



# Experience-Sampling Research Methods and Their Potential for Education Research

Sabrina Zirkel<sup>1</sup>, Julie A. Garcia<sup>2</sup>, and Mary C. Murphy<sup>3</sup>

Experience-sampling methods (ESM) enable us to learn about individuals' lives in context by measuring participants' feelings, thoughts, actions, context, and/or activities as they go about their daily lives. By capturing experience, affect, and action *in the moment* and with repeated measures, ESM approaches allow researchers access to expand the areas and aspects of participants' experiences they can investigate and describe and to better understand how people and contexts shape these experiences. We argue ESM approaches can be particularly enriching for education research by enabling us to ask new and interesting questions about how students, teachers, and school leaders engage with education as they are living their lives and thus help us to better understand how education contexts shape learning and other outcomes. In this article, we highlight the value of these approaches for addressing new and exciting questions they may help education researchers to answer as they allow us to uncover experience in new ways.

**Keywords:** experience sampling; longitudinal studies; research methods; research design

Experience-sampling methodologies (ESM), sometimes referred to as “ambulatory self-report” (e.g., Conner & Barrett, 2012) or “intensive-longitudinal designs” (e.g., Bolger & Laurenceau, 2013), examine individuals' experiences and behavior in context. The essential feature of ESM is that participants' feelings, thoughts, actions, context, and/or activities are measured repeatedly as they go about their daily lives. By allowing us to track individuals' thoughts, feelings, and actions as they are experienced, ESM takes an expanded epistemological stance regarding what we can know and learn about individuals. In ESM studies of subjective experience, participants complete brief surveys about their immediate environment, behaviors, feelings, and/or thoughts several times a day for a period of days or weeks, resulting in dozens or even hundreds of responses per participant (e.g., Csikszentmihalyi, Rathunde, & Whalen, 1997). ESM studies might also include more objective measures of participants' lives and experiences through activity monitors, random sound recordings, or other assessments of their lives as they are lived. With the exception of diary studies and a small number of innovative education researchers (e.g., Hektner, Schmidt, & Csikszentmihalyi, 2007; Schweinle, Meyer, & Turner, 2006; Schweinle, Turner, & Meyer 2008; Shernoff & Csikszentmihalyi, 2009; Strati, Shernof, & Kacker, 2012), ESM

have not been widely harnessed in education research. We argue that their broader use would allow us to expand the range of questions and ideas we can explore. Our purpose in this paper is to introduce ESM approaches to a broad readership of education researchers. We begin by reviewing what these approaches can offer researchers epistemologically and then discuss examples of ESM in the extant literature to highlight what ESM can do in terms of the questions these approaches allow us to investigate. Finally, we discuss the conceptual and technical details on the use of these methods and discuss considerations in their use for those who wish to consider using ESM in their own work.

The focus of ESM is examining experience in context—in situ—as people live their lives. Originally, ESM emerged from a phenomenological standpoint, and the focus of much ESM was and remains understanding individual subjective experience *as it is happening*. Our own work has primarily descended from this tradition, in that we have primarily relied on self-reports of subjective experience and the context in which those experiences occur (e.g., Where are you? Who are you with? What are you

<sup>1</sup>Mills College, Oakland, CA

<sup>2</sup>California Polytechnic State University, San Luis Obispo, CA

<sup>3</sup>Indiana University, Bloomington, IN

doing?). Researchers from this phenomenological tradition focus on subjective experience, typically asking participants to complete brief open- and/or closed-ended surveys at either random or specified times. These surveys often focus on elucidating individuals' subjective state through well-validated survey items and/or brief open-ended questions that can be written on a mobile device, and then these subjective experiences are combined with other data to illuminate the connections between subjective experience and other outcomes.

In our own research, for example, we have asked participants to complete surveys immediately following their college classes, and after identifying which class they were just in, participants answer a series of questions about their experience in that class. Participants' overall affective state "in the moment" was measured through a short version of a well-validated scale, such as the Positive and Negative Affect Survey (Crawford & Henry, 2004). Participants indicated how much the following statements applied to them in, for example, the class that just ended: "During the class that just ended, I felt anxious"; "During the class that just ended, I felt calm" (or sad or depressed/happy/angry/content). Another set of questions concerned the extent to which participants felt worried about whether they appeared "smart" in that class. Such questions included "During this class, I felt worried that I might say the wrong thing (may make a mistake in front of my professor/my professor may judge my intelligence during class/my professor might think I am a slow learner)," with participants making ratings on a 7-point scale with labels from *very untrue* to *very true of my experience in this class* (Murphy, Garcia, & Zirkel, 2014). Such data have allowed us to explore whether certain classes or disciplines (e.g., science, technology, engineering, and mathematics [STEM] vs. non-STEM) consistently make students feel more or less worried about their presentation of self as "intelligent" than did others. We then link those subjective, in-the-moment descriptions of their experiences to education outcomes provided through later surveys and students' transcript data. We find that such concerns are more common among women than among men, and among Students of Color than among White students. Moreover, we find that these group differences are exacerbated in STEM classes, such that women and Students of Color experience a much higher level of concern about appearing "smart" in STEM classes than do men and White students, respectively. How do these subjective experiences affect other important classroom behaviors, such as asking questions, attending office hours, or studying for exams? Preliminary data suggest that the more concern about appearing "smart" women reported, the less likely they were to ask questions or seek help during office hours. Subsequent analyses will reveal whether these experiences impact students' interest in taking more courses in a discipline or interest in certain majors. In many studies, ESM data are combined with other, often more objective, data to assess the relationships between subjective experience and more concrete actions or outcomes. In our case, our ESM study was embedded in a 4-year longitudinal study of college students' experience in which we explore the impact of early classroom experiences on later course choices and a wide range of personal and academic outcomes.

