August 2019

Multimethod Investigation of Pristine Inner Experience

Leiszle Rae Lapping-Carr

MULTIMETHOD INVESTIGATION OF PRISTINE INNER EXPERIENCE

By

Leiszle Rae Lapping-Carr
Bachelor of Arts – Psychology
Bachelor of Arts – Religious Studies
Brown University
2009

Master of Arts – Psychology
University of Nevada, Las Vegas
2016

A dissertation submitted in partial fulfillment
of the requirements for the

Doctorate of Philosophy in Psychology

Department of Psychology
College of Liberal Arts
The Graduate College

University of Nevada, Las Vegas
August 2019

UNLV | GRADUATE COLLEGE

Dissertation Approval

The Graduate College
The University of Nevada, Las Vegas

January 16, 2019

This dissertation prepared by

Leiszle Rae Lapping-Carr

entitled

Mulitmethod Investigation of Pristine Inner Experience

is approved in partial fulfillment of the requirements for the degree of

Doctor of Philosophy - Psychology
Department of Psychology

Christopher Heavey, Ph.D.
Examination Committee Chair

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Dean

Russell Hurlburt, Ph.D.
Examination Committee Member

Murray Millar, Ph.D.
Examination Committee Member

William Ramsey, Ph.D.
Graduate College Faculty Representative

ABSTRACT

Introspection can be defined as any effort to observe and report on internal experiences. As such, introspection continues to be a commonly used research method, including self-report questionnaires, experience sampling, and qualitative interviews. However, in these modern applications of introspection, the challenges of such endeavors are often not readily acknowledged or addressed. This study compared three introspective methods using a pre-test, post-test design: descriptive experience sampling (DES), the experience sampling method (ESM), and daily questionnaires (DR). Those who participated in DES, a beeper-based method designed to produce high fidelity understandings of random moments of inner experience, had lower average frequencies of common phenomena of inner experience (e.g., inner speaking) than did ESM or DR participants. DES participants also had differences twice as large, on average, as ESM and DR between *in general* reports of inner experience and momentary (ESM) or daily reports. These dramatic differences between DES and ESM or DR suggest that questionnaire-based methods (ESM and DR), regardless of timeframe, do not capture pristine inner experience. Furthermore, our results indicate that questionnaires are not a valid tool for estimating the frequency or describing the characteristics of various kinds of inner experience. Many of our psychological constructs and theories have been developed based on questionnaire data. Without valid, high-fidelity reports of inner experience, psychological science will be missing a vital piece of information: individual, lived experience.

ACKNOWLEDGEMENTS

I would like to thank my committee chair and mentor, Dr. Chris Heavey, for his support and guidance throughout this process. I would also like to thank my committee members, Drs. Hurlburt, Miller, and Ramsay, for their comments and contributions to this project. I am grateful for my fellow graduate students and research assistants, Cody, Steve, Liana, and Corinne, and their work in the various stages of recruitment and training of the participants. Finally, thank you to my family, friends, and especially my husband, Micah, for their emotional support and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Introspection can be defined as "looking into our own minds and reporting what we discover" (Boring, 1953). Essentially, introspection is aimed at understanding first person, private phenomena, or pristine inner experience. Pristine inner experience is that which is directly present before the footlights of consciousness before it is distorted by any attempt at observation or interpretation (Hurlburt, 2011). According to the above definition, many commonly used methods would be considered introspective, including self-report questionnaires, experience sampling, threshold discrimination, and qualitative interviews (Boring, 1953; Brock, 2012; Clegg, 2012; Locke, 2009).

Introspection has been discredited by psychological researchers since before the middle of the 20th century, and researchers generally deny the introspective nature of these commonly used methods (Boring, 1953; Locke, 2009). This denial results in a tendency to ignore some of the most important limitations of these modern introspective methods (Ericsson & Simon, 1993). Self-report questionnaires are a particularly common introspective method with numerous well-documented limitations. The use of retrospective reporting results in memory errors, such as a mood-congruent bias (Kihlstrom, Eich, Sandbrand, & Tobias, 2000), reconstruction errors and failures of encoding (Tourangeau, 2000), and reliance on semantic as opposed to episodic memory (Robinson & Clore, 2002b, 2002a). These errors may all impact responses such that researchers may be unable to determine the relevance of their data to the research question at hand. Introspective methods that minimize retrospective requirements, such as experience sampling methods and casual observation, are still subject to the limitations of language in reporting on pristine inner experience (Grice, 1975; Skinner, 1974), participants' unwillingness to disclose private information (Heavey, Hurlburt, & Lefforge, 2010), the tendency for analytic

self-reflection to change the experience on which participants are reporting (Halberstadt & Wilson, 2008; Mitchum, Kelley, & Fox, 2016; Wilson, Dunn, Kraft, & Lisle, 1990), and, perhaps most importantly, the presence of presuppositions (Hurlburt, 2011; Hurlburt & Schwitzgebel, 2007, 2011).

Descriptive experience sampling (DES) is a method that was designed with the aim of capturing pristine inner experience in high fidelity (Heavey et al., 2010; Hurlburt, 1990, 1993, 2011; Hurlburt & Heavey, 2006). The method limits retrospective requirements and maximizes ecological validity by having participants apprehend their pristine inner experience when prompted by a randomized beeper in their natural environment. It protects against bias and miscommunications by iteratively training participants to bracket (put aside) presuppositions about pristine inner experience and to clarify language throughout multiple days of interviewing.

When introspection is used without careful attention to the inherent hazards of studying pristine inner experience, the results consist of some indefinable mixture of pristine inner experience, assumptions about what inner experience is or should be, beliefs about the self, and memory errors (Hurlburt, 2011; Hurlburt & Heavey, 2015). This study aims to compare three introspective methods: DES, (non-DES) experience sampling, and daily report questionnaires. Participants were randomly assigned to one of these three methods. We utilized a pre-test, post-test design to compare general self-reports before and after participants engage in the three methods. We also compared the results of the three methods; DES data was coded for frequently occurring phenomena to allow comparison to the more quantitative measures. We expected that participants who engaged in DES during the experimental manipulation would report significantly lower frequencies of experience in post-test reports compared to their pre-test reports. We also expected that average frequencies of experiences obtained using DES would be

significantly lower than the average frequencies obtained by both daily reports and (non-DES) experience sampling. These findings would support the notion that DES measures different phenomena (pristine inner experience) from what is measured by other introspective methods (the indefinable mixture). Should this be true, the veracity of interpretations made by researchers employing these other introspective measures should be carefully reviewed.

CHAPTER 2: LITERATURE REVIEW

A Brief History of Introspection

In modern psychological research, "introspection" is often used pejoratively. It is usually used to refer to the highly-structured method used by Wundt and Titchener in the late 19th and early 20th centuries. Their method, which we will term Introspection (with a capital *I*), included extensive training of participants and had the aim of understanding the fundamental elements of consciousness (Boring, 1953). Participant training was of utmost importance; only trained participants could be trusted to observe their experience in a scientifically acceptable way (Boring, 1953; Brock, 2012). Wundt used Introspection for particular research questions but acknowledged the limitations of the method, such as a lack of generalizability. He believed that a full understanding of psychology necessitated diverse methodology (Brock, 2012).

A primary goal of some early psychologists, such as Wundt and his students, was to identify the fundamental elements of consciousness, or "atoms" of thought, using Introspection (Brock, 2012). Wundt's students formed two separate research groups: the Würzburg School and a group at Cornell University lead by Titchener. The Würzburg School claimed to have identified a phenomenon that is perhaps best translated as "states of consciousness," in which participants seem to have experienced a state without an identifiable sensory correlate. Titchener, believing that he had used Introspection to identify psychological atoms as exclusively consisting of sensations (Boring, 1953), publicly disagreed with the Würzburg School's interpretation. This public disagreement gave rise to the infamous "imageless thought" controversy that is often cited as the downfall of introspection. Like many other disciplines within psychology, two research groups approached a problem from different theoretical viewpoints, resulting in different interpretations of gathered data. When looking at the results of Introspection, we see that the

Würzburg School and Titchener's studies garnered similar data, and it was only in the interpretation that they disagreed (Boring, 1953; Brock, 2012; Hurlburt & Heavey, 2006; Monson & Hurlburt, 1993).

Unfortunately, Wundt and Titchener's theoretical disagreement was identified as the result of using a faulty and subjective research method: Introspection. Therefore, the method was rejected and discontinued, in word if not in deed, as behaviorism and later cognitivism became dominant forces in American psychological research. Still, introspection (with a lowercase *i*) maintained a strong presence in European psychology and never disappeared completely from American psychology (Brock, 2012). It has gone by a number of names, including phenomenological observation and self-report (about inner states; Boring, 1953). Further, introspection is necessary for other kinds of psychological endeavors, including psychophysical research (such as threshold discrimination), social psychology, and psychotherapy (Boring, 1953; Clegg, 2012).

Pristine Inner Experience

Introspection is purportedly aimed at understanding first-person, private phenomena, or pristine inner experience. Pristine inner experience is that which is directly present in awareness, before it is distorted by attempts at observation or interpretation (Hurlburt, 2011; Hurlburt & Heavey, 2015). It is "pristine in the same sense as we would say a forest is pristine…. Pristine does not necessarily mean 'clean' or 'tranquil'; much of a pristine forest is mucky, bloody, brutal and so on" (Hurlburt, 2011). It is not pure or clean, but rather is untouched. It is *not* judgments about the apprehended pristine inner experience, reactions to the apprehended pristine inner experience, or opinions about the causation of pristine inner experience, but includes hearings,

seeings, speakings, tickles, pains, feelings, etc. It also may consist of nothing, or a lack of direct experience (Heavey & Hurlburt, 2008; Hurlburt, 2011; Hurlburt & Heavey, 2015).

Pristine inner experience is often conflated with other kinds of experiences. The word "experience" can be used to describe a number of different things (Hurlburt & Schwitzgebel, 2007), including, but not limited to, beliefs about the self (I often experience anger), assumptions about inner experience (I had a bad day; thus, I was experiencing anger), generalities about situations (anyone who is unappreciated at work would experience anger), and experiential phenomena (I am currently feeling anger, which presents itself to me as heat across my chest). Rarely do we differentiate between these different kinds of experiences, but they are quite distinct in terms of meaning and interpretation. Collecting data on beliefs about one's experience might help to identify the way one's self-view is related to overall functioning. Interpreting data that includes assumptions about inner experience might help understand situational impact on self-perception. Apprehending experiential phenomena may provide insight into the everyday, lived experience of humans. When we do not distinguish between these distinct forms of experience, we lose our ability to make meaningful interpretations of our data. If we are not clear about what we are measuring, we cannot be clear about how our data relates to anything else (Hurlburt, 2011; Hurlburt & Heavey, 2015).

Research on pristine inner experiences has identified five experiential phenomena that occur most frequently across participants (Heavey & Hurlburt, 2008), termed the "5FP" for "5 frequent phenomena" (Kühn, Fernyhough, Alderson-Day, & Hurlburt, 2014). The 5FP include inner speaking (the experience of producing words internally; Hurlburt, Alderson-Day, Kühn, & Fernyhough, 2016; Hurlburt, Heavey, & Kelsey, 2013), inner seeing (the experience of seeing in the absence of the thing seen internally), feelings (the experiential component of emotion;

Heavey, Hurlburt, & Lefforge, 2012; Heavey, Lefforge, Lapping-Carr, & Hurlburt, 2017), sensory awareness (focus on internal or external sensations like taste, lights, touch, etc.; Hurlburt, 2009; Hurlburt, Heavey, & Bensaheb, 2009), and unsymbolized thinking (clear cognitive experience that does not include words, images, or any other symbol in experience; Hurlburt, 2009; Hurlburt & Akhter, 2008). Across participants, each of the 5FP occur in approximately 25% of sampled moments. However, the frequencies of the 5FP vary drastically across individuals, ranging from 0% to 100% (Heavey & Hurlburt, 2008). Also, many participants experience phenomena that do not fit into the 5FP (e.g., "just doing," or the experience of being completely absorbed in an activity) and neither are there sharp boundaries that delimit the 5FP. Pristine inner experience is often messy and unable to be neatly categorized. While the 5FP can often be used to succinctly refer to an individual's experience, it is also very common for aspects of pristine inner experience to be outside these categories, or to include aspects of more than one category. The 5FP are our best attempt at identifying nomothetic themes across participants, but are not clearly distinct or fully inclusive (Heavey & Hurlburt, 2008).

Hazards of Studying Pristine Inner Experience

As suggested by the continued use of introspection in psychology, pristine inner experience is of vital interest to the field. To study pristine inner experience adequately, one must be aware of the pitfalls and hazards of this research and plan accordingly. Hurlburt (2011) presents over 100 hazards, including the (more manageable for this paper) 10 fundamental hazards to studying pristine inner experience discussed by Heavey and colleagues (2010):

1) The first hazard is retrospection. Retrospection requires participants to reflect back on their experiences and remember what happened. There is considerable evidence that retrospection

is flawed (Kihlstrom et al., 2000; Tourangeau, 2000). Episodic memory is highly susceptible to error and alteration (Tourangeau, 2000). Memories may never have been encoded properly, meaning there is nothing there to remember. If a participant is not adequately aware of and attentive to her inner experience, it will be impossible for her to remember them and report them. Memories can also be altered through the incorporation of post-event information. A personal event must be rehearsed to enter long-term memory. During rehearsal, details are often altered or added, and the memory that is stored is different from the initial memory (Tourangeau, 2000). There is also considerable evidence that people report with a mood-congruent bias (Kihlstrom et al., 2000). If a participant happens to be in a particularly bad mood when they are introspecting about their inner experience over the past few days, they are more likely to remember and report negative events, emotions, experiences, etc.

2) The second hazard is that a participant's or experimenter's unconscious assumptions about the nature of inner experience, or presuppositions, will impact the apprehension and report of pristine inner experience. For example, societal expectations that men have less frequent and weaker experiences of feelings than do women seem to result in participants' reporting consistent with this belief when asked to reflect on their experience over an extended period of time. However, men and women do not report significantly different frequencies of feelings when they are asked to report on momentary experiences (e.g., LaFrance & Banaji, 1992). Also, when researchers are interested in a particular experiential phenomenon (e.g., negative feelings), they may ask more questions, or more detailed questions, or use a different sentence structure or tone of voice when asking questions about negative feelings. Whatever the reason, participants are then particularly likely to identify negative feelings,

leading to an overreporting of negative feelings, and potentially faulty conclusions about the frequency or nature of negative feelings (Tourangeau, 2000).

3) The third hazard is that participants will use semantic memory, instead of episodic memory, to report on their experience. Robinson and Clore (2002b) conducted a series of studies that present strong evidence in support of this claim. They found that when participants were asked to rate the intensity of feelings over long periods of time (e.g., the last few months) they tended to use semantic memory. Conversely, when participants were asked to rate the intensity of feelings over shorter time periods (e.g., the past day) they tended to use episodic memory. When the time frame becomes long enough, attempting to recall specific events is not an efficient strategy for making judgments about feelings. Instead, participants use what they know about the world and about themselves to estimate their feelings (Robinson & Clore, 2002b). An important issue related to his hazard is the length of reporting delay that leads to a shift from the use of episodic memory to semantic memory; Robinson and Clore (2002b) suggest that it occurs with delays (or reporting windows) in the range of hours, but Hurlburt and Heavey (2015) suggest that this shift may occur with much shorter delays, perhaps anything beyond a few seconds.[1]

4) The fourth hazard is that participants will report unconscious processes or assumed processes as part of pristine inner experience. For example, people have a poor ability to determine the motivations behind their actions (Nisbett & Wilson, 1977). This can be seen when researchers manipulate a social pressure variable to influence human behavior without participant knowledge. They are often successful in manipulating the behavior but the social

[1]Hurlburt and Heavey (2006, 2015) identify several other factors, in addition to reporting delay, that can lead away from reports based solely on episodic memories.

pressure influencing the behavior goes unreported; participants report a variety of motivations for their behaviors, but not social pressure. People are often mistaken about what motivates their behavior and it is difficult for them to access their own cognitive processes (Nisbett & Wilson, 1977).

