & Fudge, 2001 and others for details on focus group sampling). A variety of stimulus materials (e.g., survey questions, photos) can be used for focus group facilitation. Modern focus group modes, which capitalize on technological advances, include real-time and asynchronous on-line focus groups, traditional and computer-assisted telephone focus groups, and others. Depending on the mode, the focus group may be smaller (4-6 participants) and of shorter duration (60-90 minutes) (Krueger, 2009).

From the ‘Focused’ Interview to Understanding Diversity

The focus group is a contemporary methodological development, in contrast to other qualitative research methods such as the individual interview or participant observation. The focused interview (which evolved to become today’s focus group) was conceptualized and implemented as a research method by Robert Merton when he joined a project directed by Paul Lazarsfeld in the Bureau of Applied Social Research at Columbia University. The early focus group was intended to augment an experimental, quantitative approach to studying audiences’ responses to recorded radio programs and Army training films by scrutinizing “subjective
experiences” of the audience (Merton, Fiske, & Kendall, 1990). To assess audience opinion, a group of individuals (N=12 or so) pressed red and green buttons indicating a negative or positive response to what they listened to on the radio. As a supplement to the quantitative audience response study, the researcher conducted the “focused interview” designed to investigate “a set of hypotheses concerning the meaning and effects of the determinate situation” (e.g., what the audience heard on the radio; Merton & Kendall, 1946, p. 541).

Focused interviews were used consistently in consumer research from the 1950s and later, for modern political opinion polling (Denzin & Ryan, 2007; Stewart & Shamdasani, 1990). Remarkably, the focus group was largely absent in social science research until the 1980s, when interest in qualitative methods escalated (Bloor, Frankland, Thomas, & Robson, 2001; Morgan, 2002). From this juncture forward, the use of focus groups thrived and became increasingly varied in approach. Focus groups were utilized in market-oriented research and as a robust qualitative method to examine the meanings of participants’ experiences and to understand diversity in society (Morgan, 2002; Wilkinson, 1998).

In applied social research and evaluation, focus groups are also employed to gather different kinds of evidence that requires distinctive types of research designs. For example, focus groups are used (a) to gather basic information or in questionnaire design and development (Mitra, 1994; Poitras Duffy, 1993), (b) to yield rich description (e.g., generate program theory; Buttram, 1990; Carvalho & White, 2004), and (c) to include perspectives of marginalized and other stakeholders, (e.g., enact participatory, democratic processes; Baur, Van Eltergen, Nierse, & Abma, 2010).²

Focus Group Approaches

---

² We acknowledge the rich tradition in evaluation to train non-researchers (e.g., community members) to conduct focus groups (e.g., Krueger & King, 1997). For the purposes of this paper, we emphasize focus groups conducted by trained evaluators.
What kinds of evidence are gathered in focus groups? As Morgan (1997) notes, focus groups “provide direct evidence about similarities and differences in participants’ opinions and experience” (p. 10). However, Lezaun (2007) argues there is little attention to how opinions are “created, certified, and circulated” (p. 147) in focus groups. Building on Morgan (1997), Kitzinger (1994), and Lezaun (2007), researchers are paying more attention to the nature of knowledge generated in focus groups (Belzile & Oberg, 2012; Farnsworth & Boon, 2010; Kamberelis & Dimitriadis, 2013; Kitzinger, 1994; Markova et al., 2007). Focus group participant interaction is notably distinct in different types of research. A variety of scholars underscore that the design of participant interaction is critical to obtaining a particular kind of focus group evidence (e.g., people’s personal opinions, tacit knowledge; Belzile & Oberg, 2012; Kamberelis & Dimitriadis, 2013; Morgan, 2010, and others).

Scholars propose two distinct orientations that underpin focus group use: an individualist social psychology perspective (Type A) and a social constructionist perspective (Type B) (Belzile & Oberg, 2012; Farnsworth & Boon, 2010; Kamberelis & Dimitriadis, 2013) that guide the character of participant interaction when designing focus group research. Below we briefly summarize focus group characteristics (e.g., theoretical orientation, role of participant interaction, and type of information gathered) that differentiate these two perspectives. These orientations certainly do not reflect the totality of focus group approaches, but simply illustrate two ends of a spectrum. We acknowledge this brief description is a simplification that only partially reflects the theoretical complexities of these views.

**Type A (individualistic social psychology perspective).** Viewed from an individualistic social psychology perspective, opinions are characterized as stable personal dispositions or constructs (Eagly & Chaiken, 2007; Fazio, 2007; Markovà et al., 2007; and others). The
information gathered from a Type A focus group is primarily derived from opinions, based on a person’s thinking and reasoning that is prompted and elaborated in the focus group setting (Belzile & Oberg, 2012; Morgan, 1997). In this approach, the role of participant interaction is to elicit the prevailing range of opinions, beliefs, or preferences regarding a program or policy. The evaluator designs the structure of group interaction and how it will be standardized and managed during the focus group to stimulate and facilitate participants’ own thinking and reasoning in interaction with one another (Morgan, 1997).