Researchers from other traditions and fields have expanded the use of ESM to include other aspects of experience in recent

years, including participants' movement, behavior, subtle body language in conversation, photographs, and more. Technological advances and have enabled many innovative approaches to "sampling" different aspects of "experience" using other techniques. ESM-generated data are limited only by the imagination of investigators. For example, sociologists have had children take photographs of their neighborhood to "see" through the eyes of young people (Morrow, 2001; Strack, Magill, & McDonagal, 2004). Dieticians have asked participants to take pictures of any food consumed to more reliably assess what, when, and how much they are eating (e.g., Chang & Ko, 2007; Rollo, 2012). Estrin and Sim (2010) have been pioneering the use of data from mobile devices for health research, including the use of GPS to measure the time of day that participants leave the house (which turns out to be a useful measure of wellness for those interested in depression, chronic pain, or other health concerns). Others are turning to the use of electronic sensors to learn when and how people are moving; By measuring exercise in the moment, researchers are better able to understand not just how many calories are burned but when, how, and with whom calories are burned. This provides a fuller picture of movement and health that can inform treatment for diabetes and other health concerns (e.g., see Stanley & Osgood, 2011). Other ESM studies focus on written accounts of experience (e.g., dual accounts of marital arguments; Bolger, Davis, & Rafaeli, 2003) or different kinds of material artifacts of daily life, such as measurements of ambient sound, speech (Buchman, 2007; Mehl, Pennebaker, Crow, Dabbs, & Price, 2001), subtle body movements, and speech patterns or in-the-moment measurements of blood pressure or other physiological readings (Fahrenberg & Myrteck, 2001). The common factor in all these studies is that the researcher is "sampling" participants' experience in some way—to understand what that experience looks like *as it is happening*. As an interdisciplinary field, education researchers are highly creative in their research approaches; we expect education researchers could contribute to the development of new and exciting ways to design ESM studies.

### What Does ESM Have to Offer Education Research and Theory Development?

ESM offers several key epistemological and methodological advantages to education researchers: First, ESM provides access to settings and subjective experiences we would otherwise have no means of probing. Second, ESM offers proximity to participants' experience. Data are collected about participants' lives as they are happening, thus reducing retrospective and other biases. Third, ESM offers the opportunity to acquire descriptions of large numbers of individuals' experiences in ways that can add to that which we can learn through intensive qualitative studies, and fourth, ESM offers the ability to study intraindividual change and processes, placing thoughts, feelings, and behavior in highly specific contexts. Finally, for quantitative researchers interested in issues of statistical power, the repeated-measures nature of ESM approaches affords a level of statistical power that is difficult to acquire in other quantitative approaches.

## *Access to Subjective Experience and Behavior We Could Not Otherwise Reach*

In self-report ESM studies, participants are asked either about what was happening right when the alarm sounded (“What are you doing? Who are you with?”) or about something that just happened (“Please tell us about the class that just finished”). This provides access to information about individuals’ lives that we may have no other way to learn about, both because these actions and events take place “behind closed doors” (literally and figuratively) and because participants are not always aware of patterns in the way contexts shape their own behavior. Larson and Richards’ (1991) ESM study of whom children (ages 9–15) spend time with reveals a marked decline in time spent with parents as children grow older. Teenagers spent about half the time with parents as did younger children, with boys spending this time alone and girls with friends. Even if willing, children and adolescents would probably not have been able to give reliable estimates of time spent in any context, but through ESM, the researchers would be able to reliably access this experience.

Other measures used in ESM research tap information that participants could not reasonably be expected to know about themselves. Mehl and colleagues (2001) developed the electronically activated recording (EAR) of speech in order to record small, regular recordings of participants’ speech. Participants carry the EAR device and it records small snippets of ambient sound, including speech, at set periods of time (e.g., 30 s of speech every 5 min). These recordings are not long enough to acquire enough speech to invade participants’ privacy, but they are enough to provide interesting insights about many aspects of speech and ambient sound (Pennebaker, Mehl, & Niederhoffer, 2003). For example, Holleran, Whitehead, Schmader, and Mehl (2011) conducted an ESM study of faculty speech using the EAR, which was programmed to record 50 s of sound every 9 min between 6:00 a.m. and 11:00 p.m. (approximately 10% of the time). Analysis of speech frequency and content revealed an interesting pattern in the conversations of male and female STEM faculty. First, both male and female STEM professors were less likely to talk about research when speaking to a female colleague. Second, self-report assessments of participants at regular intervals revealed that women were more likely feel disengaged after talking about research with a male than with a female colleague. Finally, listeners’ reactions to these speech recordings of male and female STEM professors’ speech indicated that female STEM faculty were rated as being less competent when discussing their research.

### *Proximity to Experience*

By assessing participants’ experience “as it is unfolding,” ESM research affords a tremendous reduction in retrospective bias inherent in all self-report data (e.g., Kahneman & Krueger, 2006). In other words, experience-sampling data reflect participants’ in-the-moment affective, cognitive, and behavioral experience in a way that preserves the immediacy of the moment and gives participants less opportunity and motivation to present themselves as they might when reflecting backward on experiences occurring earlier in time. For example, individuals typically

misremember how much time they spent in or how they felt during different activities and contexts. When the immediacy of the moment is over, recollections of the experience are shaped and reshaped by a number of cognitive and reflective processes (Kahneman, 2011; Kahneman & Krueger, 2006; Wilson, Meyers, & Gilbert, 2003). Although self-reported ESM data cannot eliminate all of the potential biases of self-report, research suggests that by asking people to report on their activities, affect, and actions in situ and on many small occasions, we may be able to get a more accurate picture than when we ask participants to reflect backward over a period of time (e.g., Kahneman & Krueger, 2006). Asking participants how they are feeling while they are, in fact, studying, for example, might reveal frustration or boredom hidden when we ask those same participants how they felt while studying some time ago or how they feel “while studying” generically. Conner and Barrett (2012) advance the idea that the self-report time frame of ESM reports (what are you doing/thinking/feeling “right now”) taps into a different “self”—they term it the “present self”—compared to the more retrospective self-reports we often rely on in social science research, which instead capture a more “retrospective self.” Conner and Barrett (2012) note that tapping into this present self can be an informative addition to our research design repertoire, eliminating many of the filters that color our retrospections about our lives.