5) The fifth hazard is the risk of ignoring the inherent individuality of participants and treating them as nothing more than indistinguishable data points. If we lose sight of the individual person who had the experience, we miss out on the human element of pristine inner experience. Only after we have attended to the individual and the features of their experience can we can look across participants to identify group themes (Hurlburt, 2011). Conducting data analyses before considering the individual can result in misleading conclusions. This is true even with simple, descriptive statistics such as means. For example, the individuals in a sample will likely vary in the frequency with which they engage in inner speaking. If you look at the individuals, you may notice that some individuals engaged in inner speaking about 10% of the time, some engaged in inner speaking 90% of the time, that some individuals had a unique inner voice that is distinct from their outer voice, that some couldn't identify the voice of the speaker (but knew that it wasn't their own), and many other interesting, qualitative characteristics (Heavey & Hurlburt, 2008; Hurlburt et al., 2013). If the basis of the analysis is determining the mean frequency of inner speaking across participants, there is no awareness of the other important features of inner speaking. With the above example, you might even state that your participants experience inner speaking 50% of the time, on average. This statement is misleading, since no participant actually experienced inner speaking 50% of the time. Ignoring participant individuality results in a loss of the inherent individual nature of inner experience, but it is much easier to collect, analyze,

interpret, and distribute quantitative data. Attending to individuality provides a depth of understanding about each participant but produces qualitative data that is incredibly difficult to collect, analyze, interpret, and distribute.

6) The sixth hazard is that participants will purposefully hold back or be dishonest when describing their pristine inner experience. Some topics are particularly sensitive or personally embarrassing, such that participants may not want to report them to investigators. For example, it is well documented in the sexuality literature that participants in sexuality studies represent a specific subgroup of the population, one that views sex particularly positively (Saunders, Fisher, Hewitt, & Clayton, 1985; Strassberg & Lowe, 1995; Trivedi & Sabini, 1998; Wolchik, Braver, & Jensen, 1985). One interpretation of this research is that a participant who is not part of this subgroup is unlikely to disclose sexual aspects of her experiences when asked to report on her experience in general on a particular day.
7) The seventh fundamental hazard is the limitation of our language in describing pristine inner experience. Unlike language for external events, people do not have consistent opportunities to develop and shape their language for pristine inner experience. Without this shaping, people make erroneous assumptions, for example, that the word "think" reflects the same kind of experience for all people (Hurlburt & Heavey, 2001; Skinner, 1974). Participants and researchers must create the opportunities to refine and clarify language about inner experience. Creating opportunities for refinement requires an on-going ("iterative") conversation between the researchers and participants. Both must collaboratively work towards a mutual understanding (Grice, 1975). For example, when an individual says they feel sad, they may be referring to feeling an empty pit in their stomach, or a heaviness in their arms and legs, or a mental feeling that does not exist in their body, or they may be thinking

about a sad event, but not having a direct experience of a feeling of sadness at the moment in question. Without a careful conversation aimed at disambiguating such possibilities, it would be impossible to distinguish these experiences.

8) The eighth hazard is inadequately aiming for ecological validity. The assumption that pristine inner experience in a prescribed situation (such as in a research lab) approximates that which is found during everyday life is a large, and questionable, assumption. Purposefully self-initiating introspection, such that the participant begins the task unprompted and actively reflects on and analyzes her ongoing experience, is unlikely to yield results that approximate natural pristine inner experience (Hurlburt & Schwitzgebel, 2011).
9) The ninth hazard is that high fidelity apprehension of pristine inner experience requires considerable investigator skill. Essentially, an investigator must be able to hold in awareness and work to address all of the hazards presented here. This requires training, innate skill, and a willingness to be corrected by other investigators and participants (Heavey et al., 2010; Hurlburt, 2011).
10) The tenth hazard is that participants generally have little skill at apprehending or reporting their pristine inner experience. The exception may be experienced meditators, but they may only be skilled at apprehension, not report, and skilled meditators are rare in Western cultures. Without adequate training in how to apprehend pristine inner experience, we cannot be confident that our participants are giving us anything other than presuppositions and assumed processes. Or, if they are adequately apprehending their pristine inner experience, that their intent with their description matches our understanding of their description (Grice, 1975; Heavey et al., 2010; Hurlburt, 2011).

Whereas the hazards of studying pristine inner experience through introspection are well researched and documented, they are rarely identified as significant concerns (Ericsson & Simon, 1993; Heavey et al., 2010; Hurlburt, 2011; Hurlburt & Heavey, 2015; Mitchum et al., 2016; Robinson & Clore, 2002a; Skinner, 1974; Tourangeau, 2000). Researchers often give inadequate (or no) attention to the issue that their questioning about pristine inner experience may produce biased or entirely inaccurate reports about pristine inner experience. This is not an argument to stop using introspection. We, as psychological researchers, must utilize some form of introspection if we want to study pristine inner experience at all. However, it is important to take these limitations seriously when interpreting the results of introspection research.

Methods of Studying Pristine Inner Experience

The next step is identifying and describing different methods of studying pristine inner experience. With each description will be included a discussion on how each method addresses and is vulnerable to the above listed hazards.

Casual Observation. The most readily recognized introspective method is purposeful reflection on currently occurring inner experience; here we will term it casual observation. Casual observation utilizes self-initiated, present-tense, and targeted judgments of currently occurring inner experience (Siewert, 2011). Essentially, participants, when it occurs to them to do so, attend to their ongoing inner experience, ask themselves pre-determined questions about that experience, and attempt to answer those questions. By focusing on currently occurring experience, casual observation minimizes retrospective demands (hazard 1) and prioritizes gathering data on experience as it is occurring, purportedly maximizing ecological validity (hazard 8). Casual observation typically allows free-form responses to questions and participants

are encouraged to engage in further exploration of identified experiences as part of the data gathering procedure (Seager, 2002; Siewert, 2011).

The questions participants ask themselves while engaging in casual observation reflect the researcher's area of interest. When used to study emotional experience, questions might be something like "What am I feeling right now?" or "What does this emotion feel like?" or "Why am I feeling this way?" The participant is expected to elaborate on these questions, providing detailed information about any and every feeling that is occurring, the characteristics of those feelings, and the reasons for those feelings. This questioning strongly emphasizes gathering detailed information about each individual participant's experience and is thought to accurately portray the potential *depth* of inner experience (Siewert, 2011).

A major cause for concern in interpreting causal observation data is the researcher's inability to determine what part (if any) of the answers to these questions reflect pristine inner experience, presuppositions about one's inner experience (hazard 2), semantic knowledge about the self (hazard 3), and assumptions about inner experience in general (hazard 4). More concretely, the reports are some combination of reactions to engaging in a self-observational task and an evaluation of one's experience. For example, Seager (2002) discusses his attempts to introspect on emotional experience:

> …the values of things are an integral part of our conscious experience of the world right from the start. Is this phenomenologically plausible? I can only speak for myself, but it does seem clear to me that I experience the value of things as well as their perceptible appearance. (pp. 678)

Using this method, it is impossible to distinguish *evaluation* from *experience*. The very definition of casual observation identifies that it is gathering judgments of experience and not

descriptions of the experience itself. Therefore, the depth of experience that is sought by casual observation is possibly (likely?) a product only of the method and is not a true feature of pristine inner experience.

Social psychological researchers have directly investigated the disruptive impact of causal introspection on feelings and attitudes (Halberstadt & Wilson, 2008; Wilson et al., 1990). They found that when participants were asked not just how they were feeling, but why they were feeling that way, they would present a logical, rational story for their feelings. Often, this resulted in participant attitudes shifting away from affective and towards cognitive, reflecting a change from their initial reports. However, the participants' behavior was still in line with their original affective reports, creating a discrepancy between attitudes and behaviors (Halberstadt & Wilson, 2008; Wilson et al., 1990). Relatedly, there is recent cognitive research showing that evaluating one's internal processes changes the quality of those processes (Mitchum et al., 2016). Participants who continuously made judgments about their learning while learning a list of word-pairs spent their study time differently than participants who did not, and had inferior recall to participants who did not make these judgments (Mitchum et al., 2016).

We should be sure to reiterate that the introspection engaged in during those studies is casual introspection. Participants are given instructions to reflect on why they feel the way they do, or on how well they are learning, in the context of a specific experimental task. There is no attempt to measure their experience as it occurs during normal, everyday activities. Although casual introspection purports to maximize ecological validity (hazard 8), the method itself inherently limits that validity. If someone is told to analyze why they felt a certain way, part of their pristine inner experience will necessarily include the attempt or intent to analyze their experience (Hurlburt & Heavey, 2015).

A distinction can be made between those who practice casual observation in a skilled way (i.e., researchers reflecting on their own experience) and those who engage in this reflective practice solely for participation in a study (Hurlburt & Schwitzgebel, 2011; Siewert, 2011). Skilled casual observers are likely invested in understanding their own inner experience in a way that aligns with their theoretical opinions; this leaves them particularly subject to the influence of presuppositions, use of semantic memory, and the reporting of assumed processes (hazards 2, 3, 4, and 9). Furthermore, their observations about pristine inner experience are based solely on their own experience, and the limitations of our language (hazard 7) prevent any certainty about how their experiences relate to others'. Casual observers who did not engage in introspection prior to agreeing to participate in a study are not skilled at attending to their experience or particularly invested in exploring it carefully (hazard 10). Although observations by participants yield a large number of introspective reports, there is no training for the participants on how to introspect, and so participants are likely to simply report what they believe to be true or most normal (Hurlburt & Schwitzgebel, 2011). Without a clear demonstration by the experimenters that all experience is acceptable, participants are likely to present results they think are in line with experimenter or societal expectations (hazard 6). Then, although the data itself is suspect, it is usually analyzed in terms of means and group trends (Halberstadt & Wilson, 2008; Mitchum et al., 2016; Wilson et al., 1990). We have little understanding of any of the individuals involved in these kinds of studies (hazard 5).

Arguably, it is preferable to have a small number of high-quality reports of experience. This becomes problematic when researchers employing casual observation use their own introspections to make claims about the nature of inner experience for all people (hazard 9).

Quoted above, Seager (2002) states “I can only speak for myself (pp. 678)” with regard to perceiving value along with perception. However, he later comes to the following conclusion:

> Every state of consciousness, no matter how purely intellectual it might seem, contains the evaluative component (though in many cases the represented value might be virtually neutral, rather than the decisively positive or negative values characteristic of vividly experienced emotion). This is necessary if we are to care about our thoughts. It is in any case phenomenologically obvious that abstract thoughts can produce strong emotional reactions (as those engaged in intellectual pursuits know better than anyone; pp. 680).

With this statement, Seager is taking his understanding of his own inner experience (that the value of a perception is inextricable from the perception) and applying it to all conscious experience, regardless of the individual having the experience or the experience itself. His theory of conscious experience fits well with his personal experiences; therefore, this theory must be true of all people and all experiences.

Self-Report Questionnaires. Self-report questionnaires utilize introspection but are readily overlooked when considering introspective methods. Questionnaires about pristine inner experience typically ask respondents to reflect on one possible phenomenon of their inner experience (e.g., feelings or inner speaking) and report on its characteristics (e.g., rate the frequency or intensity on a scale of 1-10); this procedure clearly meets our definition of introspection. The introspective character of questionnaires may be readily overlooked in part because questionnaires are common among psychological researchers; we believe it is important to include questionnaires in this discussion for this very reason. Questionnaires should be held to the same rigorous standards as other introspective methods.

Questionnaires almost always limit response variability by providing participants with closed-ended questions (e.g., Rate feeling *X* on a scale of 1-10). Limited response variability seems to be equated with a method that utilizes a high degree of control over data collection. However, given the degree to which participants interpret instructions and questions based on their individual experience, it is possible (in fact likely) that measures that appear to have strong data collection control are simply not acknowledging the different meanings that individual participants ascribe to their responses (Ericsson & Simon, 1993; Locke, 2009). When a participant reads the question "How happy do you feel right now on a scale of 1-10?", she must comprehend the question, retrieve information about her personal experiences of happiness from memory (e.g., the relative intensities of happiness she has felt in her lifetime), make a judgment about where her current level of happiness falls along her personal spectrum of happiness, and assign it a numerical rating in order to answer the question (Lenzner & Menold, 2016). The individual variability of this process is not captured by a method that has strict limitations on response options (Ericsson & Simon, 1993; Hurlburt, 2011; Hurlburt & Heavey, 2006; Mitchum et al., 2016).

There are innumerable questionnaires that ask about pristine inner experience. Furthermore, the time-frame for questionnaires is quite variable, ranging from *in general* to *right now*. For clarity and space, let us consider one time-frame: *today*; and a questionnaire that is frequently used: The Positive and Negative Affect Scale (PANAS; (Watson, Clark, & Tellegen, 1988). While we are specifically evaluating the PANAS for the time-frame *today*, we should be clear that this evaluation could be readily applied to any questionnaire about inner experience.

Results obtained using the PANAS have formed the basis for a theoretical structure of emotional experience. Tellegen, Watson and Clark (1999) used data gathered with the PANAS to

examine different structural models, trying to determine whether positive and negative affect are bipolar (*cannot* occur simultaneously) or independent (*can* occur simultaneously). They found support for a three-level hierarchical structure: the highest order structure is a single, bipolar, Happiness vs. Unhappiness factor; the second level includes independent Positive Affect and Negative Affect factors; and the lowest order includes individual emotions. These data suggest that we first characterize our experience of emotion on a given day as either "happy" or "unhappy." Within "happy" and "unhappy," we can be independently experiencing positive affect (e.g., enthusiastic) *and* negative affect (e.g., nervous). Within both positive and negative affect, we can be experiencing many different discrete feelings (enthusiastic *and* inspired *and* nervous *and* distressed; Tellegen et al., 1999).

A more recent analysis re-examined the structure of emotional experience using the PANAS, but included an analysis of within-subjects variability along with a between-subjects analysis (Rush & Hofer, 2014). They found that Positive and Negative Affect factors within-subjects were negatively correlated. This suggests that the independence found between Positive and Negative Affect by Tellegen and colleagues (1999) is a result of averaging across subjects, and that individuals actually experience *bipolar* Positive and Negative Affect, such that they *do not* experience both positive and negative emotions simultaneously.

Both the initial (Tellegen et al., 1999) and the more recent (Rush & Hofer, 2014) interpretations of sophisticated statistical analyses make strong assertions about feelings. Both studies used prompts that required participants to retrospect about their feelings over *the past day*. This time frame may be small enough to allow participants to use their episodic memory instead of semantic memory when considering their feelings during the day (hazard 3; Robinson & Clore, 2002a, 2002b), but it does not take into account issues such as the relative salience of

different feelings, the availability of certain episodic memories at recall, or the strategy that participants use to identify all of their feelings throughout the day and determine how to rate each one on the given scale (Hurlburt & Heavey, 2006, 2015). There is also an attempt to obtain ecological validity (hazard 8), in that participants filled out the questionnaire on multiple days in their natural environment. Rush and Hofer (2014) further work towards considering the individuality of the participants (hazard 5) by using a within-subjects analysis to explore to individual differences.

The studies do not address retrospection (hazard 1). Participants are providing ratings of their experiences over the course of an entire day. There are 57,600 seconds in a 16-hour period. Feelings can rise and fall within a few seconds. It is very likely not possible for participants to accurately remember every feeling they had during the day. By not addressing retrospection, these studies become vulnerable to reports of assumed processes (hazard 4). If a participant used their episodic memory to identify a number of negative events that occurred during the day (e.g., flat tire, stained shirt, late to work, argument with spouse), they may have assumed that they were feeling angry during the day, even if they did not recall a specific instance of the experience of anger. Similarly, these studies do nothing to address the presuppositions that participants have about their feelings (hazard 2). If a participant believed herself to be an unemotional person, she would likely have rated herself low, and not have the ability to identify a potential discrepancy between her presupposition of herself as unemotional and her pristine inner experience that day. A participant's inability to differentiate between presuppositions and pristine inner experience reflects inadequate participant skill (hazard 10). This is not the participant's fault, they simply have never been required or encouraged to develop this skill.

The use of the PANAS also falls prey to the assumption of a mutual understanding of the terms being used (hazard 7). Participants are given a simple rating scale, a list of 20 feelings (e.g., *jittery*, *upset*), and instructions to report on their experiences that day. There is no way to ensure that the experimenter's definition of *jittery* was the same as Participant A's definition of *jittery*, or that either was the same as Participant B's definition of *jittery*. Therefore it is impossible to know, based on their PANAS scores, if these participants actually experienced vastly different feelings that day. The limited response options, as well as the lack of clear language and instruction, means we get the participant's interpretation of the question. The inability to clarify language and differentiate between the experiences of participants reflects inadequate investigator skill (hazard 9). There is either a lack of awareness or a lack of concern over being uncertain about the meaning of the results of a questionnaire like the PANAS.