While focus group findings are not typically characterized as generalizable, there is a scientific orientation towards replication within this perspective. The researcher/evaluator maintains an objective stance by following a standardized protocol with structured questions. The moderator takes on a more ‘scientific’ role, using robust technical skills to control bias by (a) extracting relevant information through standardized, directive questions while (b) filtering out what s/he considers to be irrelevant information by using group management techniques (e.g., identifying conforming behavior or restricting a forceful focus group member; Lezaun, 2007).

Notably, the moderator does not conduct a ‘series of individual interviews’ within the group setting. To the contrary, in this type of focus group, participant interactions are well-managed by the focus group moderator to encourage verbal exchanges between and among focus group participants (Farnsworth & Boon, 2010; Lezaun, 2007). Since this approach assumes that the information sought is opinions that are basically stable, data analysis focuses primarily on verbal content, with little attention paid to analyzing participant interactions and how knowledge might be socially constructed (Belzile & Oberg, 2012).
Type B (social constructionist perspective). From a social constructionist perspective, opinions are “socially shared knowledge” or tacit knowledge that is generated, maintained, and changed through social participation (Gergen, 1985; Hacking, 1999; Markova et al., 2007, p. 17). Type B focus groups are seen as a dynamic social process, where participants explore opinions, beliefs, and understandings about a program or policy within a group dynamic through a form of collective sense-making (Wilkinson, 1998). It is through the stories participants tell themselves and tell to each other that multiple meanings and the richness of their social world emerge, sometimes in surprising ways. Under these circumstances, knowledge or information is constructed from shared ideas, opinions, beliefs, experiences and actions.

Although norms of civil conduct and exchange are maintained by the moderator, the structure of participant interaction for this type of focus is configured by the researcher to be free-flowing, to allow participants to activate and even build collective experiences and memories about their social world. The group dynamics, social interactions, and social relations that emerge during the focus group help to clarify and reveal what is hidden, but often understood by participants and sometimes by researchers (Farnsworth & Boon, 2010). Social relations involving occupations, gender, age, etc., may significantly affect how participants engage with each other and in the group (Farnsworth & Boon, 2010).

To facilitate natural conversation among the participants and develop the group dynamics and interactions, the moderator’s role is inhibited or subordinated through the use of loosely-structured protocols composed of a few open-ended questions. The researcher/evaluator maintains an empathic—or perhaps political—stance that can vary from (a) breaking down barriers between the evaluator (as moderator) and focus group participants, (b) sharing
responsibility and authority of the focus group with participants, or (c) allowing participants to “take over” or “own” the focus group interview direction and process.

Not surprisingly, Type B focus groups call for a data analysis approach that deliberately attends to both what (content) and how (group interaction) participants talk (Farnsworth & Boon, 2010). From the perspective that people’s opinions are not stable personal constructs but something generated, maintained, and/or changed through social interaction, data analysis in this type of focus groups attends to who said what, in what context, and when.

**Designing and Implementing Focus Groups**

In practice, focus groups will reflect Type A and Type B approaches to varying degrees. In the following sections, we present three vignettes from face-to-face focus groups that we conducted in healthcare and education to illustrate how distinguishing features (e.g., types of information) are important to the design and conduct of focus groups in evaluation. These vignettes were chosen from 80 face-to-face “talk” focus groups we have conducted over the past decade, covering a variety of topics and research goals. Although we draw from traditional face-to-face focus groups, the focus group characteristics we discuss (e.g., type of evidence, plans for eliciting participant interaction) will likely be key issues to consider in designing focus groups for other focus group modes (e.g., on-line, computer-assisted telephone).

These three focus groups were each conducted for a different purpose (e.g., build theory). Vignette 1 resembles the Type A focus group approach. Designed to probe the range of participants’ responses to survey questions, the purpose was to inform our development of a survey to evaluate a statewide education policy implementation administered to teachers and principals. Vignette 2 is a hybrid blending both Type A and Type B and illustrates how focus groups can be employed as a traditional qualitative method. Rich descriptions about university
students’ health experiences were elicited to develop preliminary program theory to be utilized in developing health programs. Vignette 3 reflects the Type B approach. It was conducted to fill knowledge gaps about key issues in proposed statewide educational accountability changes. Teachers shared their knowledge and perspectives as a narrative about potential issues with a new statewide teacher evaluation system that incorporated student test scores to hold teachers accountable for student achievement. Table 1 provides a descriptive framework and summarizes distinctions between vignettes, based on characteristics we have cited (e.g., type of information, theoretical orientation). Below, we highlight a sampling of these distinctions in our discussion of the vignettes.

[Insert Table 1 about here]

Vignette 1 (Type A): Scoping Focus Group

Scoping focus groups are used in questionnaire design to study the range of participants’ responses (perceptions and understandings) of concepts being assessed (Kaplowitz, Lupi, & Hoehn, 2004). We implemented this scoping focus group to facilitate survey development for a statewide evaluation of the implementation of the No Child Left Behind (NCLB) and Race to the Top (RTTT) accountability policies (NCLB, 2002; U.S. Department of Education, 2009). After piloting and refinement, the survey was to be administered to a stratified random sample of teachers to develop a broad description of the changes in instruction, local assessment practices, etc. that teachers attribute to the state’s NCLB testing. The purpose of this scoping focus group was to generate potential hypotheses regarding how regular education, special education, and bilingual education teachers understood and interpreted survey questions. The questions we tested were selected to explore differences in how teachers in different subject areas (e.g., special education, regular education) understood and interpreted questions assessing concepts and domains (e.g., instruction vs. local assessment practices) and responded to various item formats.
We assumed that participants came to the focus group with their own ideas about changes in education due to state NCLB accountability testing. Using a Type A approach, our primary aim was to stimulate various aspects of participants’ thinking and reasoning in responding to each survey question (*type of information*; see Table 1). Survey items were used as stimulus materials. Thus, this kind of focus group is, to some extent, similar to focus groups conducted in consumer research involving product evaluations or testing, where basic information about a product is collected.