Proximity to experience also allows researchers to use ESM to acquire more accurate assessments of how people spend their time. When ESM is used to randomly sample participants’ activities at different times of day, we can observe how often we actually “find” them studying, and evidence across a wide range of studies suggests that this estimate of activity is more accurate than are retrospective assessments of how they spend their time. For example, this has proven a valuable tool in studies of how students spend their time in school (e.g., Leone & Richards, 1989; Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003; Shernoff, Knauth, & Makris, 2000; described in more detail below). Of course, ESM approaches that rely on objective, rather than subjective, measures eliminate self-report bias altogether. Such studies include, for example, those that measure location or physiological responding, such as physical activity through an activity monitor. Prince and her colleagues (2008) found that people were wildly inaccurate in their self-reports of physical activity compared to data gleaned from a physical activity monitor in an ESM study. Interestingly, self-reports of physical activity in their study revealed no consistent pattern in relation to data from the activity monitor: Sometimes people overreported physical activity and sometimes they underreported. Participants need not be actively attempting to deceive the researcher—sometimes they are simply unreliable estimators.

### *Descriptive Studies of Experience in Context*

With descriptive studies, we can explore the content of daily experience among students, teachers, or educational leaders. Descriptive studies address questions about the nature and frequency of different experiences or behavior, and they help to illuminate the contours of those experiences. The need for more phenomenologically based descriptive explorations of individuals’ subjective experience and the contexts in which those

experiences occur was, in many ways, the goal of many ESM researchers—and there is still a great deal that we can learn from careful studies of individuals' experiences and how those are shaped by context. ESM can be a wonderful tool for such explorations. Such questions include normative descriptive questions about the frequency of different experiences or actions, but they also might focus on the unique experiences of interesting subpopulations.

A rich line of descriptive research has focused on how students spend their time both inside and outside of school. ESM studies of how much time students actually spending thinking about and engaged in learning while in high school, for example, reveal that participants spent only about 15 hr a week, or an average of 3 hr a day, actively engaged in academic schoolwork and learning while at school (as opposed to participating in other, nonacademic related activities, such as physical education, lunch, chatting with friends, or daydreaming; Leone & Richards, 1989). Shernoff and his colleagues (Shernoff, 2013; Shernoff et al., 2000, 2003) explored questions about how school time is actually used: How much time is given to different pedagogical approaches (e.g., lectures, group work, individual work)? How do students experience these different pedagogies—how engaging, challenging, and enjoyable do they find different activities? In studies with hundreds of high school students, they found that students spent most of their classroom time listening to lectures or engaged in individual work and the least amount of time engaged in class discussions or group projects. Further examination revealed that although students spent the least amount of time engaged in group work (e.g., class discussions, group problem solving), these activities provided the ideal conditions for “flow,” in that students' reports of group work indicated it required high levels of skill, was challenging, and was enjoyable. Students spent the least amount of time engaged in group work in computer science (0%) and history (1.3%) classes and the greatest amount of time in group activities in science (8.6 %) and art (4.8 %), but these are very low levels across the board. Such studies can help us reflect on broad issues in curriculum and instruction. As we begin implementation of Common Core in many states, these patterns present a sobering overview of the changes required in pedagogical practices.

Basic questions about how students spend their out-of-school time have also been effectively studied with ESM. In studies of middle school students engaged in formal after-school programs, Shernoff and Vandell (2007; Vandell et al., 2005) found that students experienced the highest levels of engagement within athletics and arts enrichment after-school activities. Students also reported feeling more engaged in activities with, rather than those without, direct adult involvement. Such data offer insights about the kinds of after-school programs that are most likely to engage youth and therefore have the greatest chance of facilitating positive youth outcomes. Shernoff (2010) found that students' level of engagement, rather than the sheer number of hours present, in after-school programs better predicts improvements in social and academic competence.

Descriptive studies have also been employed to explore the frequency with which more idiosyncratic experiences occur. In a classic set of studies, Swim and her colleagues (Swim, Hyers, Cohen, & Ferguson, 2001; Swim, Hyers, Cohen, Fitzgerald, &

Bylsma 2003) used an “event-focused” approach to have participants report on racist or sexist incidents as they occurred over a period of weeks. Participants were asked to complete a structured diary entry every time they experienced a gender-related (Swim et al., 2001) or race-related incident (Swim et al., 2003). Diary entries included both open- and closed-ended items focused on asking participants to describe the event, who was involved, what actions individuals took, and reports about its emotional impact and participants' assessment of the probability that the incident was motivated by sexism or racism, respectively. Focusing their analysis only on those incidents participants reported were “likely” sexist or racist in nature, the studies provided a comprehensive descriptive understanding of the kinds of sexist or racist incidents that college students encounter as well as their frequency, nature, and emotional impact. Moreover, these data made visible the frequency and nature of a variety of more routine and “mundane” kinds of racist experiences or “micro-aggressions” (e.g., such as being stared at; Swim et al., 2003) that may often go unreported in more retrospective accounts of one's experience.