Regardless of which statistical analysis you consider to be superior, the 3-level hierarchical analysis (Tellegen et al., 1999) or the within- and between-subjects analysis (Rush & Hofer, 2014), the results are equally problematic. We do not understand the features of the inner experience of any of the participants. The language used in the questionnaire allows for large individual interpretations (e.g., rating the intensity of a single instance of emotion or rating the frequency of an emotion) and the limited response options do not allow for the capture of those individualized interpretations. Even relatively brief time frames (like *the past day*) are subject to hazards like retrospection, reports of assumed processes, and presuppositions. Asking about specific feelings likely results in an overreporting of those feelings, as non-feelings experiences are re-interpreted as including a feeling in the face of a specific inquiry. The results from the PANAS are some mixture of feelings, beliefs about the self, beliefs about feelings in general, and

the current mood of the participant. At best, the structure of emotion posed by the results of the PANAS reflect how people believe emotion to be structured.

Experience Sampling. Experience sampling methods utilize random or event-based sampling of momentary experience in natural environments. Researchers use some kind of electronic device, like a programmable wristwatch, beeper, or a smartphone, to signal participants to attend to their experience and answer questions provided. The experience sampling method (ESM) was first presented by Csikszentmihalyi and Larson (1987). In this study, participants were asked to report on a number of features related to their inner experience, including their thoughts, behaviors, attention, feelings, pain, and possible reasons for their current experience (Csikszentmihalyi & Larson, 1987). Here, we will consider a study of self-talk using such a method (Brinthaupt, Benson, Kang, & Moore, 2015).

Brinthaupt and colleagues (2015) present a series of three methodologically diverse studies aimed at assessing the accuracy and validity of the Self-Talk Scale. The Self-Talk Scale (STS; Brinthaupt, Hein, & Kramer, 2009) is a questionnaire designed to measure the occurrence of self-talk (both out loud and internal) in a specified list of situations. Brinthaupt et al.'s (2015) Study 1 compares general reports of self-talk with reports of self-talk over the past two days using the STS. Their Study 2 uses ESM to measure reports of self-talk over the last two hours; it compares these immediate reports with participants' general characterization of their self-talk. It also compares the immediate reports of participants who generally report being high and low self-talkers. Their Study 3 measures self- and other-reports of self-talk and compares close-other and stranger-other reports. The overall conclusions of the study identify no significant differences between in general, two day, and two hour reports of self-talk, significant differences between high and low self-talkers on ESM reports (only in positively-oriented situations), and

significant differences between self- and other-reports of self-talk, regardless of whether the other was a stranger or a close-other (Brinthaupt et al., 2015).

Focusing on the ESM Study 2, the authors state "A stronger test of the self-report accuracy question [than Study 1] would involve very recent experiences that should be more salient and accessible to the participants" (Brinthaupt et al., 2015). Clearly, the authors are working to address the hazards of retrospection, semantic memory, and ecological validity (hazards 1, 3, and 8, respectively). The STS presents participants with specific situations on which to reflect, the 2-hour time frame is designed to be short enough to increase the likelihood that participants are relying on their episodic memory of the situation, and participants are asked about situations that occurred during their everyday life. However, many of the same concerns that are relevant to self-report questionnaires also apply to ESM. It should be noted that the Brinthaupt et al. (2015) study is simply an example of how ESM is used to study inner experience. The critique offered here is not a specific criticism of these researchers or this single study.

Some hazards that stand out in the ESM study are presuppositions, report of assumed processes, participant reluctance to disclose, and language limitations (hazards 2, 4, 6, and 7, respectively). The prompt for the STS includes the statement "researchers have determined that all people talk to themselves, at least in some situations or under certain circumstances" (Brinthaupt, Hein, & Kramer, 2009). This prompt follows from the experimenters' presuppositions about self-talk, clearly setting the expectation that people talk to themselves in at least some of the situations listed. Should a person report *never* talking to themselves, she would be abnormal. Thus, when the participant identifies that she experienced one of the situations listed in the past two hours, she may simply assume she spoke to herself, or she may search her

experience for anything related to that situation that is at all similar to self-talk, or she may be too embarrassed to report never having engaged in self-talk. The language of the STS instructions does not leave room for participants who have experiences discordant with experimenter expectations. For example, low self-talkers reported self-talk 31% of the time in general, but 54% of the time using ESM, suggesting that when they were more concretely recalling specific situations, they were more likely to report self-talk. This result may be the result of regression to the mean, as suggested by the authors, but the high self-talkers did not report a similarly drastic decrease (79% in general to 73% during ESM; Brinthaupt et al., 2015).

The results given above do not provide sufficient evidence to make claims about the nature or frequency of self-talk. We do not know whether these averages have any meaning for individual participants (hazard 5). There were 13 participants included in the low self-talker group discussed above. A large shift for just a few participants may have been the source of the increase in average frequency. Even if the increase was a trend among most of the participants, it is unlikely they are primarily reporting on pristine inner experience. Although a two-hour time frame reduces retrospection requirements and reliance on semantic memory, the lack of training on how to attend to one's pristine inner experience (hazard 10) means they will continue to rely on presuppositions and assumptions.

Brinthaupt et al. (2015) did acknowledge the limitations of their study in terms of accessing pristine inner experience:

> A more accurate understanding of the STS and the present results would be to note that the STS is a measure of whether people notice talking to themselves and how often they recall doing so upon reflection. The present results refer more to respondents'

> interpretations of "experience and generalities" (Hurlburt and Heavey, 2015, p. 156) than actual, ongoing experiences of talking to themselves. (p. 10)

The authors cite the purpose of the study as a way to assess the validity of the STS. If we think about the STS as a way to measure participant beliefs about how often they speak to themselves, then we can view the STS as a valid measure. However, the authors' concluding statement does not seem consistent with this view:

> Combined with other research supporting the psychometric properties of the STS, the research reported here provides evidence that this measure of self-talk frequency can be used successfully to study individual differences in the phenomenon of self-talk. (Brinthaupt et al., 2015, p. 10)

The influence of experimenter and participant presuppositions, reports of assumed processes, and the other hazards discussed above on the reported occurrence of self-talk in these situations is too unpredictable. We cannot know if these participants are actually engaging in self-talk or are only reporting doing so because of the nature of the questionnaire and the study.

Descriptive Experience Sampling. Descriptive experience sampling (DES) is, as the name suggests, an experience sampling method; however, it incorporates some unique features that differentiate it from the experience sampling method. DES was created with the aim of apprehending pristine inner experience (Hurlburt, 2011). As such, the method was designed and adapted to work towards addressing all of the difficulties inherent in that task, including each of the 10 hazards discussed previously.

In DES, participants use a randomization beeper to sample their experience in their natural environment for approximately three hours. DES utilizes natural environment, random sampling to maximize ecological validity (hazard 8). Participants sample their experience

whenever is convenient for them, and the randomization prevents over-sampling experiences that the participant or the experimenter thinks are particularly interesting or important. When the beep occurs, the participant takes brief notes about what was ongoing in her experience the moment before she was interrupted by the beep. Using a beeper to cue the participant's attention to her pristine inner experience purposefully limits the degree of retrospection to a few seconds (hazard 1). Writing notes about the experience may take a few minutes, but the information included is constrained by their apprehended experiences. Participants are interviewed about their experiences within 24 hours of collecting samples. Their verbal descriptions in the interview are constrained by their written notes (Hurlburt & Heavey, 2015). This process is usually repeated five or more times for each participant (Heavey et al., 2010; Hurlburt & Heavey, 2006).

In the expositional interviews, the investigators work collaboratively with the participant to produce a high-fidelity description of the pristine inner experiences sampled. A key aspect of this interview is the collaboration: participants are explicitly told they have the deciding vote in whether to continue exploring a sample of experience (hazard 6). Investigators tell participants that if they are not comfortable answering questions about a sampled moment because they find it too embarrassing or uncomfortable, they may simply decline to discuss that moment. This approach both protects the participants' privacy, as there are certainly things they experience that are none of the investigator's business, and encourages full disclosure of other experiences. When participants are expressly given the freedom to refuse to answer questions, and can tell by the questioning that the investigators are genuinely interested in any experience they may have, they often become excited to share their pristine inner experiences (Heavey et al., 2010; Hurlburt, 2011; Hurlburt & Heavey, 2006; Turner, 2015).

Another task of the expositional interview is to increase the skills of bracketing both investigator and participant presuppositions about what should or should not be in experience at a given moment (hazard 2). During the interview, participants are asked the initial question: "What, if anything, was in your experience at the moment of the beep?" Subsequent questions are all aimed at gaining a high fidelity understanding of that momentary experience. This provides a level playing field for participants to bring anything they were experiencing to the discussion, without favoring feelings, or thoughts, or visual imagery, or any other specific phenomena. At first, participants are not skillful at bracketing presuppositions (hazard 10). They often report not the ongoing experience but the activity in which they were engaged, what they believe they usually do during that activity, or the processes they assume would have been ongoing during that activity (e.g., arguing with Mom means they must have been feeling angry; hazards 3 and 4).

To help participants improve their skills in bracketing their presuppositions and avoid reporting on assumptions and using semantic memory, participants receive iterative training in apprehending and describing their experience. As participants become more skilled in DES, they are putatively better able to describe their experience in high fidelity. These higher-fidelity descriptions can be used to provide clarity about features of experience described in previous interviews that were initially ambiguous. The primary goal of the first expositional interview is to increase the participant's ability to cleave to the moment of the beep during their next sampling day and then to describe their experience in the next expositional interview, as well as to increase the skills of the investigators to explore the experience of that particular participant.

Iterative training is also aimed at helping participants learn to apprehend only that experience that is caught in flight by the beep. The first expositional interview often includes

reports of experience or activities that occurred minutes before and after the beep. By repeatedly asking participants *only* about the experience that was ongoing and interrupted by ("caught in flight by") the beep, DES aims at truly ongoing experience. Cleaving to the moment reduces the degree to which memory errors could impact the apprehension of their experience: the experience was ongoing when the participant was cued (hazard 1). Taking notes allows the participant to more readily remember their experience in the expositional interview. Various time frames have been used between sampling and interviewing. Hurlburt and Heavey (2006) assert that as long as the interview is within 24 hours of sampling, there is no practical difference in the reporting timeframes.

The uncertainty of language is directly confronted in the interview (hazard 7). For example, participants often use the word "think" to describe any internal event (Hurlburt & Heavey, 2001). When asked for more details about what they meant by "thinking," some participants will describe the saying of a specific sequence of words to themselves with intonation and cadence in their own voice, other participants will describe visually imagining a scene, others will describe being focused on a sensory aspect of something in their environment, others will describe a clear cognitive process but say that there are no words, images, or other features to better describe it, and yet others will be unable to give a clear description (Hurlburt & Heavey, 2001), and so on. These kinds of distinctions are usually unable to be made in the first expositional interview. It is only with iterative training and practice that participants develop the language to describe their inner experience (as they have never had to put these experiences into words before) and that the investigators come to understand the individual language used by a participant to describe their experience (Hurlburt & Heavey, 2001, 2006).

The iterative training process is just as important for the investigators as it is for the participant (hazard 9). Investigators need to devote considerable time in order to become skilled interviewers, such that they become skilled at bracketing their own presuppositions and help the participant do the same. Even skilled investigators do not always get it right and must work with each new participant to hold multiple possibilities in mind while interviewing. Each participant and investigator has a unique set of presuppositions and language idiosyncrasies, and it is only with considerable experience and effort that investigators are able to adapt to each new participant and to overcome their own presuppositions (Hurlburt, 2011).

Within 24 hours of each interview, contemporaneous descriptions are written of each sampled moment and passed between investigators for disagreements, differing points of view, revisions, and so on. When a participant completes the study, an idiographic characterization is written summarizing all sampled experiences, again passed between investigators for revisions, and so on; differing points of view or understandings are encouraged and written directly into the descriptions. This is an effort to best capture the experience in high fidelity, which is sometimes messy. Only after all participants have completed the study, and idiographic characterizations are created for each participant, are nomothetic themes considered. This allows us to capture the individuality of each participant involved in DES studies (hazard 5).

DES offers a combination of strict methodological structure and open response format. Participants are provided with concrete procedures to follow, a specific cue to attend to their experience, and a predictable format in which to describe that experience. They are also provided minimal suggestions as to what *should* be in experience at a given moment, instead being encouraged to report on *whatever* is present, if anything. In the interview, the first question asked about each moment is “What, if anything, was in experience at the moment of the beep?” Once a

participant provides an initial description, the investigators ask more detailed questions in an attempt to better describe the apprehended experience. If a participant reports they were talking to themselves, the investigators may ask “Exactly what words, if any, were you saying to yourself?” If a participant reports they were feeling happy, the investigators may ask “How did this feeling happy present itself, if at all?” In the early interviews, many participants find these questions difficult or impossible to answer. The investigators normalize this difficulty and encourage the participants, on the next sampling day, to try and apprehend those details if they are present. The investigators stress that they have no opinion about thoughts, or feelings, or images, or the specific features of those experiences, or anything else being a particularly interesting or important kind of experience. No experience is just as interesting as any particular experience; the investigators want only descriptions of whatever happens to be present, if anything. This open-beginninged line of questioning allows participants free-range in reporting any and all experiences they apprehend.

Whereas DES was designed with the specific aim of describing pristine experience in high fidelity, it of course has limitations. It is possible that the iterative training, bracketing of presuppositions, careful use of language, and open-beginninged questioning are no better at protecting against memory errors and cognitive biases than the other methods; some critics argue that the training may have a negative impact, resulting in reports further from pristine inner experience. In general, we believe this is not true. DES has demonstrated good inter-observer reliability (Spearman-Brown median typical reliability estimate = .98, range from .91 to .98; Hurlburt & Heavey, 2002). Furthermore, inner speaking identified using DES was concordant with brain activation indicative of language production obtained using fMRI (Hurlburt, Alderson-Day, Fernyhough, & Kühn, 2017; Kühn et al., 2014).

DES is, however, highly reliant on investigators' interviewing skill and their ability to bracket their own presuppositions. An interviewer who starts to believe that they know exactly what a participant is saying, especially early in a description or early in training, runs a serious risk of imparting their bias to that participant. Also, over many years DES has identified five categories of phenomena that occur more frequently than others. Perhaps the expectation that these categories occur more frequently have contributed to their more frequent occurrence. This hypothesis is somewhat questionable, as the rates of occurrence have not increased across the years, but have remained approximately the same since the initial identification of the five frequent phenomena (Heavey & Hurlburt, 2008; Kelsey, 2016).

One of DES's greatest challenges is in the time intensive nature of the procedure. A single participant completing a study requires approximately 10 hours of at least two investigators' time. Running a study with 20 participants would require at least 400 hours of investigator time. Thus, sample sizes are generally small. The data that results from DES (i.e., idiographic characterizations) are interesting and valuable, but are not easily quantifiable or distributed to other researchers. And the majority of the data is idiographic: it is about that individual and does not necessarily tell us anything about any other person's pristine inner experience. Small sample sizes make nomothetic characterizations difficult, but at times we have pursued this goal, like with the identification of the 5FP (Heavey & Hurlburt, 2008). Still, to date we believe it is the best available method to provide data on pristine inner experience, and thus is a valuable tool within the repertoire of psychological researchers.

Reflections

Given the above discussion, it is clear I am approaching this project with some preconceived notions about the quality of the various methods included in my study. I believe

that (non-DES) experience sampling and daily report questionnaires do not adequately account for the hazards inherent to the study of pristine inner experience. I want to clearly state that I do not think (non-DES) experience sampling and daily report questionnaires are invalid methods in and of themselves. They are well-suited for numerous research tasks, including behavioral, attitudinal, and observational measures; however, these methods are not able to describe pristine inner experience. DES was developed with the goal of apprehending pristine inner experience, and thus I believe it is far better suited to this task.

These beliefs have the potential to introduce bias into my data collection and analysis. However, any experimenter comparing different methods is likely to have beliefs about the strengths and weaknesses of the methods under consideration. In an effort to protect the results of this study from experimenter bias, I incorporated blinding procedures into the data collection and analysis such that the DES interviewers did not have access to questionnaire data until all DES data analysis was completed. By acknowledging and clearly stating my beliefs, maintaining an awareness of this potential influence, and including protections in my experimental design, the overall impact of this issue hopefully has been reduced.

CHAPTER 3: AIMS OF THE STUDY

The current study aimed to directly compare three introspective methods using a pre-test, post-test design: descriptive experience sampling (DES), (non-DES) experience sampling, and daily report. Participants filled out an initial set of questionnaires including measures assessing their inner experience *in general* and their personality. This initial set of questionnaires about inner experience *in general* constituted the pre-test data. The experimental manipulation included one group that participated in DES, one group that participated in (non-DES) experience sampling, and one group that participated in daily reporting. All participants then completed the post-test: a set of questionnaires about their inner experience *in general* (the same as at pre-test).