This focus group took place in a small Midwest community, at Gere Elementary School (grades K-5). Five teachers who exhibited the characteristics of interest were recruited to participate. The group included two bilingual education, two special education, and one regular education fifth-grade teacher. At the time the focus group was conducted, 44% of the school’s population were students of color, with Hispanic students representing the largest minority subgroup; 10% were English language learners; and 10% were in Special Education. The school had not met Adequate Yearly Progress (AYP) since 2006 because of subgroup test performance (e.g., Limited English Proficiency students), and was now at risk for sanctions (e.g., students would be allowed to transfer to another public school in the district).

An evaluator conducting a Type A focus group typically takes on an objective *stance*. In our example, the *structure* of the focus group was largely managed by a moderator who used a protocol composed of standardized questions and probes to enhance replicability. To successfully elicit the full range of item responses, the management of the focus group was critical to ensuring participants could express their views about the items if they wanted to do so.

After the moderator distributed the survey questions (stimulus materials), participants were given ten minutes to answer these questions as if they were alone at home. The moderator
then initiated discussion of a pre-determined topic with the directed question, “What were you thinking about when you decided how to answer these [survey] questions?” The moderator’s follow-up questions were similarly constructed. The initial focus group instructions were standard—encouraging participants to ‘pass the conversation ball’ among themselves so participants’ views were not prompted just by the moderator. When necessary, the moderator probed or interrupted the conversation to ensure that all participants’ views were heard, to elicit additional perspectives, or to explore possible conformity, while otherwise maintaining a neutral distance.

The note-taker’s main role in a Type A focus group is to record each participant’s verbal responses; however, salient issues would also be noted, such as the moderator’s intervention to directly or indirectly manage group dynamics. Although the focus group interaction is important for elaborating individuals’ opinions as Table 1 suggests, data analysis is primarily focused on the content of participants’ statements. Routine non-verbal communication and participant interactions are presumed to have limited impact on stable personal opinions, which evaluators seek to learn about, and are thus not the subject of data analysis.

**Scoping focus group analysis.** Closer examination of our scoping focus group reveals how the group dynamic elicited personal opinions (*type of information*). There was both agreement and disagreement in teachers’ opinions about the consequences of accountability testing and individual survey items. However, differences (or agreements) of opinion in answering survey questions appeared to reflect participants’ own personal perspectives more than teaching subject expertise. Similar patterns of agreement and disagreement were found across survey questions, which assessed views of changes in instruction, the teaching profession, local assessment practices, and other topics.
Figure 1 presents an example of the type of survey questions that each participant answered prior to the focus group discussion. Following Figure 1 are excerpts of the moderator’s prompt and participants’ brief discussion of the Part A question, which asked whether the participants saw an increase in the use of benchmark assessments.

Moderator: Moving on to question 2, what came through your mind when you were answering? Anyone can start and we’ll go around.

Mr. S. (special education): We just had training for the new Discovery assessment that is a predictive assessment for the accountability test.

Others: uh-huh

Ms. J. (bilingual education): It seems we do a lot of benchmark assessments.

Others: uh-huh

Ms. H. (bilingual education): I would agree with that.³

As this excerpt illustrates, focus group participants across regular, special, and bilingual education provided basic information in the spoken or ordinary language of their everyday life as teachers. With benchmark assessment explained through example (e.g., Discovery assessment is a predictive assessment), the teachers concurred in their opinions about what defines benchmark assessment. Further, they agreed that benchmarking was increasing as a component of local assessment practices utilized in their school.

Reflecting the structured design of this focus group, the excerpt shows how the moderator’s initial prompt not only set the parameters of what participants would discuss about an item, but also, the manner in which the discussion would proceed. In stating, “we’ll go around,” the moderator essentially directed participants to use a turn-taking approach and

³ Data analysis of this kind of focus group data typically focuses on individual quotations. However, this conversational excerpt was used to illustrate the similar interpretations and agreements across these teachers.
determined how turn-taking would take place. Such a directive approach was intended to prompt all focus group participants to participate and to discourage dominance by select participants.

In contrast to the agreement captured above, teachers expressed a wide range of perspectives regarding reasons for the increase in local assessments (see Figure 1, Part B). In the example below, we see some considered it an aspect of good teaching while others thought it was more or less driven by accountability demands. Differences in interpretations, however, did not seem to be a function of the teachers’ diverse teaching backgrounds but rather, differences in their opinions.

Moderator: OK, question 5, to what extent were these changes a result of the state NCLB accountability test?

Mr. S. (special education): I think assessment is just a component of good teaching. I focus more on the district curriculum when making instructional decision.