ESM studies could similarly be applied to explore the frequency and nature of interactions students have with teachers, parents, and peers that might otherwise be hidden from researchers. Building on the work of Swim and her colleagues (2001, 2003), education researchers could explore students' everyday experiences with racism or sexism in schools and illuminate which contexts are the most problematic, who perpetrates these actions, and what students do in response. ESM would allow us not only to learn more about the frequency of those experiences but also to explore context: How often do racial or sexist incidents occur in school versus out of school? How often with adults (such as teachers, staff, or parents) versus with peers (friends, acquaintances, or other schoolmates?). Alternatively, we might ask, “How often do students have conflictual interactions with teachers? What is the ratio of conflictual to positive interactions that students have with teachers?” or “How much time do students spend with peers of their own or a different racial or ethnic group?” or “What is the frequency of their interactions with peers about schoolwork-related issues/relationships/feelings/problem solving/engaging in a task together?” A study of bullying might investigate how often students report feeling bullied or harassed by peers, while also assessing the circumstances and social context of this bullying: Who is bullying whom, in what ways, who is witnessing it, and what are their reactions? These are questions that are often studied using intensive qualitative methods, such as ethnography, that focus on a specific school site or the experiences of a relatively small number of children. The ethnographer cannot be everywhere, however, and some actions—such as bullying or other negative behavior—may be hidden from view. ESM can provide a complementary approach, through which we might gain access to different experiences, and a larger number of children in a wider range of contexts might be studied and their experiences compared.

ESM studies can also help to identify patterns in the daily activities and affective experiences of teachers and school leaders. For example, school leaders generally plan and desire to be deeply involved in instructional leadership—but how much of their time is really given to these activities, and how much gets



taken up with other demands? Descriptive ESM studies could be employed to address theoretically meaningful questions about time allocation, but they could also be used as a professional development tool with teachers or school leaders to help them learn about how their time is spent and whether that matches their professional goals.

### *Studying Intraindividual Change and Processes*

ESM studies focused on change or unfolding processes capitalize on the “nested” nature of ESM data to enable us to explore questions about how individual experience is shaped by contexts and time. In other words, by examining multiple instances of individuals’ experiences over time, researchers can disentangle which effects are due to individual differences and which can be attributed to be contextual factors. ESM used longitudinally also allows us to study how individuals’ thoughts, feelings, and actions are shaped by time—perhaps at different points in a school term or in response to events (e.g., an exam, a grade) along the way. Through repeated assessments of individuals over time in a number of contexts, ESM data are particularly well suited to answering questions about how and why individuals change over time and particularly to help us pinpoint the sources of those changes. Careful orchestration of the timing and data collection mechanisms of experience-sampling studies allows us to look for both naturally occurring changes and changes that occur as the result of interventions.

ESM studies allow us to demarcate how aspects of experience are linked to contexts and to individual differences, thereby allowing us to understand more directly. For example, Schmidt and her colleagues (Schmidt, Lyutykh, & Shumow, 2012; Schmidt, Shumow, & Zaleski, 2012; Shumow, Lyutykh, & Schmidt, 2011) have studied high school students’ experiences during science classes in ways that lead to some interesting insights about what engages students’ motivation and interest in science. Schmidt, Lyutykh, et al. (2012) found that students’ ratings of their experience of challenge, learning, and enjoyment (key components of “flow”; see Csikszentmihalyi, 1991) were higher when science teachers’ discourse was focused on content and encouraged higher-order thinking and questioning than at other times. Their research also revealed that students reported less engagement and a reduced sense of relevance but more enjoyment and interest during lab lessons than during lectures (Schmidt, Shumow, et al., 2012), even though those same lab lessons were rated by scientists as meeting few of the goals of science education. Similarly, Schweinle and her colleagues (2006, 2008) explored the ways that elementary and middle school students’ affective experience in the mathematics classroom related to the level of challenge, importance, efficacy, and skill level they report in order to understand students’ motivational processes. In one set of studies, Schweinle et al. (2006) demonstrate that the positivity of students’ affect emerged from a combination of feelings of personal efficacy and student assessments that the material is challenging and important. A subsequent study revealed that when students report lower personal efficacy (or skill) in the moment and also report that math is important and challenging, positive affect drops precipitously. In contrast, positive affect during math remains high when student-reported

personal efficacy is moderate or high no matter how challenging the material (Schweinle et al., 2008). Importantly, students’ reports of personal mathematics efficacy were highest in classrooms in which teachers cultivated cooperation and support between students.

ESM can also help us to answer questions about relationships between a number of cognitive variables and performance. Kane and his colleagues (2007) explored the relationship between working memory capacity (WMC) and mind wandering among undergraduates. Over a 7-day period, participants responded eight times daily at random intervals to questions about the task they were engaged in and a number of subjective assessments, including whether their mind was wandering from the task at hand. Kane et al. found that when engaged in more difficult tasks, students with a higher WMC were more able to stay focused on the task than students with a lower WMC. These findings help us to understand how WMC might operate to improve academic performance over time—by facilitating time on task, especially during complex or demanding activities.

ESM is ideally suited to allowing us to understand the processes by which change takes place: What are the experiences, activities, actions, or events associated with change? For example, explorations of students’ daily experiences can help us make sense of changes in student engagement, motivation, or performance over time. We have been exploring how women’s earliest experiences in STEM courses in college assessed through an ESM study of participants’ experiences in class are linked to participants’ later academic performance and persistence in STEM majors (Murphy et al. 2014). We found that female college students participated less frequently in STEM classrooms when they perceived their professor to hold a fixed theory of intelligence than when they perceived the professor to have a “growth” theory of intelligence (Murphy et al., 2014). We hypothesize that this reduced participation will be linked to reduced enjoyment, lower performance, and ultimately, reduced likelihood of persisting in a STEM major. Collier, Schernoff, and Strati (2011) found that engineering students were substantially more likely to be engaged in their learning and to pursue advanced course work when early course work included video games that taught key concepts. In a different vein, London, Downey, and Mace (2007) studied beginning law school students at an elite urban law school and found that decreases in engagement, motivation, and confidence among Students of Color in the cohort could be tied to the law students’ reports of (a) the frequency with which they experienced of racially charged or gendered incidents and (b) the extent to which those experiences were described as “isolating” by participants over the course of their first term.