With this study, we hoped to answer two questions: 1) Do different methods of investigating inner experience produce different results? 2) Does iterative training in apprehending and describing pristine inner experience as provided by DES impact characterizations of experience *in general*? We expected significant differences between frequency of 5FP as identified using DES and the 5FP as measured by questionnaire using the other two methods of studying inner experience. We also expected participants who engaged in DES to show significantly different *in general* reports pre- and post-experience sampling. We did not expect participants who engaged in (non-DES) experience sampling or daily report to demonstrate any changes in their pre- and post-test *in general* reports. We believe that these differences would support the notion that DES is studying pristine inner experience, while (non-DES) experience sampling, daily report, and *in general* questionnaires are studying an indefinable combination of pristine inner experience, memory errors, presuppositions, semantic knowledge, assumed processes, etc.

CHAPTER 4: METHOD

Participants

Participants were 60 students from the University of Nevada, Las Vegas (UNLV) Psychology Subject Pool. A power analysis had indicated 60 participants would provide enough power to detect significant differences between phases given a large effect size (effect size partial-eta squared = .2, power = .8; Faul, Erdfelder, Lang, & Buchner, 2007). We chose a large effect size given the prior research that has demonstrated very large effect sizes when directly comparing DES and questionnaire reports (Heavey et al., 2019). Participants were randomly assigned to one of the three groups such that 12 participants were assigned to the descriptive experience sampling (DES) group, 24 to the experience sampling method (ESM) group, and 24 to the daily report (DR) group. Due to the amount of time required by researchers for each member of the DES group, we were unable to achieve adequate power and equal sample sizes for each group. Three participants dropped out of the study during the experimental phase, one from the ESM group and two from DR group, leaving a total of 57 participants included in the analyses (DES: n = 12; ESM: n = 23; DR: n = 22).

The final sample of 57 participants (15 men, 41 women, 1 transgender individual) ranged in age from 18 to 57 (M = 21.41, SD = 6.90). Participants were ethnically diverse: 1.7% Arabic, 12.3% Asian, 15.8% Black/African-American, 29.8% Hispanic/Latino, 3.5% Native Hawaiian/Pacific Islander, 22.8% White/Caucasian, and 14.0% selected multiple races/ethnicities. There were no significant differences between groups in demographics ($p > .05$).

All participants who completed the study received four hours of research credit. They also received up to $15 (DR), $20 (ESM), and $25 (DES) for completing study activities.

Participants were required to be 18 years of age or older, speak and read English fluently, and have sufficient availability to complete two interviews a week. There were no other inclusion or exclusion criteria for the study.

Measures

Demographic Questionnaire

The demographic questionnaire was created for this study and includes questions about participant age, enrollment status, year in school, employment status, handedness, gender, race/ethnicity, relationship status, sexual orientation, religious affiliation, and prior experience with formal mindfulness practice which consisted of a single item rated on a scale of 0 (*No experience*) to 5 (*Expert/daily mindfulness practice*).

Big Five Inventory-2 (BFI-2)

The BFI-2 consists of 60 items assessing normal-range personality characteristics using a five-point agreement scale from 1 (*strongly disagree*) to 5 (*strongly agree*). This is the newest revision of the BFI and shows improved psychometric properties (Soto & John, 2017). The BFI-2 measures personality on five domains and has demonstrated acceptable internal reliability: Negative Emotionality ($\alpha = .91$), Extraversion ($\alpha = .87$), Open-Mindedness ($\alpha = .84$), Conscientiousness ($\alpha = .88$), and Agreeableness ($\alpha = .83$). This measure was included as part of an additional phase of the study that is not relevant to the current study. Results from this questionnaire will not be reported.

Five Facet Mindfulness Questionnaire (FFMQ)

The FFMQ consists of 39 items assessing five facets of mindfulness: Observing, Describing, Acting with Awareness, Non-judging, and Non-reactivity (Baer, 2006). Participants rate how true statements are for them on a scale of 1 (*never or very rarely true*) to 5 (*very often*

or always true). The FFMQ has demonstrated acceptable internal reliability: Observing (α = .83), Describing (α = .91), Acting with Awareness (α = .87), Non-judging (α = .87), and Non-reactivity (α = .75; Baer, 2006). This measure was included as part of an additional phase of the study that is not relevant to the current study. Results from this questionnaire will not be reported.

Nevada Inner Experience Questionnaire (NIEQ)

The NIEQ (Heavey et al., 2019) is a 10-item questionnaire that asks participants to rate the frequencies of characteristics of their own inner experience with regard to the five frequent phenomena (5FP: inner speaking, inner seeing, unsymbolized thinking, feeling, sensory awareness). For example, the two items on inner seeing are: "How frequently do you mentally see or visualize something?" rated on a 10-point visual-analog scale between *Never* and *Always*; and "Generally speaking, what portion of your inner experience is in images (seeing things in your imagination)?" rated on a visual-analog scale between *None* and *Always*. The scale has demonstrated acceptable psychometric properties. Cronbach's alpha values were somewhat low, especially for sensory awareness (inner speaking α = .50, inner seeing α = .66, unsymbolized thinking α = .60, feeling α = .65, and sensory awareness α = .34; Heavey et al., 2019), but this is understandable as the scales consist of only two-items, making Cronbach's alpha a less than ideal measure of reliability (Briggs & Cheek, 1986). Inter-item correlations, an alternate measure of internal reliability, were acceptable (inner speaking r = .33, inner seeing r = .49, unsymbolized thinking r = .43, feeling r = .48, and sensory awareness r = .21; Heavey et al., 2019).

Nevada Inner Experience Questionnaire, At-The-Moment Version (NIEQ-ATM)

Designed for this study, the NIEQ-ATM is identical to the NIEQ except that instead of instructing participants "to indicate the **characteristics of your own inner experience**"

(emphasis in original), participants are instructed "to indicate the **characteristics of your own inner experience at the moment of the beep**" (emphasis in original). Also, participants are only presented with one question for each of the 5FP: "At the moment of the beep, what portion of your inner experience was…" as the frequency question would not make sense for an *at the moment* timeframe.

Nevada Inner Experience Questionnaire, Today Version (NIEQ-T)

Designed for this study, the NIEQ-T is identical to the NIEQ except that instead of instructing participants "to indicate the **characteristics of your own inner experience**" (emphasis in original), participants are instructed "to indicate the **characteristics of your own inner experience today**" (emphasis in original). Each question was also modified to orient participants to the past day. For example, the two inner seeing items are: "How frequently did you mentally see or visualize something today?" and "Overall, what portion of your inner experience today was in images (seeing things in your imagination) today?" Results from the NIEQ-T will not be reported in this study.

Positive and Negative Affect Schedule (PANAS)

The PANAS (Watson et al., 1988) consists of a list of 20 emotions and was used to assess positive and negative affect. Ten of the emotions are positive (e.g., enthusiastic, proud), and 10 of the emotions are negative (e.g., upset, ashamed). Participants rated their emotions on a 5-point scale from 1 (*very slightly or not at all*) to 5 (*extremely*). The PANAS has shown good internal consistency when used to rate emotions *in general* (positive affect $\alpha = .88$, negative affect $\alpha = .87$) and with the timeframe *today* (positive affect $\alpha = .90$, negative affect $\alpha = .87$; Watson et al., 1988). Participants were given the prompt "Indicate to what extent you generally feel this way, that is, how you feel on the average" for pre- and post-test assessment and the prompt "Please

indicate the extent to which you have felt this way today" for DR group participants. Results from the *today* timeframe will not be reported.

Self-Talk Scale (STS)

The extent to which participants talk to themselves (silently or out loud) is often assessed with the Self-Talk Scale (STS; (Brinthaupt et al., 2009). The STS presents the participants with situations and asks them to rate the veracity of each statement on a 5-point frequency scale from 1 (*never*) to 5 (*very often*). For example, "I talk to myself when I need to figure out what to do or say". It has shown adequate test-retest reliability ($r(99) = .66$, $p < .001$) and internal consistency across 4 subscales (α ranged from .79 to .89). All participants completed this measure at pre- and post-test.

Self-Talk Scale-Today (STS-T)

Designed for this study, the STS-T is identical to the STS except that participants were instructed to "Determine how true each item was for you personally **today** by circling the appropriate number next to each item" (emphasis in original). Also, the prompt for each question was "Today, I talked to myself when…" and participants were able to endorse an "N/A" option if they did not encounter the situation that day. This adaptation procedure was consistent with that used by Brinthaupt et al. (2015) in their validity study. Results of the STS-T will not be reported in this study.

Apparatus

Participants in the DES and ESM groups used a random signal generator (or beeper) to sample their experience. The beeper was carried in a pocket or clipped to a belt. Once the beeper was turned on, it emitted a 700 MHz tone at random intervals uniformly distributed between 0 and 60 minutes with a mean duration of 30 minutes between signals. The signals were delivered

through an earpiece and the volume was adjustable. The signal was stopped and the next random interval programmed by pressing a button on the top of the beeper. Participants in the DES group were also given a pocket-sized (3 in x 5 in) notebook in which to record notes about their inner experience.

Procedure

We used a pre-test post-test design with three experimental groups to compare data gathered through interview-based experience sampling (DES), questionnaire-based experience sampling (ESM), and questionnaire-based daily report (DR). We examined both between and within group differences, comparing pre- and post-test data within and between groups as well as the data obtained using each method.

In the first (pre-test) phase, participants provided informed consent, were assigned a group, and completed the initial packet of questionnaires. In the second (experimental manipulation) phase, the DES group sampled their experience using DES, the ESM group sampled their experience using the NIEQ-ATM, and the DR group reported on their experience using the NIEQ-T, PANAS (*today* timeframe), and STS-T at the end of the day. In the third (post-test) phase, all participants again completed the same questionnaires as in the pre-test phase.

Phase 1: Pre-test

The initial meeting with participants included informed consent, an explanation of the three phases of the study, randomization into one of the three groups, and an initial set of questionnaires: demographics, the BFI-2, the FFMQ, NIEQ, PANAS, and STS. After completing the questionnaires, participants were awarded 1.5 hours of research credit through the Psychology Subject Pool. Participants were fully consented to all conditions before group

assignment. During the informed consent procedure, participants were provided with information about the differences in compensation, procedure, and time commitment for the different conditions. Participants were encouraged to ask any questions about the procedure during this meeting and throughout the study.

Immediately following informed consent, participants were block-wise-randomly assigned to one of three conditions: the descriptive experience sampling (DES) group, the experience sampling method (ESM) group, and the daily report (DR) group. Participants were randomly assigned in six blocks of 10 participants. Within each block, two participants were randomly assigned to the DES group, four to the ESM group, and four to the DR group. This method of randomization resulted in the 12 DES participants being staggered throughout the recruitment period. The significant time commitment (approximately 5 hours/week for each participant) required by researchers to conduct DES procedures limited the number of DES group participants able to be run simultaneously.

Phase 2: Experimental manipulation

This phase constituted the experimental manipulation of the study. Participants assigned to different conditions (DES, ESM, and DR) reported their experience in one of three ways: through DES sampling, through ESM sampling, or through daily report as described below.

Participants assigned to the DES group received in depth training on observing and describing their experience as part of the Phase 2 procedure. Because this was part of the experimental manipulation, participants assigned to the ESM and DR groups did not receive the same degree of experimental-manipulation phase training.

Descriptive experience sampling (DES) group. Participants sampled their experience using standard descriptive experience sampling on five days. Participants used the randomized

beeper to collect six samples of experience on a given day, which took approximately three hours and was completed at the discretion of the participant. On average, DES participants completed experimental manipulation phase study procedures in 17.67 days (SD = 4.67, ranging from 11 to 28 days). The first day was considered training and was not included in later data analysis.

Within 24 hours of sampling their experience, participants took part in an expositional interview aimed at producing a high-fidelity description of their inner experience as it was ongoing the moment before the beep interrupted. The expositional interview continued for one hour or until all samples were described, whichever was shorter. The first expositional interview was used to help the participant cleave to the moment of the beep, clarify their language in describing their experience, and begin to put aside (bracket) presuppositions about what should be present in experience. Subsequent expositional interviews continued to build and reinforce these skills. Expositional interviews were video-taped to allow for review during data analysis, if necessary.

The expositional interviews were collaborative, and the researchers explicitly stated the necessity of the participants' contribution of their experience to the shared endeavor. As part of this collaboration, participants were encouraged to only describe a sampled moment if they were willing to describe every detail. Both to protect privacy and encourage full disclosure of non-sensitive samples of experience, participants were given the choice to not discuss any samples about which they were uncomfortable answering questions.

Within 24 hours after each expositional interview, one of the investigators wrote a contemporaneous description of the sampled experiences discussed during that interview. The descriptions were circulated to the other investigators who had been present at the interview for comment. Disagreement about the details of experience was encouraged; sometimes discussions

about different points of view led to consensus about the characteristics of a particular sample, but sometimes they did not. Both were considered valuable. When disagreements persisted after review and discussion, they were maintained in the written description to best capture the messiness of the experience in high fidelity.

DES group participants were awarded one hour of research credit through the Psychology Subject Pool for each expositional interview, given immediately after each interview. They also received $15 for completing all sampling days and interviews, given immediately after the final interview.

Experience sampling method (ESM) group. Participants sampled their experience using ESM on five days. Participants used the beeper to collect six samples of experience on a given day; this took approximately three hours and was completed at the discretion of the participant. When the beep sounded, participants completed the NIEQ-ATM online using a convenient device of their choosing (e.g., smartphone, laptop). Participants were asked to sample their experience on five separate days over a three-week period, with no more than two days sampled in a single week. This procedure approximated the DES group participants' sampling experience. On average, ESM participants completed experimental manipulation phase procedures in 15.65 days (SD = 4.61, ranging from 5 to 30 days). Participants chose on which days they preferred to sample their experience and received an electronic reminder (text or email) at 8:30 AM on each day they chose. The reminder included the link to the questionnaire. ESM group participants received typical ESM training in how to use the beeper during their initial meeting with the researchers; questions about the procedure were answered at any time.

ESM group participants were awarded one-half hour of research credit through the Psychology Subject Pool for each sampling day, given at the end of each day they completed the

sampling. They also received $10 for completing all sampling days, given after the final sampling day.

Daily report (DR) group. Participants in the daily report group completed the PANAS (*today* timeframe), STS-T, and NIEQ-T about their inner experience on five days. On average, DR participants completed experimental manipulation phase procedures in 13.73 days (SD = 2.39, ranging from 7 to 17 days). Participants chose on which days they would prefer to complete the questionnaires. They received two electronic reminders via email or text (8:00 PM and 9:00 PM) to complete the daily report questionnaires. The reminders included a link to the online survey. Participants were awarded one-half hour of research credit through the Psychology Subject Pool for every two completed daily reports and $5 after completing all five days of daily reports.

Phase 3: Post-test

In the final (post-test) phase, all participants again completed the NIEQ, PANAS, and STS. On average, participants took 19.60 days (SD = 4.04, ranging from 7 to 30 days) to complete all study procedures. After completing the final questionnaires, participants were debriefed on their participation, given the opportunity to ask any additional questions, and received $10. Participants in the DR group were also given one hour of research credit through the Psychology Subject Pool (the DES and ESM groups had already been awarded sufficient research credit for their course requirements).

Data Analysis

DES Analysis

The DES analysis followed the usual DES procedure. The first step in analyzing DES results was to create an idiographic characterization for each participant. Once a participant in

the DES group completed the expositional interview for the fifth sampling day, and the contemporaneous descriptions were written, investigators scheduled a meeting to start developing the idiographic characterization for that participant. In the meeting, each contemporaneous description was reviewed and recurring, salient characteristics for that participant were identified. Often, but not always, a participant's sampled experiences predominantly fell in one to three categories of phenomena. Besides identifying which phenomena occurred most frequently for participants, qualitative features, including clarity, intensity, and anything else of note were discussed as salient characteristics. After the meeting, each investigator independently summarized the salient characteristics. One investigator then referred to each investigator's summary in writing a detailed idiographic characterization, including frequency counts of different phenomena, qualitative features, and the contemporaneous descriptions that exemplified these salient characteristics.

Additionally, all contemporaneous descriptions were independently coded for the five frequent phenomena (5FP; inner seeing, inner speaking, feelings, sensory awareness, and unsymbolized thinking) by at least two investigators. For each sample, each of the 5FP were coded as 1 (present), .5 (partially or possibly present), or 0 (absent). So, a single sample may have a 1 for each category, or a 0 for each category, or anything in between. After independent coding, the investigators rectified discrepancies by reviewing each sample in which there was a discrepancy, referencing beep descriptions, personal notes, and video recordings of interviews as necessary. The investigators discussed each discrepant coding to reach a consensus; if a consensus was not possible, a coding of .5 was awarded to indicate the phenomena in question as partially or possibly present.