Ms. B. (regular education): I think you might be naïve

Ms. J. (bilingual education): Big time! I think it’s a lot to do with preparing for the state test.

Ms. R. (special education): Well, it depends on how you look at it I guess…if you are cynical everything is because of accountability assessment, which is how I think. But maybe some of the changes are just what’s best for kids.

This focus group revealed a range of personal opinions about the extent to which teachers attributed changes in local assessment practices to accountability testing. The information yielded was taken to be straightforward and needing minimal theorizing. As Ms. R. said, “it [your opinion] depends on how you look at it”; perceptions about the changes due to NCLB accountability were largely about persons’ viewpoints. The range of views was due to participants’ different ways of viewing the link between changes in local assessment practices and accountability testing.
Strengths and limitations. As illustrated by Vignette 1, the Type A approach is valuable for obtaining evidence based on personal opinions and perceptions. This approach is a fairly efficient means for gathering basic qualitative information about issues of interest and for generating hypotheses for further testing. For example, in evaluation, this kind of focus group (Type A) is recommended for gathering information about participant satisfaction, to generate hypotheses about the effects of programs and policies being evaluated, etc. (Centers for Disease Control and Prevention, 2008).

The appropriateness of this focus group approach depends on the research and/or evaluation purpose and questions. The focus group setting allows for conversations that encourage elaborations, agreements, and disagreements among participants that reveal the range of responses to a specific issue. The moderator closely manages the focus group structure to ensure that pertinent information is obtained, while allowing naturalistic interactions to gather abundant data from multiple participants quickly. Focus groups like Type A have been cited as a cost-efficient means of gathering and analyzing information without the need for individual interviews that would require more time and labor-intensive analysis (Krueger, 1994; Stewart & Shamdasani, 1990).

Analysis of the data from this kind of focus group will primarily examine speech content, so important information might be missed. Even the most skilled moderators might not be able to successfully manage group dynamics, so the full range of participants’ views are not revealed. The structured character of the Type A approach, which includes standardized protocols and directive questions, also has drawbacks. There are concerns about the extent to which the moderator and focus group participants will attribute the same meanings to the concepts referenced in the interview protocol (Fontana & Frey, 2005). To address this issue, meticulous
pilot testing of the interview protocol (which would require additional expenditures), becomes essential for producing credible focus group evidence.

**Vignette 2 (Hybrid): Theory-building Focus Group**

Focus groups are often used to gather rich descriptions about meanings, processes, and experiences from participants’ points of view (Jarrett, 1993). For example, in a recent mixed-methods assessment of the health education needs of engineering and science (STEM) students at a large Midwestern land-grant university, we conducted focus groups to follow up on preliminary results from a large-scale survey about students’ health practices. Specifically, our purpose was to develop rich description to learn more about STEM students’ common health experiences in the three areas that their survey sub-scores differed significantly from non-STEM students’ scores: nutrition, physical activity, and attitudes about depression.

What is it about being a STEM student in a highly competitive Research I institution that generally leads to students’ poor nutrition and sedentary lifestyle when compared to students in other fields at the same university? In conducting these focus groups, we employed specific techniques to elicit rich descriptions from participants that would help us develop a more elaborated understanding or preliminary theory about STEM students’ health experiences. A richly described preliminary theory was intended to inform the creation of campus health education programs targeted toward STEM students.

Rich description includes both participants’ personal opinions and their collective experiences that are articulated together during a focus group (*type of information*; see Table 1). Focus groups that are designed to yield rich description often reflect features of both Type A and Type B approaches. Focus group researchers have proposed that mechanisms for eliciting the meanings of participants’ subjective experiences include (a) structuring focus groups to enhance
disclosure and (b) creating interview protocols to access participants’ own language and concepts, especially around sensitive topics (Jarrett, 1993; Morgan, 2012; Wilkinson, 1998).

Findings from previous STEM research suggest that female students’ educational experiences—and therefore, discussion of these topics—might be of a sensitive nature. STEM educational paths and careers have long been considered to be male-dominated domains (e.g., http://www.nacada.ksu.edu/Clearinghouse/advisingissues/STEM-gender.htm). In consideration of these issues, we took efforts to safeguard participants’ comfort and security in discussing these topics. Four focus groups were planned, each with an identical design but conducted with different segments of the STEM student population: undergraduate men, undergraduate women, graduate men, and graduate women. The example presented below is drawn primarily from the focus group conducted with undergraduate women (N=9).

In contrast to the scoping approach, where the moderator tries to remain objective to minimize biasing the group, the moderator in a theory-building focus group tends to take an empathic stance to purposefully break down barriers between the “researcher and researched” (Jarrett, 1993). As Table 1 suggests, the structure of this hybrid focus group is a mix of the Type A and Type B approaches with the goal of encouraging a semi-structured conversation among participants.

For example, the moderator began with a question about participants’ general views about health: “What does being healthy mean to you? What are some examples of being healthy/unhealthy?” A broad question such as this allowed focus group participants to describe their experiences using their own language. At the same time, the facilitator maintained some control over the conversation and asked planned semi-structured questions; she probed for more detail about key topics, and directed the flow of conversation to include a variety of voices.
As with Type A focus groups, the note-taker’s role was to record what was said. In addition, the note taker in this hybrid focus group was alert to group interactions that could be relevant to the analysis. Reflecting the Type B approach, as Table 1 shows, some of the interactions between participants were considered to be ‘data’ (such as when participants emphasized others’ comments) to underscore the shared nature of participants’ experiences.