When combined with experimentally designed studies of interventions, ESM approaches can help us better understand the effects of intervention efforts in schools or with youth. A study of a primary school intervention to increase physical activity among 10- and 11-year-olds revealed that the intervention led to a marked increase in physical activity (as measured by a physical activity monitor) and reduced body mass index, an effect that was still evident 6 months after the intervention (Eather, Morgan, & Lubans, 2012). The ESM approach to studying physical activity helped the researchers to learn that the use of tools developed by the researchers to help kids think about

how to use playground equipment in ways that increased physical activity were particularly important. These tools, combined with teacher support for the program, were important predictors of students' ongoing increased physical activity 6 months after the intervention (Eather, Morgan, & Lubans, 2013).

### *Statistical Power and Complexity*

For those engaged in quantitative research, there are two ways to increase statistical power: increasing the number of participants or increasing the number of data points per person. The multiple repeated measures that characterize ESM add statistical power that enhances our ability to detect patterns of interest without enormous sample sizes. Repeated reports from each individual participant increase the reliability of the data, allowing researchers statistical power far greater than would normally be available without very large sample sizes. A "typical" but fairly modest ESM study might include five to 10 samplings per day for 1 to 3 weeks, resulting in 35 to 210 data points per participant, enabling a high degree of reliability and statistical power. Because ESM involves many repeated measures of the same individual in specific contexts, it allows researchers greater opportunity to observe effects resulting from individual differences (e.g., gender or level of motivation), immediate context (e.g., teacher or classroom characteristic), and possible institutional factors (e.g., school type or context). It also enables the addition of variability or variance as a focus of study. As an example, we can study the variability of adolescents' mood, not just average levels of mood. Witte, Fitzpatrick, Joiner, and Bradley Schmidt (2005) found through ESM data that it was the variability in depressive affect, rather than the intensity, that best predicted suicidal ideation and actual suicide attempts with a college sample. Thus, ESM allows for the statistical power to examine complex relationships that influence student experiences that might otherwise require hundreds if not thousands of participants.

ESM data can be complex and are typically multilevel: Individual reports are nested within individuals, individuals within classrooms, and in some cases, classrooms within schools. Multilevel statistical techniques, such as hierarchical linear modeling, are needed to tease apart these different influences within the data. This complexity is also an advantage: ESM's design of repeated measures per individual has statistical advantages over more traditional longitudinal research. Inevitably, in all longitudinal research, some participants drop out along the way or participate somewhat irregularly. In quantitative research, statistical power drops precipitously with each lost participant. With ESM and its focus on repeated measures, statistical analyses can accommodate participants who participated for less than the full study (within certain parameters), and this allows us to explore the experiences of a greater range of participants (e.g., those who are more and less conscientious). Although certainly repeated measures can be obtained through designs that do not involve ESM, the nature of ESM means that a far greater number of repeated measures are possible. ESM designs enable the researcher to examine data from participants with variable response rates, whereas in a repeated measures design, each participant needs to complete the exact same number of reports for the data to be valid.

### **Sampling Approaches for ESM**

The *sampling* part of *experience sampling* is a central feature of the method, and thus it is important to consider different approaches to this sampling. There are three primary means of organizing an ESM study: random sampling, fixed sampling, or event-focused sampling (see Bolger & Laurenceau, 2013, for a more expansive discussion). *Random sampling* of experience describes ESM studies in which, as the name implies, participants' experiences are assessed at random times, generally by means of randomly set alarms on a mobile device. In such a study, when the alarm goes off, participants complete a prescribed measure or some measurement is taken (e.g., speech, activity), and researchers have a spontaneous assessment of experience. Random sampling is most appropriate if a researcher is not targeting specific contexts (e.g., comparing experiences in different classes or college courses; comparing social vs. academic experiences) or does not expect targeted behaviors to occur during a specific time frame (e.g., always at night, at the start or end of the day, in the morning, etc.). The spontaneity of random sampling presents advantages in "finding" participants engaged as they are. For example, Sanchez and Garcia (2009) used this type of methodology to examine how the racial composition of one's immediate context affected perceptions of public regard for one's racial identity and psychological well-being. Random sampling was important in this study because these experiences were presumed to be idiosyncratic. Of course, with random sampling, the researcher may opt to choose to constrain the possible sampling times to typical waking hours, and one can easily restrict the range with which participants will be randomly beeped using an iPod or other type of computerized device.

*Fixed sampling* ESM studies are those in which the measurement of experience occurs at specified times. For example, participants are asked to complete a brief survey at certain times or at specific intervals during the day. This more structured approach allows researchers to assess experience along whatever parameters of time make sense for the research question and makes it more amenable to low-technology approaches to ESM. Daily diaries are a fixed sampling approach that might be used to explore questions about particular kinds of experiences (e.g., "Tell me about your experience with teachers in your school today" or "Tell me about any times you felt bullied at school today"). The emphasis here may be on greater reflectivity in responses in exchange for less spontaneity. Alternatively, some electronic monitors may be set on a fixed timetable to collect data at specified intervals, say, every 5 or 10 min for a measure of activity or speech or every 2 to 4 hr for assessments of participants' subjective experiences.

Finally, with *event-focused sampling*, researchers may set data collection to occur in response to particular events. Researchers might ask a high school or college student to report on his or her experience after every time he or she studies with friends or, as in a study in which we are engaged, after every college class (Murphy et al., 2014). Researchers might ask 1st-year teachers to complete a survey after every interaction they have with parents or their school principal. Those interested in exercise patterns might ask participants to turn on an activity monitor before every exercise session, for example. Such an approach is particularly useful

when the researcher is interested in the effects of a particular kind of event or incident in participants' lives (e.g., experiences in STEM classes). They can also be quite useful when we are interested in low-frequency events, such as interactions between parents and teachers. Event-focused sampling allows the researcher to keep the focus only on the events of interest.