For this study, six participants' samples were independently coded by two investigators, five participants' samples were coded by three investigators, and one participant's samples were coded by four investigators. The number of raters depended on the number of investigators involved in the sampling interviews. Prior to the rectification procedure, the investigators showed acceptable inter-observer reliability in their codings (Spearman-Brown median typical reliability estimate = .82, range from .72 to .97). After rectification, all investigators agreed on the assigned codings.

The coded 5FP frequencies of the DES analysis were the primary results used for comparison to the other methods. All DES analysis was completed prior to data analysis for the questionnaire data from the other phases and conditions. The interviewers and coders were blinded to the results of the pre- and post-test data. This helped protect against potential researcher bias in coding and idiographic characterization creation.

Quantitative Comparisons

We utilized a mixed between-within, multivariate analysis of variance (MANOVA) to examine effects of experimental condition (between-subjects independent variable) and study phase (within-subjects independent variable) on the average frequencies of each of the 5FP (five dependent variables). When appropriate, we conducted Bonferroni post hoc tests to determine the source of significant main effects and interaction effects. We set the following *a priori* hypotheses to guide our data analysis:

Hypothesis 1: In the experimental manipulation phase, we expected participants in the DES group to report significantly lower frequencies of the 5FP based on DES sampling than would participants in the other groups based on ESM sampling or DR reports. We did not expect any statistically significant differences in frequencies

of the 5FP between the ESM sampling and DR reports. We expected this difference because we believe the questionnaire-based reporting of ESM and DR does not adequately account for the hazards of studying pristine inner experience, resulting in an indefinable mixture of pristine inner experience, memory errors, miscommunications, presuppositions, etc.

Hypothesis 2: We expected participants in the DES group to report significantly lower frequencies of the 5FP based on DES sampling during the experimental manipulation phase compared to their pre-test and post-test reports. We did not expect any statistically significant differences in frequencies of the 5FP across phases for ESM sampling and DR reports. We expected this pattern of results because we believe that the hazards that impact *in general* reports of inner experience are still present and problematic for short-timeframe questionnaires.

Hypothesis 3: We expected post-test NIEQ, PANAS, and STS scores to be significantly lower than pre-test scores for the DES group, whereas we did not expect these differences for the ESM and DR groups. We expected these differences for the DES group as a result of in-depth training in apprehending to momentary experience during the experimental manipulation phase, whereas the participants in the ESM and DR groups did not receive such training as part of their experimental manipulation phase experiences.

CHAPTER 5: RESULTS

We will first describe the process used to get the data ready for comparisons between groups, including accounting for missing data and for multiple time points within the experimental manipulation phase. Next, we will describe the results of the omnibus mixed between-within subjects MANOVA looking at main effects of condition and time, as well as interaction effects. Then, we will describe the results of the specific post hoc analyses that examine each of the three hypotheses outlined above. Finally, we will describe some qualitative differences observed between participants in the DES group.

Data Cleaning

Data collection during Phase 2 was monitored for adherence to procedures for all groups. For the DES group, the total number of samples collected was targeted at 288 (excluding Day 1 samples, 24 samples × 12 participants). In actuality, 270 samples were described (the participants were interviewed about their experience and the data were included in the study), while 280 samples were collected (10 samples were collected by participants but excluded from analyses due to methodological concerns, e.g., the patient had walked away from the beeper and returned to find the beep going off). This would give a conservative adherence rate of 93.75% and a liberal adherence rate of 97.22%. For the ESM group, the total number of surveys completed was targeted at 690 (30 surveys × 23 participants). Of those surveys, 30 (4.35%) were missing, giving an adherence rate of 95.75%. For the DR group, the total number of surveys completed was targeted at 110 (5 surveys × 22 participants). Of those surveys, 1 (0.91%) was missing and 9 (8.18%) were completed the day following the reminders (instead of at the end of the day), giving an adherence rate of 90.91%. DR data completed the following day was retained, as analyses excluding the data did not yield different results.

If a participant was missing an entire day of data in any group, the mean across all participants was imputed to minimize impact of missing data on between group differences. If (as was more often the case for DES and ESM participants), a single beep was missing from a day, the average for that day was computed based on the other four samples/surveys.

Before examining our data across experimental conditions or phases, we needed to determine how Phase 2 data should be used in our overall analysis. Note that we exclude Day 1 reports from all groups, because DES procedure requires excluding Day 1 samples to allow sufficient training in the method; following this procedure for ESM and DR data allows for a more consistent comparison between methods, providing an equal number of data points for each participant. After excluding Day 1 data, we conducted a repeated measures MANOVA to investigate if there were significant within-subject differences between Days 2-5 (IV) in the various 5FP (five DVs). All results were non-significant (alpha set to .01, to correct for family-wise Type I error). As there were no significant differences within subjects across Days 2-5, we collapsed the results from these days into a single average Phase 2 percent for each of the 5FP for every participant.

Quantitative Comparisons

Descriptive statistics for each of the dependent variables are presented in Table 1 (located in the Appendix). The first column of the table lists the dependent variable (5FP) and the way that dependent variable was measured (NIEQ, DES, NIEQ-ATM, NIEQ-T, PANAS, or STS). The second and third columns list the phase of data collection and experimental condition, respectively. Columns four through seven provide the mean, standard deviation, minimum, and maximum average frequencies of the 5FP, respectively. For example, the first row of Table 1 shows that participants assigned to the DES group reported experiencing inner speaking 64.17%

of the time, on average, when using the NIEQ during the pre-test. Figures 1-5 provide a visual representation of this data, in which each dot on the graphs represent an average frequency of that 5FP, for that group, during that phase.

Figure 1. Inner Speaking across Phases by Group.

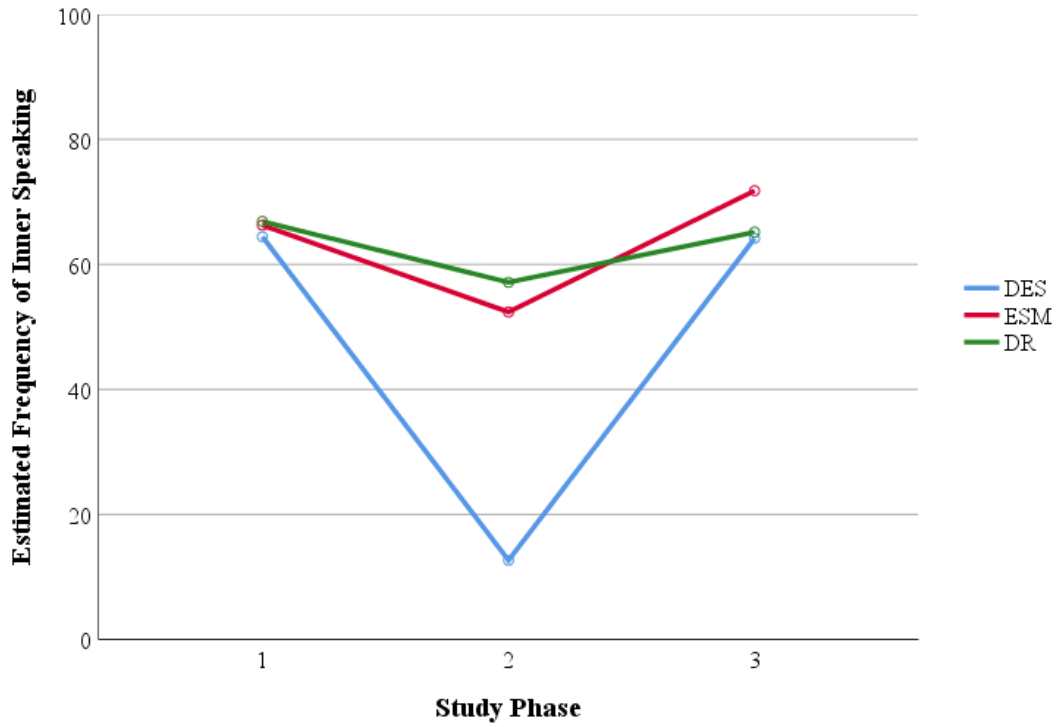


Figure 2. Inner Seeing across Phases by Group.

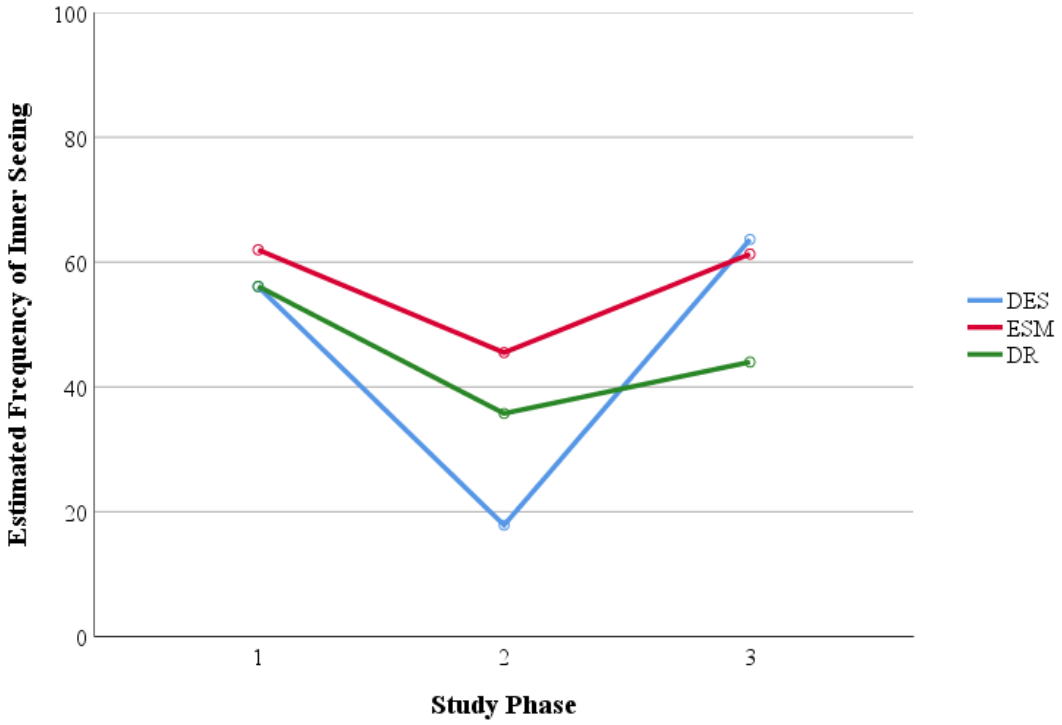


Figure 3. Feelings across Phases by Group.

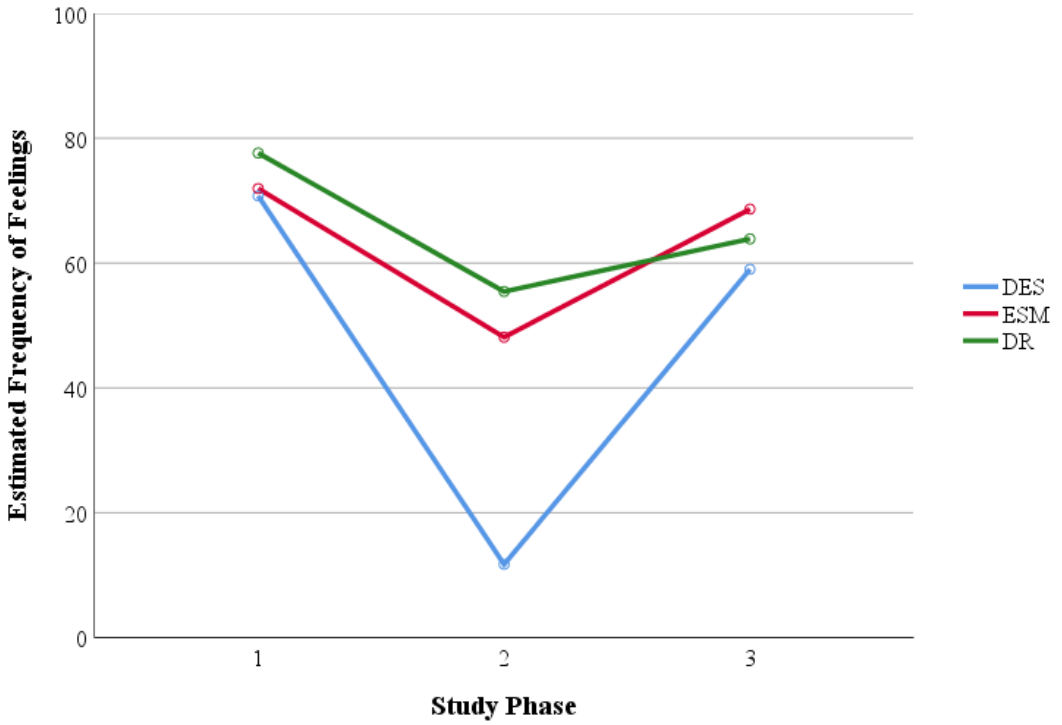


Figure 4. Sensory Awareness across Phases by Group.

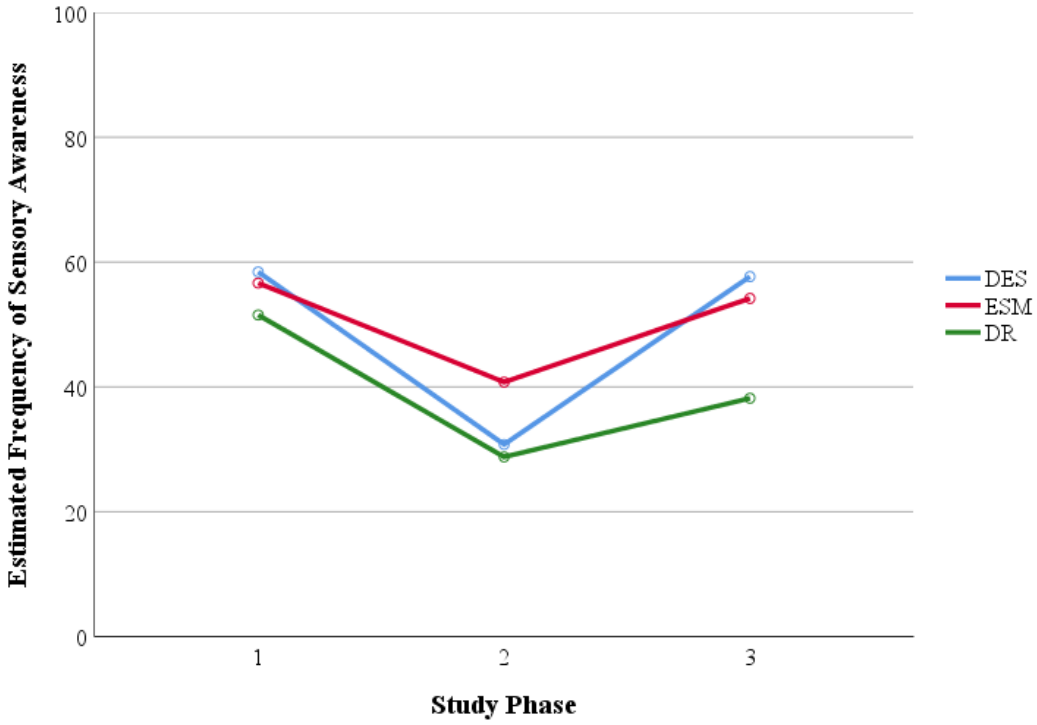
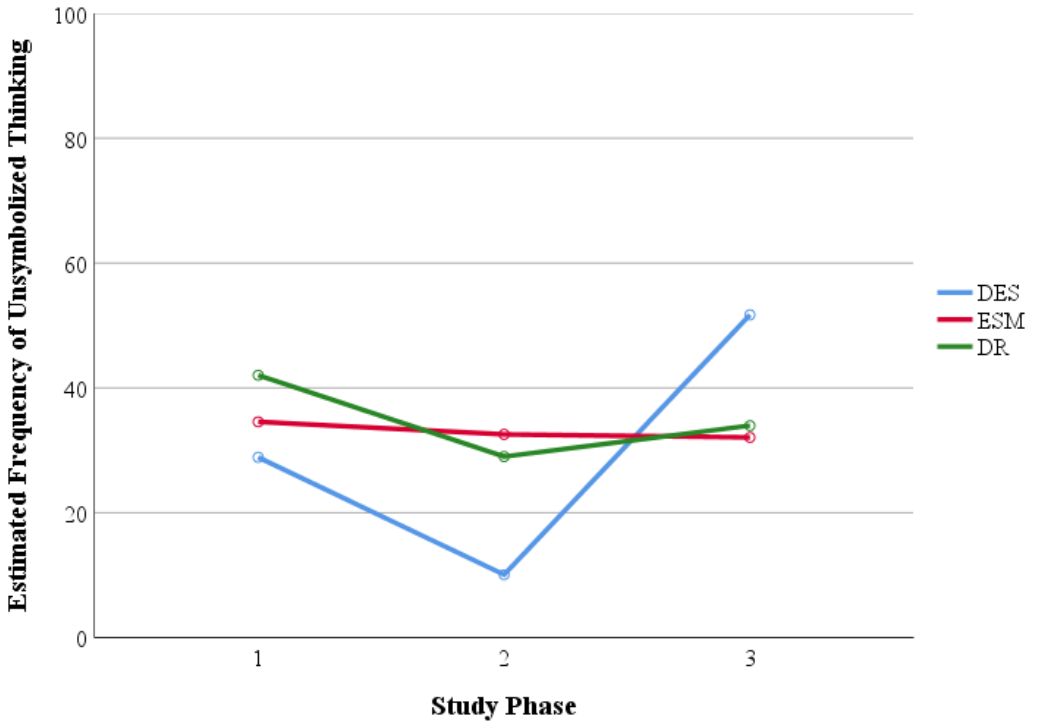


Figure 5. Unsymbolized Thinking across Phases by Group.