**Theory-building focus group analysis.** Overall, findings from the four hybrid focus groups revealed at least one reason why STEM students generally had poorer health experiences than non-STEM students: A rigorous curriculum and pressures to perform led students to spend a significant amount of time on school work, crowding out healthful behaviors. As one student said, “Usually being in Engineering you tend to get overwhelming workloads, and it may be hard to take the time to relax and not just be studying.” We found that the female undergraduate focus group was distinguishable from other groups by the intensity of the discussion and the rich examples the women shared about their lives. From the undergraduate female perspective, the choices they made about time demands were guided by an implicit (and sometimes explicit) hierarchy of priorities. Work, personal hygiene, sleep, food, and physical activity were their priorities, in that order. The short excerpt below provides a glimpse of their thinking about work and food over physical activity:

Moderator: What do you think would lead you to engage in physical activity (not just exercise) more consistently?

G: It’s hard [to make it to the gym]. If you’re not committed, then you’re not going to go ever, and […] school trumps working out, and you’re like “[…] I feel like I am being sucked into a black hole.”

D: Yeah, I feel like if I don’t have strict commitment to someone else […] then I am not [going to go]. And it’s like an hour and a half lost, because then I have to shower […] so it’s like two hours lost that I could be finishing up my lab report, or I could be eating dinner, so….
Below a female undergraduate described how she justifies working out at the gym only by studying at the same time suggesting “work trumps all”:

C: If I’m going to do something like jogging on a treadmill, I have to take time to make sure I am […] actually doing work. So, I’ll take my notebook and have it open while running and be like, “Okay, I am studying Chemistry while being active.” So, as long as I can multitask—because if I feel like I am doing physical activity with no other benefit […] I am far too worried about classes to keep doing it.

The empathic role of the facilitator was particularly evident during the discussion of depression when she departed from the pre-written protocol, which broke down barriers between herself and the participants:

Moderator: Just listening to you talk—the best thing I have ever used [the Counseling Center] for is to call at 7:15 and make the appointment […] it’s a safe place to just get it out.

L: That’s kind of what I do with my advisor—I just go and vent once a week.

Moderator: That is fortunate to have an advisor like that… I just wanted to point that out. But on that note, [a focus group participant] said [Counseling Center staff] haven’t been through these [STEM] classes […] do you resonate with that idea?

Although a focus group moderator often provides only enough commentary to keep the conversation going, in the excerpt above, the empathic interviewer interjected her own experiences (e.g., about calling the Counseling Center) to connect with the participants as a fellow student, who also had all the stresses of academic life. Nevertheless, she maintained an authoritative position by not letting the discussion stray too far away from the semi-structured format and redirecting the conversation with a probe about perceived differences between the life experiences of counseling staff and STEM students.

**Strengths and limitations.** This focus group design blends features of the Type A and Type B approaches. The f vignette offers only a small sampling of the thick description derived from conducting focus groups with different segments of STEM students when using a semi-
structured interview protocol. Consonant with the hybrid approach, our analysis integrated participants’ personal opinions with the common grounds of their shared experiences, going beyond the simple content analysis of the Type A approach but not relying heavily on narrative as in the Type B approach.

This approach led to evidence in the form of rich description suitable for developing general understandings or preliminary theories of social phenomena. The descriptions mined in focus groups like this one are well suited for disclosing tensions or contradictions in participants’ opinions about complex issues such as health status. These kinds of descriptions reveal how individuals talk about particular concerns and can show people’s reasoning about their experiences and choices. In combination with relevant research, this kind of focus group can be utilized for developing and refining program theory. By contributing to theory-building efforts, this approach is useful for designing programs that target a particular population or population subgroup by using language familiar to the population and addressing the needs inferred from their own life descriptions.

The evidence derived from this kind of information comes at some cost. Simple content analysis (as in the Type A approach) will not be sufficient to develop a detailed understanding of participants’ experiences. Data analysis will require a substantial time commitment from a moderately-skilled qualitative data analyst. Empirically grounded theory-building is typically based on a grounded theory approach (or similar qualitative analysis technique) that requires sophisticated analysis to develop and identify concepts from the data that will be accepted as credible evidence (Strauss & Corbin, 1988).

**Vignette 3 (Type B): Narrative Focus Group**
This narrative focus group was conducted as part of the four-year mixed methods NCLB and RTTT policy implementation evaluation previously mentioned (NCLB, 2002; U.S. Department of Education, 2009). The survey was employed to help us describe the actions of a large number of teachers (asking, for example, what are teachers’ instructional and assessment practices in response to NCLB). Using the Type B approach described in Table 1, the purpose of this type of focus group was to investigate how or why questions and to fill in gaps in knowledge about key issues (e.g., why are teachers using these assessment and instructional practices?). In the survey we conducted in Spring 2011, teachers reported increased stress and fears about RTTT teacher evaluations that were incorporating student performance gains (specifically, students’ test scores) to make judgments about teaching quality. For example, they expressed concerns over teacher dismissals, which the teachers viewed as due to circumstances beyond their control (e.g., state test performances that do not accurately reflect student learning).