### Considerations in ESM Research

As with all research methods, ESM has its practical and conceptual challenges, and we note a few of those considerations here. The potential "intrusiveness" of ESM is something that must be taken into account. Often, ESM studies are organized such that sleep time and other personal time are protected. A researcher might typically set time constraints on when participants can be asked to respond to reports (e.g., between the hours of 9 a.m. and 10 p.m.). This necessarily limits what can be learned outside of those limits, but it is often essential. The level of psychological "burden" of having to complete reports should inform the frequency and length of reports—typically, reports are kept short in order to minimize intrusiveness, and the desire for more reports must be balanced against participant fatigue.

Participation in these types of studies can be time-consuming and requires a high level of commitment from the participants. Participant fatigue is thus an important consideration, because the researcher may be asking participants to answer the same questions over and over. Researchers will want to ensure that participants are still deeply engaged in the work toward the end of the study. It is important to think about how many times per day and for how many days one can reasonably expect participants to complete surveys—like most research, this becomes a trade-off of more data versus reduced participation. Some participants will inevitably be more conscientious and involved than others, and anything one can do to add to that conscientiousness is worth thinking about in the beginning. Participant compensation is one way to do this; we have found that higher levels of compensation increase participation and engagement. In some studies, we have stratified compensation to match participation levels or provided bonus compensation for those who complete the study (adding their name to a raffle to win a prize, etc.). Regular reminders to complete the surveys also improve compliance rates. One can also attempt to increase investment in the study by highlighting that the participants are "co-researchers" in the study, to enhance the sense that they are learning and creating knowledge with the experimenters.

A central consideration in ESM research is that the research itself really must be theory driven. Although theory is valuable in the design of any study, ESM demands that the researcher make choices that only theory can guide. Initially, the researcher needs a theoretically driven model of the experience he or she is interested in sampling. ESM approaches can be stand-alone studies of experience in a particular constrained time period, but they are often combined with other data collection methods to develop a longer time frame for a longitudinal study of the influence of experiences (e.g., within classrooms) on later outcomes (e.g., grades, student well-being, school persistence, choice of a major in college). The researcher also needs theoretically driven conceptualizations of whom to include in the study and what they

are asked. Theory about the problem under study will drive details, such as when the study will be done (for example, might this be best in a specific time of year; of a term; before, during, or after midterm exams or quarter grades are given?), for how long (3 days, 1 week, 2 weeks?), and how often participants will be contacted (once a day, five times a day? EAR studies might sample speech every few minutes; how often and how much speech to record?). Theory will also guide which approach to sampling (random, fixed, event focused?) is best used to answer specific research questions.

The frequency and timing of the sampling is important: How frequently is the event of interest likely to occur? Is this something that is likely to happen several times a day or once a week? Studies of experiences of racist or sexist events typically are given longer time frames—maybe a month or a semester—to allow for a greater chance that participants will have experienced a number of these events. Studies focused on bullying might also need a longer time frame. Alternatively, studies focused on frequent events, such as students' experiences in the classroom, might need to explore only a week or two of students' lives. Are the events of interest likely to happen in the daytime (e.g., classes) or in the evenings (e.g., studying, socializing)? ESM can be used for more exploratory questions, but it is important for meaningful ESM research design to follow from a basic understanding of the phenomenon under study.

Once the researcher has ideas about the organization of the sampling (length of study, time of day, frequency and method of sampling), attention needs to be given to the chronological placement of the study. Conducting this study at the wrong time could well mean that a researcher "misses" the very events of interest. In a study of students' experiences in classes, for example, we have waited until students were well into the term to ensure that start-of-year enthusiasm might have waned, leaving more individual differences in motivation and engagement (Murphy et al., 2014). Even descriptive studies will need careful attention to planning the timing of data collection so as best to be able to describe the phenomenon under study. Recall the London et al. (2007) study of beginning law students: This study was conducted at the start of the law school experience to explore the impact of early experiences on later law school engagement, performance, and persistence. Alternatively, researchers might need to weigh participant fatigue in timing their studies. During times of the year when participants are especially busy might be a conceptually interesting time to conduct a study—during finals week, for example—but it might also be a particularly difficult time to get good participant compliance.

A core practical challenge of ESM research is that it can be somewhat labor-intensive and expensive for the researcher as well. Other intensive research, such as ethnography, can also be labor-intensive, but such approaches often center that labor in one person. ESM approaches often need a team of research assistants at key times of the study. Programming device alarms, managing devices and participants, and providing technological support can require several research assistants and considerable time from the researcher during the intensive period of the study. Participants in one of our studies were texting research assistants for technological support at 1:00 a.m.! Moreover, electronic



ESM studies utilize equipment that can be expensive—although the equipment can be used for multiple studies. Strategies for minimizing the costs of ESM can be found in Conner, Tennen, Fleeson, and Barrett (2009). Additionally, some of the newer technologies measuring behavior and engagement without a focus on participants' reporting subjective experience can be less labor-intensive: Participants wear a device or play a game and the data collection is just programmed into the device.

Beyond these technical challenges, we have experienced more conceptual challenges as well. Some participants mentioned learning about their own behavior by participating in the study. In our own study of gender and STEM, one woman noted that she observed she was more likely to raise her hand to ask questions or contribute answers in her humanities and social science classes than she was in her STEM classes. This caused her to reflect on her class participation: Why in some classes and not in others? What inhibits me from participating? Can I change this? This insight was of great interest to us and was, in fact, part of the focus of our study. From one perspective, then, we were finding that the phenomenon we were seeing felt "real" to participants, and this served as a kind of "member checking" on our interpretation of their experience. However, this self-awareness may have influenced our participants' behavior over the 2 weeks of the experience-sampling portion of the study. The nature of ESM data allows researchers to detect such patterns (by comparing patterns early versus later in the data collection process or empirically analyzing how previous responses affected subsequent responses). Moreover, such insights and their impact on student behavior might in fact become a subject of study on its own. However, this does mean that it is important to consider how participating in the study might alter students' experience, insights, or behavior.