To determine if randomization was effective in eliminating group differences on outcome measures, we conducted a between-subjects MANOVA on Phase 1 NIEQ, PANAS, and STS data. No significant differences were found between the groups (F's ranged from .056 (Inner Speaking) to 2.324 (PANAS Positive Affect), p = .946, .108, respectively). Alpha was set to .05 to provide a conservative test for positive evidence of randomization. Thus, we concluded that randomization had indeed been effective.

Next, we conducted a mixed between-within subjects MANOVA on average frequency of the 5FP (measured by the NIEQ, NIEQ-ATM, NIEQ-T, or DES) to determine if overall differences between experimental conditions and phases were present, which would allow for more targeted analyses testing each hypothesis. When appropriate, we conducted post hoc t-tests to determine the source of significant main effects and interaction effects. Alpha was set to .01 for these analyses to decrease risk of Type I error. Results of the omnibus MANOVA are presented in Table 2. This first column provides the kind of comparison being conducted, with Phase indicating within-subject, Group indicating between-subject, and Group × Phase indicating the interaction. The second column shows which of the 5FP are being compared. The third through fifth columns provide the degrees of freedom for the analysis, the F statistic, and the effect size. Significance is indicated with asterisks. For example, the first row of Table 2 shows there was a significant main effect of study phase for Inner seeing, such that there are significant differences with a very large effect size in the average frequency of inner experience across the three study phases ($F(2,108) = 82.12$, $p < .001$, $\eta^{2=}$ .60). We are reporting partial-eta squared for effect sizes. Note that we are aware there are some concerns about the generalizability across studies of partial-eta squared (Olejnik & Algina, 2003); however, the main concern is in

underestimating effect size due to the absence of a blocking component in the research design. This is something for future studies to consider when examining our results.

Overall, we found omnibus effects showing significant within-subjects differences of time for each of the 5FP. Significant interactions between group and time were found for all 5FP except for Sensory Awareness. These results provide the basis to more closely examine each of our hypotheses and identify the source of these differences.

Table 2

Main effects and interaction effects of mixed MANOVA of 5FP

Comparison	**5FP**	**df**	**F**	**Partial-eta squared**
Phase	Inner Seeing	2, 108	82.12**	.60
	Inner Speaking	2, 108	38.11**	.41
	Feeling	2, 108	101.93**	.65
	Sensory Awareness	2, 108	31.69**	.37
	Unsymbolized Thinking	2, 108	10.782**	.17
Group	Inner Seeing	2, 54	4.58	.15
	Inner Speaking	2, 54	2.65	.09
	Feeling	2, 54	4.43	.14
	Sensory Awareness	2, 54	2.22	.08
	Unsymbolized Thinking	2, 54	0.28	.01
Group × Phase	Inner Seeing	4, 108	18.27**	.40
	Inner Speaking	4, 108	5.79**	.18
	Feeling	4, 108	12.62**	.32
	Sensory Awareness	4, 108	2.083	.07
	Unsymbolized Thinking	4, 108	6.92**	.20

*$p < .01$
**$p < .001$

Hypothesis 1: In Phase 2, we expected participants in the DES group to report significantly lower frequencies of the 5FP than participants in the ESM or DR groups. We did

not expect any statistically significant differences in frequencies of the 5FP between the ESM and DR group participants.

Given the omnibus results showing a significant interaction between experimental condition and study phase, we conducted an independent samples t-test to compare groups during Phase 2 to test our first hypothesis. Results of this analysis are presented in Table 3. The first column indicates which of the 5FP is being compared. The second, "Comparison" column indicates which measures of the 5FP are being compared as well as the direction of the comparison (*A-B*). The third through fifth columns provide the degrees of freedom, t statistic, and the effect size (post hoc t-test effect sizes are reported using Cohen's d). For example, the first row of Table 3 shows that DES reports of inner speaking were significantly lower than NIEQ-ATM reports with a very large effect size ($t(33) = -6.15$, $p < .001$, $d = -2.19$).

Overall, hypothesis 1 was supported for all 5FP except Sensory Awareness. DES estimates of 5FP frequencies were significantly lower on average than NIEQ-ATM and NIEQ-T estimates of 5FP frequencies for Inner Speaking, Inner Seeing, Feeling, and Unsymbolized Thinking. There were no statistically significant differences between average frequencies collected using ESM and DR for any of the 5FP (Sensory Awareness came closest to a significant difference: $t(43) = 1.90$, $p = .065$, $d = .57$).

Table 3

Mean differences between Groups for Phase 2 data

5FP[a]	Comparison		df	t	Cohen's d
	A	*B*			
Inner Speaking	DES[b]	NIEQ-ATM[c]	33	-6.15**	-2.19
	DES	NIEQ-T[d]	32	-7.01**	-2.53
	NIEQ-ATM	NIEQ-T	43	-.82	-0.24
Inner Seeing	DES	NIEQ-ATM	33	-4.32**	-1.52
	DES	NIEQ-T	32	-2.88*	-1.03
	NIEQ-ATM	NIEQ-T	43	1.85	0.55
Feeling	DES	NIEQ-ATM	33	-5.25**	-1.87
	DES	NIEQ-T	32	-6.52**	-2.34
	NIEQ-ATM	NIEQ-T	43	-1.11	-0.33
Sensory Awareness	DES	NIEQ-ATM	33	-1.23	-0.44
	DES	NIEQ-T	32	.28	0.10
	NIEQ-ATM	NIEQ-T	43	1.90	0.57
Unsymbolized Thinking	DES	NIEQ-ATM	33	-4.33**	-1.54
	DES	NIEQ-T	32	-2.95*	-1.06
	NIEQ-ATM	NIEQ-T	43	.64	0.19

Note. Negative t-values indicate column *A* reports were smaller than column *B* reports.
[a]5 frequent phenomena
[b]Descriptive experience sampling-based reports (DES group)
[c]Nevada Inner Experience Questionnaire-At the Moment (ESM group)
[d]Nevada Inner Experience Questionnaire-Today (DR group)
*$p < .01$
**$p < .001$

Hypothesis 2: We expected participants in the DES group to report significantly lower frequencies of the 5FP during Phase 2 compared to their in general reports during Phases 1 and 3. We did not expect any statistically significant differences in frequencies of the 5FP across timepoints for participants in the ESM or DR groups.

Given the omnibus results showing a significant main effect of study phase, we conducted a dependent samples t-test to phases within groups to test our second hypothesis. No significant differences were found between Pre-test and Post-test reports (more fully explained in relation to hypothesis 3 below), so to simply the comparison we averaged Pre-test and Post-test

reports of the 5FP for each participant. We then compared Experimental manipulation phase reports with the averaged *in general* reports using the dependent samples t-test. Results of the dependents samples t-test are presented in Table 4. The first column indicates which measure of the 5FP (and thus which group) is being compared to the averaged *in general* reports. The second column indicates which of the 5FP is being compared. The third through fifth columns provide the degrees of freedom, t statistic, and the effect size (measured with Cohen's d). For example, the first row of Table 4 shows that DES-based reports of the frequency of inner speaking were significantly lower than *in general* questionnaire-based reports of inner experience, with a very large effect size ($t(11) = -8.53$, $p < .001$, $d = -3.43$).

Overall, hypothesis 2 was partially supported. 5FP frequency gathered using DES during Phase 2 was significantly different from the DES group's averaged *in general* NIEQ reports for all 5FP except Unsymbolized thinking (Phase 1 and Phase 2 were not significantly different; Phase 2 and Phase 3 were significantly different) and Sensory Awareness (while there was a significant main effect of time, post hoc tests were non-significant with alpha set to .01.

Table 4

Comparing Experimental Phase to *In General* Questionnaires by Group

Measure (Group)	**5FP**[a]	**df**	**t**	**Cohen's d**
DES-based reports[b]	Inner Speaking	11	-8.53**	-3.43
(DES)	Inner Seeing	11	-7.01**	-2.56
	Feeling	11	-9.24**	-3.30
	Sensory Awareness	11	-2.82[ǂ]	-1.23
	Unsymbolized Thinking	11	-3.81*	-1.76
NIEQ-ATM[c]	Inner Speaking	22	-6.32**	-0.87
(ESM)	Inner Seeing	22	-3.77**	-0.77
	Feeling	22	-6.12**	-1.06
	Sensory Awareness	22	-4.50**	-0.61
	Unsymbolized Thinking	22	-.19	-0.04
NIEQ-T[d]	Inner Speaking	21	-3.40*	-0.49
(DR)	Inner Seeing	21	-4.79**	-0.79
	Feeling	21	-5.36**	-0.75
	Sensory Awareness	21	-4.10**	-0.96
	Unsymbolized Thinking	21	-2.49	-0.40

Note: Frequencies reported in Pre-test and Post-test phases were averaged and compared via paired-samples t-test to Experimental manipulation reports: DES-based, NIEQ-ATM, or NIEQ-T. Negative numbers indicate Experimental manipulation reports were lower than *in general* questionnaire reports.

[a]5 frequent phenomena

[b]Descriptive experience sampling-based reports (DES group)

[c]Nevada Inner Experience Questionnaire-At the Moment (ESM group)

[d]Nevada Inner Experience Questionnaire-Today (DR group)

*$p < .01$

**$p < .001$

[ǂ]$p = .017$

Contrary to hypothesis 2, NIEQ-ATM and NIEQ-T estimates of 5FP frequency during the Experimental phase also showed significant differences from *in general* reports. For NIEQ-ATM and NIEQ-T reports, significant differences were observed for all of the 5FP except Unsymbolized Thinking. Please see Table 4 for statistics.

Importantly, the effect sizes of the differences in estimated 5FP frequency between the Experimental phase and *in general* reports were quite different between groups. As you can see in Table 4, DES group effect sizes for significant differences ranged from d = -1.76 to -3.43;

ESM group effect sizes for significant differences ranged from d = -.61 to -1.06; DR group effect sizes for significant differences ranged from d = -.49 to -.96. For unsymbolized thinking (non-significant differences for all groups) the effect sizes for DES, ESM, and DR groups were d = -1.76, -.04, and -.40, respectively. In each instance, the minimum effect size observed in DES was much larger than the maximum effect size observed in the ESM and DR groups. On average, DES effect sizes were 3.5 times as large as those observed for the ESM and DR groups. This difference in effect size is depicted visually in Figures 1-5. For example, in Figure 1 depicting the average frequencies of inner seeing, we can see that the magnitude of the difference between the Experimental phase and the Pre-test and Post-test phases for the DES group is much larger compared to these differences for the ESM and DR groups.

Thus, the differences in reported 5FP frequencies between the Experimental phase and the *in general* reports of Pre-test and Post-test were not insignificant for the ESM and DR groups, as was hypothesized. Still, the differences in reported 5FP frequencies between the Experimental phase and the *in general* reports of Pre-test and Post-test for the DES group were consistently larger (based on effect size) than those found for the ESM and DR group, regardless of statistical significance.

Hypothesis 3: We expected Phase 3 (post-test) in general reports of 5FP frequency, PANAS Positive and Negative Affect, and STS scores to be significantly lower than Phase 1 (pre-test) in general reports for the DES group, whereas we did not expect these differences for the ESM and DR groups.

Hypothesis 3 was not supported for any of the measures except for the NIEQ estimates of the frequency of Feelings (mean difference = 9.60, $p < .001$, 99% CI [3.14, 16.06]). DES and DR participants reported similar decreases in the average frequency of Feelings from Phase 1 to

Phase 3 (mean difference = -11.75 and -13.75, respectively), while ESM show a much smaller change (mean difference = -3.31).

Summary of Quantitative Comparisons

During the Experimental manipulation phase, participants reported significantly lower average frequencies of each of the 5FP using DES (except Sensory Awareness) compared to NIEQ-ATM and NIEQ-T reports. All participants, regardless of experimental condition, reported significantly lower average frequencies when reporting on their experience using *at the moment of the beep* or *today* timeframes (DES reported 5FP, NIEQ-ATM, and NIEQ-T) compared to an *in general* timeframe (NIEQ), with a few exceptions. The effect sizes for the differences between timeframes were about twice as large for DES-based reports compared to NIEQ-ATM and NIEQ-T reports. There were not consistent differences between pre-test and post-test NIEQ, PANAS, or STS reports of inner experience after participants had practice attending to and reporting on their experience using shorter timeframes, regardless of whether they received intensive training with DES or participated in ESM or DR.

Descriptive Comparisons

The quantitative results reported above demonstrate significant differences between methods in terms of the average frequency with which participants report experiencing specific phenomena. However, even more striking is the difference in the kind of the information gathered.

ESM participants were asked to provide a brief, free text description of what they were doing at the moment of the beep in addition to answering the questions about the presence of the 5FP. An example of such a response set is one participant's entry for Day 2 Beep 6: "on social media looking through information"; inner seeing—14%; 21% inner speaking—21%; feeling—

26%; sensory awareness—38%; and unsymbolized thinking—52%. This is the extent of the data collected for that beep, and the method of collection is in line with standard ESM procedures.

On the other hand, DES allows for a rich description of participants' inner experience. Here is an example from "Liana" (name changed to protect confidentiality) who was also on social media on Day 5 Beep 2 (information in brackets indicate context that was not directly present in experience at the moment of the beep):

> [Liana was reading the caption of an Instagram post of one man asking another man wearing a prosthetic arm, "Can you give me a hand?" in which the man with the prosthetic arm throws it.] At the moment of the beep, Liana was innerly seeing a prosthetic arm floating mid-air to the left of a grocery store freezer with Styrofoam coolers on top of the freezer. The seeing is in color; the freezer and coolers are blurry, while the prosthetic arm is clear but not very detailed. [The grocery store freezer/cooler background was created by Liana in her imagination as the original Instagram post had nothing to do with a grocery store]. The inner seeing accounts for approximately 60% of Liana's experience.
>
> Simultaneously, Liana is reading the word "*practiced*" [which is italicized in the original]. Liana is reading "*practiced*" separate from the rest of the sentence that it is a part of—that is to say, she is reading the word "*practiced"* on its own and not part of a larger sentence. The reading is present without an inner speaking, in that Liana is simply reading the word.

This kind of detail and differentiation is not possible to gather using ESM or DR reports. To further exemplify the kind of information that can be uniquely gathered by DES, we will take a closer look at two DES participants.

Kendra and Isaac

Kendra and Isaac were recruited and engaged in sampling at approximately the same time. They were different from each other demographically (gender, race, and age), but were demographically similar to other participants (i.e., neither was clearly part of a different population). Both Kendra and Isaac were consistent in how they reported their inner experience throughout sampling as well as clearly engaged in the procedure. Both openly expressed enjoying the experience and feeling as though they learned something about themselves. Yet, they were strikingly different from each in the qualitative nature of their inner experience.

Throughout sampling**,** Kendra's experience was consistently unclear and seemed difficult for her to apprehend and describe. On the first four days we were never sure she grasped the concept of the moment of the beep, or that she grasped the concept of experience. It seemed that her response to each beep was to survey her environment and report about the result of that survey as if it had been her experience at the moment of the beep. Here is an example from sampling day 4:

> 4.4 – [Kendra was at work and had just ordered a burger for lunch.] In the interview, she initially said that at the moment of the beep, she was somehow cognitively aware that she was about to eat, and that she knew that she was hungry. However, she could not provide any details of cognitive experience. Then she said she was feeling hungry, which she described as a bodily feeling in her stomach (but without any specifics about those sensations). Then she said she

> experienced a low energy level, again without detail. Then she reported that she was watching her burger being made, seeing the burger on the grill, hearing the sizzle of the pan, watching the bread being toasted, watching the tomato and lettuce being dressed, and watching the sandwich being assembled. In reality, those things would take place on the order of minutes apart, not all at a simultaneous moment. Then she reported that watching the cooking was making her hungrier.