Unlike the Vignette 1 scoping focus group that gathered basic information, in the Fall 2011 narrative focus group we studied the local community version of teachers’ knowledge about the emergent teacher evaluation policies. To uncover this local knowledge, the same interview protocol was administered to teachers in several schools, each contributing different historical, social, cultural, and achievement dimensions to our study. Although no new teacher evaluation models were actually being implemented, we were interested in teachers’ efforts at collective sense-making through their social interaction about these emerging teacher evaluation policies (type of information). We wanted to hear the stories teachers were telling themselves and each other as they dealt with uncertainty about new teacher evaluations. In addition, we wanted to compare how these stories were similar or different within and across schools.
This particular focus group took place at Big Grove Middle School (grades 5-8) and included five experienced white English and math female teachers; one was current co-president of the local teacher’s union. The school is located in a small Midwestern town with a population of 6,000 that is 85% white, with approximately 50% low-income students. For the first time, Big Grove Middle School did not meet 2011 NCLB annual yearly progress targets.

The focus groups were structured to reveal the interactions (e.g., shared understandings, tensions) and dynamics among focus group participants, not just the content of conversations. The moderator opened the focus group by saying, “Let’s now get to this teacher evaluation. What do you think about it?” Subsequently, although the moderator occasionally made a comment or asked a question, he played in a secondary role so that participants could interact naturally. Food and drink were shared during ice-breaking activities to create an informal, festive occasion. As Table 1 suggests, the emergent nature of the conversation (structure) and the ancillary role of the moderator in the Type B focus group were notably different from interactions in the Type A and hybrid focus group approaches.

In addition to audio-taping, the note-taker paid close attention to non-verbal communication (e.g., tone, group agreement, humor). As Table 1 suggests, unlike the Type A focus group approach that prioritizes content, a group’s interactional and relational dynamics are crucial sources when analyzing Type B focus group data. Therefore, in our data analysis and interpretation, we noted how participants added to, or ignored certain comments to help us understand the group’s dynamic and/or how dominant and collective opinions were formed while other perspectives were suppressed.

**Narrative focus group analysis.** While teachers expressed some fear and uncertainty about the new teacher evaluation model, they also constructed a distinct narrative that
externalized teacher evaluation on two dimensions: (a) teacher evaluation is seen as beneficial for other teachers, but a stressor for themselves and (b) teacher evaluation is necessary because of the shortcomings of university teacher programs. Overall, teachers built on each other’s opinions and beliefs through their conversations that made transparent the emerging group dynamics and social relations (e.g., power-related status). Both externalization and social relations issues are illustrated in the excerpts below.

Teacher M: I don’t think…we’ve really been given enough information to have a complete opinion on that [teacher evaluations]… a lot of us are sitting back, maybe fear’s not the right word, maybe it is [laughs]. We don’t know what’s coming….

Teacher K: You know, I don't know if they [teachers] had them [teacher evaluations] when I was in school. We had teachers who definitely didn’t need to be teachers…I had one guy who read a newspaper during class and he would tell you ‘read chapter 2 of the book’ while he sat with his feet up on the desk…as far as learning anything, it was self-taught.

Teacher J [Union co-president]: Does the bar need to be high? Yea. [agreement from the other teachers]…You need to keep abreast of what’s going on… they’re tryin’ to go after the people who do not evolve.

Teacher L: And you and I need to learn how to use our SmartBoards [laughs]

Although brief, this excerpt hints at how the narrative revealed a collective knowledge and identity relevant to teacher evaluations; for example, when Teacher M says, “I don’t think we’ve…” Further, the excerpt reveals how teachers (e.g., Teachers K and J) respond to each other’s comments as they elaborate on why teacher evaluations are beneficial for some teachers by sharing stories. Note how the moderator did not direct the discussion but instead allowed teachers to share and build on their own rich descriptions and examples.

Importantly, the teachers discussed needs for their continuing education and for removing teachers who were unwilling to change or unprepared to teach. While acknowledging their own limited SmartBoard skills (an educational technology) through laughter, they focused on the need for other teachers to continue their education and to improve. The focus group dynamic was
critical. Teacher J, the Union co-president, played a major role in framing the teacher evaluation issue when she made the statement, “people who do not evolve.” These dynamics are also evident in the following example, where, distancing themselves further, teachers moved the discussion to issues with teacher education. Again, Teacher J named teacher education as a university’s responsibility; Teacher K elaborated on the idea with examples.

Teacher J [Union co-president]: But… it’s making us scramble…it goes back down to Universities’ responsibility where the bar needs to be set higher there…

Teacher K: …You don’t really get into hands-on in classroom until you’re student teaching … They need to push down the experience in classroom earlier so they [teachers] know what they’re getting into.

The social dynamics within a Type B approach are seen as providing insights or information, in contrast with a Type A approach, where such interactions are controlled as a source of bias. The above excerpt suggests that Big Grove Middle School and the teachers’ union (and its representatives) may well be influencing teacher beliefs and opinions about the emerging teacher evaluation policies—as is happening in some other districts (e.g., the Chicago Public School District; Rossi, Fitzpatrick, Esposito, & Spielman, 2012). Interestingly, we did not find this same narrative in focus groups we conducted at an urban (lower-achieving) and suburban (higher-achieving) school.