## Summary and Conclusions

Experience-sampling research methods offer education researchers an innovative approach to investigate questions about the daily lives of educational actors (students, parents, teachers, leaders) that we are not otherwise able to access, thus greatly expanding the areas of experience we can learn about. It offers us access to information and experiences that would otherwise remain unavailable to research—either because they are hidden, because of retrospective biases, or because these experiences occur outside of participants' awareness. ESM research provides additional means of describing participants' experience and placing that experience in detailed context in ways that is otherwise elusive in quantitative research. ESM used longitudinally can allow us to explore questions about change over time and enable the study of educational processes as they unfold, in situ, in "real time." ESM offers education researchers the ability to "get closer to experience," providing more accurate answers to questions about how participants spend their time and with whom as well as providing opportunities to study their affective experience in new ways. ESM also enables time itself to become a new variable of study and affords greater statistical power because of its repeated-measures design. We argue that ESM is underutilized in education research and has much to offer education researchers who want to learn more about how education is experienced. We

also feel that education researchers have much to offer ESM: The creativity of education researchers can expand our conceptions of what these methods can do. We look with anticipation to seeing ways that ESM might help us to gain insights in education.

## NOTE

The authors wish to thank two anonymous reviewers for their insightful and constructive suggestions for how to improve the paper. Support for this paper was received from National Science Foundation Grant No. 1307977 to all three authors and from a sabbatical leave provided by Mills College to the first author. Correspondence should be addressed to the first author at szirkel@mills.edu.

## REFERENCES

- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. doi:10.1146/annurev.psych.54.101601.145030
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal research methods: An introduction to diary and experience-sampling research*. New York, NY: Guilford Press.
- Buchman, M. (2007). The science of subtle signals. *Strategy and Business, 48*, 1–10.
- Chang, U. J., & Ko, S. A. (2007). A study on the dietary intake survey method using a cameraphone [Abstract]. *Korean Journal of Community Nutrition, 12*, 198–205.
- Coller, B. D., Schernoff, D. J., & Strati, A. J. (2011). Measuring engagement as students learn dynamic systems and control with a video game. *Advances in Engineering Education, 3*, 1–32.
- Conner, T. S., & Barrett, L. F. (2012). Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. *Psychosomatic Medicine, 74*, 327–337. doi:10.1097/PSY.0b013e318
- Conner, T. S., Tennen, H., Fleeson, W., & Barrett, L. F. (2009). Experience-sampling methods: A modern idiographic approach to personality research. *Social and Personality Psychology Compass, 3*, 292–313. doi:10.1111/j.1751-9004.2009.00170.x
- Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*, 245–265.
- Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience*. New York, NY: Harper Perennial.
- Csikszentmihalyi, M., Rathunde, K., & Whalen, S. (1997). *Talented teenagers: The roots of success and failure*. Cambridge, UK: Cambridge University Press.
- Eather, N., Morgan, P. J., & Lubans, D. R. (2012). Improving the fitness and physical activity levels of primary school children: Results of the Fit-4-Fun group randomized control trial. *Preventative Medicine, 56*, 12–19.
- Eather, N., Morgan, P. J., & Lubans, D. R. (2013). Social support from teachers mediates physical activity behavior change in children participating in the Fit-4-Fun intervention. *International Journal of Behavioral Nutrition and Physical Activity, 10*, 1–15.
- Estrin, D., & Sim, I. (2010). Health care delivery: Open mhealth architecture. An engine for health care innovations. *Science, 330*, 759–760.
- Fahrenberg, J., & Myrtek, M. (Eds.) (2001). *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies*. Kirkland, WA: Hogrefe and Huber.
- Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience-sampling methods: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage.