It seemed that Kendra was reporting events that happened in the general vicinity of the beep, not that which was directly experienced—caught in flight—at the moment of the beep. Furthermore, it seemed that she was giving plausible descriptions of what experience might be like when those events occurred, rather than noticing experience that actually presented itself to her. This is relatively unusual after 4 days of sampling.

By contrast, Isaac had clear experience. Starting on his second sampling day, his reporting of his experience was straightforward. He had a variety of inner experiences, including inner seeing, unsymbolized thinking, feeling, inner speaking, nothing/doing, and perhaps one or two instances of sensory awareness. He seemed readily able to apprehend and describe his experience and was nuanced and careful in his descriptions. Here is an example from Isaac's day 4:

> 4.5 – [Isaac was getting a rocking chair out of his garage to give to someone. In the way was a storage chest. He wondered if his lost nerf gun could be in the chest, and then] he innerly saw the drawer of the chest open [in actuality it was closed] with the nerf gun and some other stuff in the drawer. The gun was pointing to the left, the barrel under a piece of cloth, and other stuff seen

indistinctly. There was some hopeful/excited sense of this seeing, but he was not sure that was in his experience. [During the interview, Isaac expressed slight worry that the cloth that he recalled seeing was a confabulation because when he actually opened the drawer there were some sheets in it. He thought the after-the-beep seeing may have distorted his recollection of the at-the-moment-of-the-beep experience. He's confident he saw the drawer and the gun and some other stuff, just not whether the stuff included the fabric.]

Isaac was particularly aware of and able to express how after-the-moment experience may have impacted his at-the-moment recollection, something not often seen in sampling. He was attuned to his inner experience and able to clearly differentiate between aspects in which he has full confidence and aspects of which he is less sure.

Because of the descriptive difficulties up to day 4, we speculated that Kendra had no inner experience at most or all of her beeps, and that it was this lack of experience that had led to her difficulty or impossibility of cleaving to the moment and to experience. On day 5 we became more confident of that no-experience interpretation, as she now seemed to be better able to cleave to the moment of the beep and now reported mostly nothing as present in her experience. We worked hard in the interviews to leave room for experience, even after we began to suspect that Kendra does not have inner experience, at least not in the way most people do. We never suggested to Kendra that perhaps she had no inner experience to report—we simply asked for the details of what she described as experience.

After day 5, in an informal experiment, the investigators showed her some well-known ambiguous (alternating) figures (faces/vase, Neckar cube, Jastrow's duck-rabbit, Boring's old woman/young woman). The figures did not alternate for her; she reported being able to see both

figures, but only if she closed one of her eyes or covered up part of the image. The investigators believe that alternation is a characteristic of experience—the objective external reality (retinal display, etc.) of, say, Jastrow's duck stays constant when the experience becomes of Jastrow's rabbit. If one does not have experience, then ambiguous figures will not alternate. The lack of alternation for Kendra thus supports the no-experience characterization of Kendra. Therefore, we think the best interpretation of all her samples is that she did not have inner experience in the way that most people have inner experience, as measured by DES.

By contrast, an unusual aspect of Isaac's experience was that it occasionally included "deep thoughts" rarely captured by DES. These thoughts were about the nature of society and/or people. For example, on sampling day 2, beep 3, Isaac was thinking without words or symbols a thought that could be expressed: *In a society we have rules and beliefs, formal and informal. Going against those, even if the belief is wrong, upsets society.* Most experiences by most DES participants are substantially shallower and mundane, as characterized by the participants themselves (and by DES investigators).

Also unique to Isaac in this group of participants was his curiosity about the nature of the study. At the end of his participation, he inquired about the difference between using questionnaires and DES to ask about experience. He spontaneously reported (after completing his Phase 3 questionnaires) that he felt very uncertain whether his responses on the questionnaires would be understood correctly. He stated that he felt there was a good chance that someone would misinterpret the meaning of his questionnaire responses and thus not have a good understanding of the nature of his inner experience based on those questions and his answers. This uncertainty seemed to arise (Isaac thought) as the result of his efforts in Phase 2 at

carefully describing his experience; that is, he had learned in Phase 2 about the ambiguity of his own reports about his own experience.

Table 5

Average Frequency of the 5FP for Kendra and Isaac

		Inner Speaking	Inner Seeing	Feeling	Sensory Awareness	Unsymbolized Thinking
Pre-test	Kendra	100.00%	86.50%	77.00%	92.00%	100.00%
	Isaac	59.00%	68.00%	39.00%	56.50%	1.00%
Experimental	Kendra	0.00%	0.00%	10.4%	8.3%	2.1%
	Isaac	14.29%	28.57%	26.19%	7.14%	40.48%
Post-test	Kendra	91.50%	88.00%	68.00%	100.00%	100.00%
	Isaac	37.50%	76.00%	42.50%	60.50%	37.50%

Note. Pre-test and Post-test average frequencies are measured by the NIEQ. Experimental phase average frequencies for the 5FP are average scores across days, not including Day 1.

Table 5 presents Kendra's and Isaac's frequency of experiencing each of the 5FP across the three phases. The table shows that there seems to be something different about their inner experience. Kendra reports high frequencies of inner experience using questionnaires and has very low frequencies of inner experience as obtained using DES. Isaac shows a similar, but less drastic, pattern of results. Both participants show a similar pattern of results to the rest of the DES group, in that their pre- and post-test *in general* reports of inner experience are high and similar to each other, and their DES reports are lower. An exception, also in line with the DES group results, is Isaac's report of unsymbolized thinking: low on the pre-test and much higher on the post-test. However, there is nothing in Table 5 that suggests that Kendra and Isaac have

drastically different inner experiences (or lack thereof). That kind of result depends entirely (at the current state of the art) on the ability to apprehend experience in high fidelity.

CHAPTER 6: DISCUSSION

This study compared reports of inner experience for 57 undergraduate students randomly assigned to one of three methodological groups: the descriptive experience sampling group (DES), the (non-DES) experience sampling group (ESM), and the daily report group (DR). We used a pre-, post-test design to directly compare the three methods to each other and to *in general* reports of inner experience. The primary phenomena compared were those commonly found using DES: the five frequent phenomena (5FP: inner speaking, inner seeing, feelings, sensory awareness, and unsymbolized thinking; Kühn et al., 2014)

Our first hypothesis, that participants in the DES group would report lower frequencies of the 5FP than would participants in the ESM and DR groups, was supported for four of the 5FP (no difference for sensory awareness). No significant differences were found between the ESM and DR groups for any of the 5FP. If we accept the DES reports as the most valid given the care taken in DES to confront the challenges of introspection, then these findings suggest that simply shortening the reporting delay and reducing retrospective requirements does not sufficiently limit the impact of the other hazards of studying pristine inner experience. Thus, even short-timeframe questionnaires are subject to the influence of presuppositions and other biasing factors.

Our second hypothesis, that participants in the DES group would report lower frequencies of the 5FP when using DES as compared to *in general* pre- and post-test reports, was supported for four of the 5FP (no difference for sensory awareness). Contrary to our hypothesis, participants in the ESM and DR groups also reported lower frequencies of the 5FP using *at the moment* and *today* timeframes compared to their *in general* reports. However, the magnitude of the differences in the reported frequency of phenomena between *in general* reports and *at the moment* or *today* reports were on average about twice as large for participants using DES than

for participants using (non-DES) experience sampling or daily report. This suggests that a short timeframe helps to focus their reports on recent experiences when compared to *in general* reports, but that questionnaires with any timeframe likely measure some combination of inner experience and presuppositions.

Our third hypothesis, that post-test *in general* reports of inner experience would be significantly lower than pre-test reports for DES group participants, was not supported. This finding suggests that *in general* reports primarily rely on presuppositions and expectations about experience. Furthermore, *in general* descriptions are not readily amendable to change, even when participants are trained in careful apprehension and description of pristine inner experience.

Based on the results of our study, we have three main take-aways: 1) DES is measuring something different from questionnaires, regardless of the timeframe used; 2) Qualitative descriptions of experience provide valuable information not able to be gleaned from quantitative measures; and 3) These results have implications for how we understand existing psychological research based on self-report questionnaires, the development of psychological theories, and clinical applications of research.

DES Measures Something Different than Questionnaires

Our results support the assertion that DES measures something different from self-report questionnaires, regardless of the timeframe. First, the direct comparison between the *in the moment* and *today* questionnaire timeframes and DES showed significant differences in the expected direction, with very large mean differences (the average mean difference was larger than 25%). We believe the unique training offered by DES in apprehending and describing inner experience as it was occurring *at the moment of the beep* minimizes the degree to which

presuppositions, memory errors, and assumed processes influence the reported experience. Because participants in the ESM and DR groups received no such training, they were unable to distinguish between what they generally think about themselves and their experience, what they expect to be present based on the situation, what they remember about the moment or the day when they get around to filling out the questionnaire, and pristine inner experience (Hurlburt & Heavey, 2015).

Second, the lack of difference between frequency of inner experience reported using *at the moment* and *today* questionnaires further supports the assertion that DES is measuring something different. If it were only a matter of timeframe, we would expect the ESM group and the DES group reports to be essentially equivalent, as they both ask participants to report on experience *at the moment of the beep*. No difference between *at the moment* and *today* timeframes has been shown in previous research (Robinson & Clore, 2002b). Robinson and Clore (2002b) argued that a long timeframe results in participants' reliance on semantic, as opposed to episodic, memory in reporting on their recent experiences. However, this study shows that simply having a short timeframe is not sufficient to protect against influences on self-report, such as the intrusion of semantic memory, or any of the other hazards of studying pristine inner experience.

Relatedly, Robinson and Clore (2002b) also found that self-report questionnaires using *at the moment* and *today* timeframes obtained significantly different results than did questionnaires with *in general* timeframes. Our results replicated this finding; however, we also went further to show that DES reports were more consistently significantly different and had effect sizes about twice as large as the questionnaire methods. Therefore, although DES and short-timeframe questionnaires all seem to measure something distinct from what is measured using an *in general*

timeframe, DES seems to measure something different from questionnaires regardless of timeframe. Essentially, *in general* reports seem to primarily result in participant reports based on semantic memory and presuppositions; short-timeframe questionnaires seem to result in a mix of episodic memory for specific instances of inner experience, semantic memory, and presuppositions; and DES seems to result in primarily reports of pristine inner experience. These differences are likely due to a number of factors, such as training in apprehending and describing experience, the relative impact of semantic memory on reports, as described earlier, and presuppositions about experience, the self, the situation, etc. (Hurlburt & Heavey, 2015).

The Exceptions of Sensory awareness and Unsymbolized Thinking

Notably, sensory awareness and unsymbolized thinking did not consistently show the same pattern of results as did the other 5FP. As shown in Figure 4, participants in all groups showed a relatively similar decrease in frequency of reported sensory awareness from Phase 1 to Phase 2: ESM had the smallest mean difference between phases at 15.89%, whereas DES had the largest mean difference at 27.63%, a range of about 12%. This is strikingly different from the much larger decrease found in DES reports from Phase 1 to Phase 2 compared to ESM and DR reports for other of the 5FP. Taking inner speaking as an example, shown in Figure 1, DR had the smallest mean difference between phases at 9.75%, while DES had the largest mean difference at 51.76, a range of about 42%.

We can think of three possible explanations for the lack of significant differences across groups for sensory awareness. First, the Sensory Awareness scale on the Nevada Inner Experience Questionnaire (NIEQ) used to measure the 5FP for the ESM and DR groups is the least psychometrically sound. It has the highest correlations with other scales and the lowest internal consistency. It is possible that these issues with measurement impacted the reports.

Second, there may be something about sensory awareness that makes participants less likely to endorse it on questionnaires compared to the other 5FP. Perhaps participants confuse sensory awareness with feelings, as feelings frequently include bodily sensations; perhaps participants are embarrassed by their specific instances of sensory awareness, and so ignore them or do not want to report on them; perhaps it is something totally different. Third, it is possible that there really is a difference in the population that was not reflected in this particular sample. Our DES sample reported a somewhat higher average frequency of sensory awareness (31%) than what has previously been found (e.g., 22% reported in Heavey & Hurlburt, 2008). Also, we performed a large number of comparisons and thus used a conservative alpha to protect against Type I error, potentially increasing our risk for Type II error.

Unsymbolized thinking did not show a significant difference between Phase 1 (pre-test) and Phase 2 (experimental condition) for the DES group, unlike the other 5FP. To understand this lack of difference, we looked closer at the average frequencies of unsymbolized thinking reported by the three groups. Phase 1 reports of unsymbolized thinking were quite low across all groups and were lowest for the DES group (still not a significant difference between groups in Phase 1). The ESM and DR group participants did not show any change in reports of unsymbolized thinking across timepoints, while the DES group showed a small (non-significant) decrease between Phase 1 and 2, and then a large (significant) increase in Phase 3. We believe this increase is related to the training that DES participants received during Phase 2.

Unsymbolized thinking is an unfamiliar concept to most people, so when someone reads a simple description on a questionnaire, they are probably not entirely sure what it is. Due to various presuppositions (e.g., experimenters being unlikely to ask about something that does not exist, worries about being seen as abnormal if they report no experience of unsymbolized

thinking, not wanting to mislead experimenters by reporting frequent experiences of unsymbolized thinking), participants may have endorsed the item even though they did not understand it. DES participants received training in how to apprehend and describe their inner experience, and some encountered the concept of thinking without words. Thus, when they were again asked to report on it *in general*, they may have no longer had the same presuppositions influencing their responses as they did in Phase 1. As a result, their reported frequencies were similar to the other 5FP, resulting in the sharp increase in unsymbolized thinking observed in Phase 3. The ESM and DR group participants received no such training, and so answered the question the same way throughout the study, resulting in the lack of significant differences between timepoints that were observed for many of the other 5FP.

In General Reports are Not Pristine Inner Experience

Our results showed that all participants, regardless of group, reported similar experience on the pre- and post-test *in general* timeframe questionnaires. Even with intensive training in how to apprehend and describe pristine inner experience, and significant drops with very large effect sizes in reported experience with DES, participants still apparently primarily rely on presuppositions about experience when asked to report *in general*. In hindsight, this is not surprising. The instructions for the *in general* questionnaires require a participant to use semantic, not episodic, memory. Therefore, questionnaires with an *in general* timeframe are fundamentally measuring something other than pristine inner experience. Participants cannot provide reports of inner experience as it actually occurs when asked to report *in general*; they can only provide judgments, which are influenced by presuppositions, memory errors, and cognitive biases about their inner experience.

Considerations of Descriptive Results

By gathering descriptive information about Kendra's and Isaac's inner experience, we were able to describe in detail unique features of their experience that would not be known to exist if we used any other method. If we ignored the descriptive data, and focused only on the coded frequencies from DES, we would not have been able to surmise the extreme differences in the qualitative nature of their inner experience. Perhaps, simply based on the numbers (see Table 5), we would say that Kendra drastically overestimates the amount of her inner experience using *in general* questionnaires relative to what is found by DES. Perhaps we would also say that Isaac became more aware of his tendency to have unsymbolized thoughts by engaging in DES. However, I am unsure if anyone would make those statements without first having the descriptive accounts of their respective inner experiences to compare.

These unique characteristics would not have emerged if either Kendra or Isaac had been randomized to the ESM or DR group instead of the DES group. We have shown that ESM and DR questionnaires are more similar to *in general* reports than to DES. So if we imagine what Table 5 would have looked like had Kendra and Isaac been the ESM group, we would not see anywhere near the change in Kendra and Isaac's reported frequencies of experience, and Kendra likely would have reported higher frequencies of experience during Phase 2 than did Isaac. With these results, one might have said that Kendra has *more vivid* inner experience than does Isaac—her 5FP numbers are higher across the board. There would have been absolutely no clue that Kendra had less or no inner experience.

Despite the benefits of this approach, adequately conveying the descriptive characteristics of each of the 12 participants would be difficult and the meaning could be lost in the length of the details. Thus, it requires careful thought to decide how to best convey the important characteristics captured with DES. This is a far more onerous task than conducting a MANOVA

and summarizing the results. Still, given the kind of information available, we believe it is valuable to include this kind of idiographic analysis in psychological research.

The benefits of questionnaires lie mainly in their ease of administration, scoring, analysis, and distribution of the results. These benefits are a direct result of limiting response options. Supposedly, this also helps to standardize participant response to allow for direct comparison between participants, between groups, and between studies. However, when questionnaires ask about inner experience, we are not getting a high-fidelity description of that inner experience, as evidenced by the wealth of additional information gathered using DES. We also believe we are not getting veridical reports of inner experience frequency, as evidenced by the large differences in the frequencies of experience as measured by questionnaires and DES. Therefore, the standardization of responses leads to an illusion of accuracy, resulting in years of research and theories of psychological constructs based on reports of inner experience that likely are primarily reports of presuppositions, memory errors, and likely even miscommunications.