**Strengths and limitations.** The Type B focus group approach supports the uncovering of key issues firstly, by paying equal attention to what people say and what they do not say. Especially when participants have a great deal of tacit knowledge about an issue, what goes unsaid may be as revealing as verbalized ideas. Secondly, the ancillary role of the moderator, and the unstructured protocol, expedited disclosure of how participants (teachers, in this case) make collective sense of, as well as their knowledge about, issues they identify as vital.
The flexible nature of the Type B focus group approach creates a democratizing environment that encourages participants to bring up and discuss topics that might not be elaborated on or even verbalized within other, more structured focus group designs. In many cases, participants can provide context-based reflections based on first-hand experience with the problem that policymakers are intending to address (e.g., teacher evaluation). In adding to our understandings of how participants’ perspectives are shaped, findings from this approach can provide insights about the potential consequences of controversial policies and point to critical stumbling blocks or needed incentives to facilitate policy or program implementation.

The type of evidence collected in narrative Type B focus groups may be particularly useful in policymaking and other domains that have traditionally relied heavily on statistical and other forms of data that are valued for their technical qualities (e.g., reliability, validity; Epstein, Heidt, & Farina, 2012). Stakeholders may have deep knowledge about “facts, causes, interrelationships, and likely consequences” of a local or national policy (p. 7). Moreover, stakeholders use narratives, not technical data, as the primary means of supporting their positions when evaluating policies. Researchers are just beginning to investigate how non-standard forms of evidence such as narratives may be utilized as evidence for making claims in policy-making (Epstein et al., 2012). Currently, narratives are recognized for their value in helping researchers acquire deeper and richer perspectives about prevailing expert knowledge (Collins & Evans, 2007; Sole & Edmondson, 2002).

The data yielded from the Type B approach are rich; however, it is both time-consuming and expensive to analyze discourse or narratives. Narratives may be collected as a data pool, used to develop taxonomies and categories and analysis may also involve gathering events to construct a single narrative or set of explanatory stories (Polkinghorne, 1995). Employing the
narrative focus group approach should be carefully weighed in relation to the financial and material resources that are typically available for conducting evaluations. Nevertheless, as a distinct form of local knowledge (e.g., social, practitioner), the character of the information gathered is difficult to access with other methods.

**Discussion**

Hollander (2004) argues that “focus groups may be best conceptualized as a ‘research site,’ not a research instrument” (p. 631). Focus groups, which rely on group processes, offer a different view of social interaction than do individual-oriented methods (Solano, 1988). Nevertheless, how to capitalize on the choice to work with groups instead of individuals in gathering evidence remains underdeveloped. Focus group theorists’ debates on how to circumscribe the role of participant interaction in focus groups are well-rehearsed (Kitzinger, 1994; Morgan, 2012). Yet, the character of participant interactions are typically not discussed in research or evaluation studies that report on focus group findings despite their importance in determining the nature of evidence that focus groups generate (Belzile & Oberg, 2012; Webb & Kevern, 2008).

As we have illustrated in this paper, the researcher/evaluator is strategic in configuring focus groups to generate a distinct kind of evidence. Building on recent developments in social science methodology (e.g., Belzile & Oberg, 2012; Farnsworth & Boon, 2010; Morgan, 2012), we describe how different research designs are required to gather different *types of information* with focus groups. Decisions about the focus group design such as moderator’s *stance* (objective, empathic, or ancillary), and *data analysis approach* (content- and/or interaction-focused) are critical for obtaining a particular type of evidence (see Table 1). The vignettes presented show how the kinds of interactions between and among focus group participants (and moderator) signal notable differences in the kinds of research conducted and information being collected.
To maximize what evaluators learn from focus groups and improve the credibility of focus group evidence, we propose that evaluators critically reflect on some initial core questions to guide their decisions about focus group design and implementation. Aligning focus group features to best match the inquiry purpose can enhance the rigor and value of this method. Evaluators will be able to explicate the logic and reasoning behind the focus group approach they select, as well as justify their design decisions and the kind of evidence they gather. Some core questions for the evaluator to consider when conceptualizing a focus group and how the interaction will be handled in analysis and reporting include:

- What is the purpose of the focus group (gathering basic information, theory building, empowering stakeholder participants, etc.)?
- What capacities (e.g., evaluator skills) and resources (human and financial) are available?
- What type(s) of information are to be obtained (personal opinions, collective experience, etc.)? What is the nature of participants’ knowledge about the topic of interest?
- What is the role of participant interaction?
  - How is the focus group to be structured?
  - What will the moderator’s stance be in relationship to the participants (e.g., objective, empathic)?
- How will focus group data be analyzed and reported?

Reflecting critically on these questions will help evaluators identify the approach most consistent in purpose and type of knowledge, implementation, and analysis.