- Holleran, S. E., Whitehead, J., Schmader, T., & Mehl, M. R. (2010). Talking shop and shooting the breeze: A study of workplace conversation and job disengagement among STEM faculty. *Social Psychological and Personality Science*, 2, 65–71. doi:10.1177/1948550610379921
- Kahneman, D. (2011). *Thinking fast and slow*. New York, NY: Penguin Press.
- Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20, 3–24. doi:10.1257/089533006776526030
- Kane, M. J., Brown, H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of working memory and executive control in daily life. *Psychological Science*, 18, 614–621.
- Larson, R., & Richards, M. H. (1991). Daily companionship in late childhood and early adolescence: Changing developmental contexts. *Child Development*, 62, 284–300. doi:10.1111/j.1467-8624.1991.tb01531.x
- Leone, C. M., & Richards, M. H. (1989). Classwork and homework in early adolescence: The ecology of achievement. *Journal of Youth and Adolescence*, 18, 531–548.
- London, B., Downey, G., & Mace, S. (2007). Psychological theories of educational engagement: A multi-method approach to studying individual engagement and institutional change. *Vanderbilt Law School*, 60, 455–481.
- Mehl, M. R., Pennebaker, J. W., Crow, M., Dabbs, J., & Price, J. (2001). The electronically activated recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavioral Research Methods, Instruments, and Computers*, 33, 517–523.
- Morrow, V. (2001). Using qualitative methods to elicit young people's perspectives on their environments: Some ideas for community health initiatives. *Health Education Research*, 16, 255–268. doi:10.1093/her/16.3.255
- Murphy, M. C., Garcia, J. G., & Zirkel, S. (2014). *Women, STEM, and the impact of professors' theories of intelligence on performance and persistence over time* [Unpublished data]. Indiana University, Bloomington.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54(1), 547–577.
- Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 5, 56–79. doi:10.1186/1479-5868-5-56
- Rollo, M. E. (2012). *An innovative approach to the assessment of nutrient intake in adults with type 2 diabetes: The development, trial, and evaluation of a mobile phone photo/voice dietary record* (Unpublished doctoral thesis). Queensland University of Technology, Brisbane, Australia.
- Sanchez, D. T., & Garcia, J. A. (2009). When race matters: Racially stigmatized others and perceiving race as a biological construction affect biracial people's daily well-being. *Personality and Social Psychology Bulletin*, 35, 1154–1164.
- Schmidt, J. A., Lyutykh, E., & Shumow, L. (2012, April). *A study of teachers' speech and students' perezhivanie in high school physics classrooms*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, BC.
- Schmidt, J. A., Shumow, L., & Zaleski, D. (2012, April). *Multiple perspectives on student motivation, learning, and engagement in high school biology labs*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, BC.
- Schweinle, A., Meyer, D. K., & Turner, J. C. (2006). Striking the right balance: Students' motivation and affect in elementary mathematics. *Journal of Educational Research*, 99, 271–294.
- Schweinle, A., Turner, J. C., & Meyer, D. K. (2008). Understanding young adolescents' optimal experiences in academic settings. *Journal of Experimental Education*, 77, 125–146.
- Sherhoff, D. J. (2010). Engagement in after-school programs as a predictor of social competence and academic performance. *American Journal of Community Psychology*, 45, 325–337.
- Sherhoff, D. J. (2013). *Optimal learning environments to promote student engagement*. New York, NY: Springer.
- Sherhoff, D. J., & Csikszentmihalyi, M. (2009). Flow in schools: Cultivating engaged learners and optimal learning environments. In R. C. Gilman, E. S. Heubner, & M. J. Furlong (Eds.), *Handbook of positive psychology in schools* (pp. 131–145). New York, NY: Routledge.
- Sherhoff, D. J., Csikszentmihalyi, M., Schneider, B., & Sherhoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18, 158–176. doi:10.1521/scpq.18.2.158.21860
- Sherhoff, D. J., Knauth, S., & Makris, E. (2000). The quality of classroom experiences. In M. Csikszentmihalyi & B. Schneider (Eds.), *Becoming adult: How teenagers prepare for the world of work* (pp. 141–164). New York, NY: Basic Books.
- Sherhoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: Quality of experience from the perspective of participants. *Journal of Youth and Adolescence*, 36, 891–903. doi:10.1007/s10964-007-9183-5
- Shumow, L., Lyutykh, E., & Schmidt, J. A. (2011). Predictors and outcomes of parental involvement with high school students in science. *School Community Journal*, 21, 81–98.
- Stanley, K. G., & Osgood, N. D. (2011). The potential of sensor-based monitoring as a tool for health care, health promotion, and research. *Annals of Family Medicine*, 9, 296–298.
- Strack, R. W., Magill, C., & McDonagal, K. (2004). Engaging youth through photovoice. *Health Promotion Practice*, 5, 49–58.
- Strati, A. D., Sherhoff, D. J., & Kackar, H. Z. (2012). Flow. In R. Levesque (Ed.), *Encyclopedia of adolescence* (pp. 1050–1059). New York, NY: Springer.
- Swim, J. K., Hyers, L. L., Cohen, L. L., & Ferguson, M. J. (2001). Everyday sexism: Evidence for its incidence, nature and psychological impact from three daily diary studies. *Journal of Social Issues*, 57, 31–53.
- Swim, J. K., Hyers, L. L., Cohen, L. L., Fitzgerald, D. C., & Bylsma, W. C. (2003). African American college students' experiences with everyday racism: Characteristics of and responses to these incidents. *Journal of Black Psychology*, 29, 38–67. doi:10.1177/0095798402239228
- Vandell, D. L., Sherhoff, D. J., Pierce, K. M., Bolt, D. M., Dadisman, K., & Brown, B. B. (2005). Activities, engagement, and emotion in after-school programs (and elsewhere). *New Directions for Youth Development*, 105, 121–129. doi:10.1002/yd.111
- Wilson, T. D., Meyers, J., & Gilbert, D. T. (2003). "How happy was I, anyway?" A retrospective impact bias. *Social Cognition*, 21, 421–446. doi:10.1521/soco.21.6.421.28688
- Witte, T. K., Fitzpatrick, T. E., Joiner, T. E., Jr., & Bradley Schmidt, N. (2005). Variability in suicidal ideation: A better predictor of suicide attempts than intensity or duration of ideation? *Journal of Affective Disorders*, 88, 131–136.

## AUTHORS

**SABRINA ZIRKEL**, PhD, is a professor of education in the School of Education at Mills College, 5000 MacArthur Boulevard, Oakland, CA

94613; [szirke@mills.edu](mailto:szirke@mills.edu). Her research focuses on the critical study of the ways that race, ethnicity, gender, and class influence students' educational experiences and also explores the ways that raced and gendered narratives shape the social construction of "merit" in public conversations about education.

**JULIE A. GARCIA**, PhD, is an associate professor of social psychology and associate chair in the Psychology and Child Development Department at California Polytechnic State University, 1 Grand Ave, San Luis Obispo, CA 93407-0387; [jgarc31@calpoly.edu](mailto:jgarc31@calpoly.edu). Her research focuses on how contexts shape social identity meanings and consequently well-being and motivation with respect to multiple groups, including multiracial people and women in science, technology, engineering, and mathematics.

**MARY C. MURPHY**, PhD, is an assistant professor of social psychology in the Department of Psychological and Brain Sciences at Indiana University, 1101 E 10th Street, Bloomington, IN 47405; [mcmpsych@indiana.edu](mailto:mcmpsych@indiana.edu). Her research focuses on prejudice, stereotyping, and the situational cues in organizations and educational institutions that trigger social identity threat—examining how these cues elicit cognitive, affective, and motivational processes that influence the behavior of advantaged and disadvantaged groups.

Manuscript received March 12, 2013  
Revisions received July 7, 2014, and December 1, 2014  
Accepted December 4, 2014