A thoughtful focus group design, however, does not ensure seamless implementation. As practitioners, we freely acknowledge that focus groups are interactive and communicative events that sometimes do not unfold as planned by even the most skilled evaluators. The dynamic nature of this method can yield results that are inconsistent with the intended goals of the selected approach. We have found it highly beneficial to conduct a critical reflection on how the focus group was actually deployed to determine how it may have diverged from the original design and
what results were actually obtained. We identify three issues that we have found helpful in scrutinizing focus group implementation, and present them as questions:

- Did the focus group participants establish common ground in conversation or primarily act as individuals?
- What were the power dynamics between the moderator and participants, both as a group and as individuals? What were the relations among the participants—collective or dominant?
- What were the participants using the focus group for?

For example, a moderator who tries to draw out participants’ shared experiences could fail to establish common ground (e.g., teachers from the same school end up speaking as individuals; see Hydén & Bülow, 2003 for fuller discussion about this challenge). In addition, as Belzile & Oberg (2012) note, researchers often overlook participants’ use of focus groups, which may not correspond with the research goal. One of the authors (Gandha) experienced “losing control” over a theory-building focus group about NCLB assessment consequences—the participants “completely took over.” Although the moderator came with a semi-structured protocol, participants’ responses to the questions posed were superficial. The group chose to use this time to discuss topics that were not targeted in the protocol. They reminisced about the kinds of rich curriculum they used to offer to the (high-achieving) students and why their current (low-achieving) students could not benefit from a rich curriculum. The focus group did yield rich data, but not on the topics the evaluation team aimed to learn about.

Methodological decisions in a substantial number of evaluations are also influenced by resource availability and limitations (i.e., budgets and contracts determine how many focus groups are possible and how much time is available for analysis). Other practical issues such as access to individuals that meet selection criteria and the time constraints of those individuals may also restrict the type and depth of information collected. As illustrated in this paper, some
evaluations might warrant the greater money and time expenditure required in gathering and analyzing both the content and interaction dimensions of focus group data (e.g., when investigating a more abstract topic such as stakeholder values in a multiracial evaluation context). Regardless of the focus group choices made for a particular evaluation, understanding the potential complexity and nuances of different focus group approaches, as well as focus group evidence, is an important component of analyzing and interpreting data and augmenting evaluators’ learning from focus groups.
References


Farnsworth, J., & Boon, B. (2010). Analysing group dynamics within the focus group. Qualitative Research, 10, 605-662.


Kamberelis, G., & Dimitriadis, G. (2011). Focus groups: Contingent articulations of


Kitzinger, J. (1994). The methodology of focus groups: The importance of interaction between research participants. *Sociology of Health & Illness*, 16, 103–121.


Table 1

*Descriptive Framework for Focus Group Design Characteristics and Evidence*

<table>
<thead>
<tr>
<th>Theoretical Perspective</th>
<th>Type A approach</th>
<th>Hybrid approach</th>
<th>Type B approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[Scoping focus group]</td>
<td>[Theory-building focus group]</td>
<td>[Narrative focus group]</td>
</tr>
<tr>
<td>Purpose or use</td>
<td>Individualistic social psychology</td>
<td>Mixed-opinion/ experiences</td>
<td>Social constructionist</td>
</tr>
<tr>
<td></td>
<td>Pretest, hypothesis-generating</td>
<td>Build mid-range theory, constructs</td>
<td>Fill in gaps—how and why questions, empowerment</td>
</tr>
<tr>
<td>Type of information</td>
<td>Stable personal opinion</td>
<td>Mixed-opinion/ experiences</td>
<td>Social and/or tacit knowledge</td>
</tr>
<tr>
<td></td>
<td>Stimulate and elaborate personal opinion</td>
<td>Generate mix of personal opinion and collective experiences</td>
<td>Facilitate collective knowledge building</td>
</tr>
<tr>
<td>Role of participant interaction</td>
<td>Standardized, replicable, directive, predetermined</td>
<td>Mixed with semi-structured conversation</td>
<td>Non-standardized, variant, emergent, spontaneous, natural conversation</td>
</tr>
<tr>
<td></td>
<td>Scientific neutrality/ perhaps technician</td>
<td>Empathic interviewer with authority</td>
<td>Ancillary and/or political</td>
</tr>
<tr>
<td>-Structure</td>
<td>Content-oriented analysis</td>
<td>May mix or merge interaction with content; grounded theory analysis</td>
<td>Narrative analysis</td>
</tr>
<tr>
<td>-Evaluator stance/role</td>
<td>Basic information Simple qualitative description</td>
<td>Rich description Preliminary program/policy theory</td>
<td>Narratives</td>
</tr>
<tr>
<td>-Data analysis</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Since NCLB testing began in grades 3-8, how have local assessment practices changed? To what extent were the changes a result of NCLB testing? [Please rate each item on both scales.]

<table>
<thead>
<tr>
<th></th>
<th>(Part A)</th>
<th>(Part B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>This practice has...</td>
<td>To what extent was it</td>
</tr>
<tr>
<td></td>
<td>Increased a lot</td>
<td>a result of NCLB</td>
</tr>
<tr>
<td></td>
<td>Increased somewhat</td>
<td>testing?</td>
</tr>
<tr>
<td></td>
<td>Not changed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreased somewhat</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decreased a lot</td>
<td></td>
</tr>
<tr>
<td></td>
<td>We don't use this strategy</td>
<td></td>
</tr>
<tr>
<td>c. Benchmark</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>assessments</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>□</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 1.* Example survey question used as scoping focus group stimulus material.