Implications for Psychology

Veracity of Questionnaire Reports

The discrepancy between DES and questionnaire-reported frequencies of inner experience found in this study raises questions about the validity of conclusions based on results of questionnaire studies. To further clarify these concerns, we will use inner speaking as an example, although we could discuss any of the 5FP. Many researchers claim that inner speaking is a predominant part of pristine inner experience (Klinger and Cox, 1987). However, multiple DES studies have consistently found large individual differences in frequency of inner speaking and an across participant average of roughly 25% (Heavey and Hurlburt, 2008; Hurlburt et al., 2013). Further, DES has directly compared questionnaire data on inner speaking to DES results,

and found no correlation (Heavey et al., 2019; Hurlburt & Heavey, 2015). This finding is strengthened by research validating DES identified inner speaking using fMRI imaging (Hurlburt et al., 2016; Kühn et al., 2014). This suggests that DES is able to identify inner speaking when it is present, but memory errors and cognitive biases inflate estimates of inner speaking given on self-report questionnaires. Thus, there is substantial evidence in this study and elsewhere that questionnaires do not gather sufficiently veridical data on inner experience to allow for accurate conclusions about the frequency or nature of inner experience.

Comprehensive Psychological Theories

Theories about psychological constructs that include aspects of pristine inner experience, such as theories of emotion, may be faulty if their development primarily relied on data from self-report questionnaires. As discussed above with the theory of the independence versus bipolar nature of positive and negative affect, the interpretation of the data may be misleading if the responses are not actually representative of participants' pristine inner experience. If we look at DES data (Heavey et al., 2017), there is evidence that positive and negative affect can be present in someone's experience simultaneously, even if it is a relatively uncommon occurrence. Without the incorporation of high-fidelity descriptions of pristine inner experience, we, as a field, are likely missing a vital piece of the puzzle.

As another example, theories of emotional development have recently been re-evaluated based, in part, upon considerations from DES data (Picker, 2017). A commonly accepted theory of emotion development is the Levels of Emotional Awareness model that posits feelings are first recognized in the self before being recognized in others and that certain phenomenological experiences are less advanced than others (e.g., bodily sensations are less sophisticated than blended emotions; Lane and Schwartz, 1987). However, this theory does not consistently

correlate with age, as expected, and does not build on other developmental models. A recently proposed perceptual differentiation of feeling model posits that feelings development parallels perceptual development, such that feelings are first recognized in others before they are recognized in the self and that emotional breadth, as opposed to specific phenomena, represents emotional sophistication (Picker, 2017). Besides the support found in correlational data with age, DES results also support this theory. Results from two young adolescents showed that even when there were emotional aspects of their experience (e.g., saying "I am sad" repeatedly), these adolescents rarely, if ever, had a direct experience of a feeling (Hurlburt, 2011). Adults, on the other hand, vary widely in how frequently they experience feelings. Perhaps adults with less emotional breadth have less pristine inner experience that includes feelings (Heavey et al., 2012). Continuing to incorporate careful observations of pristine inner experience will help develop a more complete understanding of emotional development and other psychological constructs.

Clinical Considerations

Psychotherapists generally operate on the assumption that clients have access to their pristine inner experience and can describe it. This study has shown that even *at the moment* reporting does not adequately capture a person's pristine inner experience without sufficient training. Regardless of whether they are queried on questionnaires or simply asked in the therapy room about their inner experience, clients will most likely only recognize and relate that which is in line with their presuppositions. Thus, if clients see themselves as depressed, they will more readily attend to and remember aspects of their experience that are in line with being depressed both throughout the day and when sitting in the therapy room. This suggests that psychotherapy clients may benefit from focused training in apprehending and describing their pristine inner experience. This notion is supported by the ongoing prevalence of mindfulness-based therapies.

In psychotherapy, psychology is strongly in favor of using questionnaires to help assess a client's primary symptoms and improvement overtime (American Psychological Association, 2006). This use of questionnaires is concerning, given the results of the current study. Treatment planning is often based on the specific symptoms a patient endorses on questionnaires spanning time frames much longer than *at the moment* or *today*. This is not to say that our diagnostic categories are meaningless, or that increases on psychotherapy outcome measures are unrelated to therapeutic improvements. However, when we focus on questionnaire endorsement, we do not know how these measures relate to clients' actual lived experience. Perhaps this disconnect is related to the substantial overlap between our diagnostic categories or our inability to identify the mechanism of change for many interventions.

Limitations and Future Directions

As with any study, it is important to recognize the limitations of this project. First, the Nevada Inner Experience Questionnaire (NIEQ) was the primary questionnaire used for comparison between groups. While it has shown acceptable psychometric properties, there are limitations, especially with regard to the sensory awareness scale. This could have impacted our sensory-awareness results. Similarly, the NIEQ-ATM and the NIEQ-T, developed for this study, have unknown psychometric properties, and may have impacted the results. Still, the NIEQ is approximately equivalent to other questionnaires in its psychometric adequacy. However, this study leads us to question the ability of the NIEQ and other questionnaires to measure pristine inner experience. This study also had unequal, relatively small sample sizes. However, given the *a priori* power analysis, and the large effect sizes observed, it is unlikely that the sample sizes significantly impacted the main conclusions of the study.

In the future, it would be interesting to investigate whether the intensive training in DES impacts participants' ability to report on their experience using short-timeframe questionnaires. Perhaps *in general* reports are resistant to change due to their explicit reliance on semantic memory and presuppositions, but the short-timeframe reports might be influenced by DES training. This could potentially allow for obtaining higher fidelity data with a less time-consuming method. A shorter period of DES training, followed by a longer period of short-timeframe questionnaire reporting would allow for larger sample sizes.

Conclusion

This study found that a method aimed at exploring inner experience in high fidelity (DES) found substantially lower frequencies of inner experience than would be predicted by questionnaire self-report, and substantially lower than would be indicated by other experience-sampling or diary methods. We conclude that questionnaire-based measures of inner experience produce results that include some indefinable mixture of pristine inner experience, memory errors, cognitive biases, miscommunications, beliefs about the self or the world, etc. Many of our psychological constructs and theories have been developed based on questionnaire data. A more complete psychological science requires the integration of descriptive methods that arguably avoid the hazards of studying pristine inner experience.

APPENDIX

Table 1

Descriptive Statistics for Outcome Measures by Group and Phase.

Measure	Phase	Group	Mean	St Dev	Min	Max
Inner Speaking[a]						
NIEQ	1 (Pre-test)	DES	64.17	19.08	45.00	100.00
NIEQ		ESM	66.26	20.81	24.00	100.00
NIEQ		DR	66.87	20.84	25.00	96.50
DES-based 5FP	2 (Experimental)	DES	12.66	13.99	0.00	43.75
NIEQ-ATM		ESM	52.38	19.87	20.56	95.21
NIEQ-T		DR	57.13	19.17	6.75	83.00
NIEQ	3 (Post-test)	DES	64.17	16.41	37.50	91.50
NIEQ		ESM	71.78	18.24	24.50	98.50
NIEQ		DR	65.16	19.04	35.50	93.00
Inner Seeing[a]						
NIEQ	1 (Pre-test)	DES	56.04	21.82	24.50	86.50
NIEQ		ESM	61.96	24.84	19.00	95.50
NIEQ		DR	56.11	27.04	11.50	95.50
DES-based 5FP	2 (Experimental)	DES	17.85	17.52	0.00	60.84
NIEQ-ATM		ESM	45.49	18.19	16.25	81.00
NIEQ-T		DR	35.74	17.17	8.00	65.75
NIEQ	3 (Post-test)	DES	63.63	19.25	35.00	94.50
NIEQ		ESM	61.28	25.75	12.00	97.00
NIEQ		DR	43.98	19.38	2.50	81.50
Feeling[a]						
NIEQ	1 (Pre-test)	DES	70.75	22.77	35.00	100.00
NIEQ		ESM	71.93	18.02	36.50	100.00
NIEQ		DR	77.61	20.50	32.00	100.00
DES-based 5FP	2 (Experimental)	DES	11.74	11.03	0.00	35.00
NIEQ-ATM		ESM	48.12	22.52	6.17	86.63
NIEQ-T		DR	55.43	21.64	13.75	100.00
NIEQ	3 (Post-test)	DES	59.00	20.53	30.00	85.50
NIEQ		ESM	68.63	22.85	16.50	96.50
NIEQ		DR	63.86	21.91	27.50	100.00
Sensory Awareness[a]						
NIEQ	1 (Pre-test)	DES	58.42	21.74	29.00	92.00
NIEQ		ESM	56.63	25.80	18.50	90.50
NIEQ		DR	51.52	17.84	21.50	91.00
DES-based 5FP	2 (Experimental)	DES	30.78	21.56	2.50	75.00
NIEQ-ATM		ESM	40.77	23.37	4.38	85.67
NIEQ-T		DR	28.78	18.72	0.00	58.50
NIEQ	3 (Post-test)	DES	57.67	24.48	18.00	100.00
NIEQ		ESM	54.17	25.84	6.00	99.00

NIEQ		DR	38.18	17.35	12.00	74.50
Unsymbolized Thinking[a]						
NIEQ	1 (Pre-test)	DES	28.89	26.50	1.00	100.00
NIEQ		ESM	34.57	25.81	3.00	82.00
NIEQ		DR	42.03	28.85	0.00	90.00
DES-based 5FP	2 (Experimental)	DES	10.07	10.21	0.00	39.58
NIEQ-ATM		ESM	32.55	16.35	1.25	56.08
NIEQ-T		DR	28.99	20.77	0.00	62.50
NIEQ	3 (Post-test)	DES	51.71	24.62	10.00	100.00
NIEQ		ESM	32.07	19.75	1.50	78.50
NIEQ		DR	33.95	24.61	0.00	78.50
Positive Affect[b]						
PANAS	1 (Pre-test)	DES	36.00	6.11	28.00	50.00
PANAS		ESM	34.87	6.61	22.00	50.00
PANAS		DR	31.05	7.91	14.00	46.00
PANAS	3 (Post-test)	DES	34.33	7.02	23.00	49.00
PANAS		ESM	35.13	7.05	21.00	50.00
PANAS		DR	27.45	7.55	12.00	42.00
Negative Affect[b]						
PANAS	1 (Pre-test)	DES	19.92	6.10	11.00	32.00
PANAS		ESM	22.26	6.59	14.00	36.00
PANAS		DR	23.05	8.54	10.00	40.00
PANAS	3 (Post-test)	DES	19.25	5.34	11.00	28.00
PANAS		ESM	22.96	5.66	13.00	33.00
PANAS		DR	20.14	7.16	11.00	42.00
STS[c]						
STS	1 (Pre-test)	DES	73.65	18.33	47.50	100.00
STS		ESM	78.86	12.03	55.00	100.00
STS		DR	74.60	14.03	48.75	100.00
STS	3 (Post-test)	DES	71.77	14.19	48.75	96.25
STS		ESM	79.35	11.26	57.50	100.00
STS		DR	68.18	13.45	36.25	88.75

[a] calculated mean based on cleaned data for all groups and all phases. Phases 1 and 3 used scaled scores based on NIEQ data. Potential range is 0-100.

[b] calculated mean of scale score. Potential range of scale score is 10-50. Higher scores indicate higher levels of positive or negative affect. Not measured in Phase 2.

[c] calculated according to the procedure utilized by Brinthaup, et al (2015), such that a total score was calculated for each participant, then divided by 80 (the maximum possible score on the measure) to provide a number indicating the percentage of time the participant reported engaging in self-talk. Potential range is 0-100. Not measured in Phase 2.

REFERENCES

American Psychological Association. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271–285. doi: 10.1037/0003-066X.61.4.271

Baer, R. A. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, *13*(1), 27–45. doi: 10.1177/1073191105283504

Boring, E. G. (1953). A history of introspection. *Psychological Bulletin*, *50*(3), 169–189. doi: 10.1037/h0090793

Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, *54*(1), 106–148. doi: 10.1111/j.1467-6494.1986.tb00391.x

Brinthaupt, T. M., Benson, S. A., Kang, M., & Moore, Z. D. (2015). Assessing the accuracy of self-reported self-talk. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.00570

Brinthaupt, T. M., Hein, M. B., & Kramer, T. E. (2009). The Self-Talk Scale: Development, factor analysis, and validation. *Journal of Personality Assessment*, *91*(1), 82–92. doi: 10.1080/00223890802484498

Brock, A. C. (2012). The history of introspection revisted. In J. W. Clegg (Ed.), *Self-observation in the social sciences* (pp. 25–43). New Brunswick, N.J: Transaction Publishers.

Clegg, J. W. (2012). Developing an adequate theory of self-observation. In J. W. Clegg (Ed.), *Self-observation in the social sciences* (pp. 3–21). New Brunswick, N.J: Transaction Publishers.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease*, *175*(9), 526–536.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: The MIT Press.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). GPower 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi: 10.3758/BF03193146

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics III: Speech acts* (pp. 41–58). New York, NY: Academic Press.

Halberstadt, J., & Wilson, T. D. (2008). Reflections on conscious reflection. In J. E. Adler & L. J. Rips (Eds.), *Reasoning: Studies of Human Inference and Its Foundations* (pp. 548–565). Cambridge: Cambridge University Press.

Heavey, C. L., & Hurlburt, R. T. (2008). The phenomena of inner experience. *Consciousness and Cognition*, *17*(3), 798–810. doi: 10.1016/j.concog.2007.12.006

Heavey, C. L., Hurlburt, R. T., & Lefforge, N. L. (2010). Descriptive experience sampling: A method for exploring momentary inner experience. *Qualitative Research in Psychology*, *7*(4), 345–368. doi: 10.1080/14780880903009274

Heavey, C. L., Hurlburt, R. T., & Lefforge, N. L. (2012). Toward a phenomenology of feelings. *Emotion*, *12*(4), 763–777. doi: 10.1037/a0026905

Heavey, C. L., Lefforge, N. L., Lapping-Carr, L., & Hurlburt, R. T. (2017). Mixed emotions: Toward a phenomenology of blended and multiple feelings. *Emotion Review*, *9*(2), 105–110. doi: 10.1177/1754073916639661

Heavey, C. L., Moynihan, S. A., Brouwers, V., Lapping-Carr, L., Krumm, A. E., Kelsey, J. M., … Hurlburt, R. T. (2019). Measuring the frequency of inner-experience characteristics by

self-report: The Nevada Inner Experience Questionnaire. *Frontiers in Psychology, 9*. doi: 10.3389/fpsyg.2018.02615

Hurlburt, R. T. (1990). *Sampling normal and schizophrenic inner experience*. New York, NY: Plenum Press.

Hurlburt, R. T. (1993). *Sampling inner experience in disturbed affect*. New York: Plenum Press.

Hurlburt, R. T. (2009). Unsymbolized thinking, sensory awareness, and mindreading. *Behavioral and Brain Sciences*, *32*(02), 149. https://doi.org/10.1017/S0140525X09000673

Hurlburt, R. T. (2011). *Investigating pristine inner experience: Moments of truth*. New York: Cambridge University Press.

Hurlburt, R. T., & Akhter, S. A. (2008). Unsymbolized thinking. *Consciousness and Cognition*, *17*(4), 1364–1374. https://doi.org/10.1016/j.concog.2008.03.021

Hurlburt, R. T., Alderson-Day, B., Fernyhough, C., & Kühn, S. (2017). Can Inner Experience Be Apprehended in High Fidelity? Examining Brain Activation and Experience from Multiple Perspectives. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00043

Hurlburt, R. T., Alderson-Day, B., Kühn, S., & Fernyhough, C. (2016). Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech. *PLOS ONE*, *11*(2), e0147932. https://doi.org/10.1371/journal.pone.0147932

Hurlburt, R. T., & Heavey, C. L. (2001). Telling what we know: describing inner experience. *Trends in Cognitive Sciences*, *5*(9), 400–403. https://doi.org/10.1016/S1364-6613(00)01724-1

Hurlburt, R. T., & Heavey, C. L. (2002). Interobserver reliability of Descriptive Experience Sampling. *Cognitive Therapy and Research*, *26*(1), 135–142.

Hurlburt, R. T., & Heavey, C. L. (2006). *Exploring inner experience: the descriptive experience sampling method*. Philadelphia: John Benjamins Pub.

Hurlburt, R. T., & Heavey, C. L. (2015). Investigating pristine inner experience: Implications for experience sampling and questionnaires. *Consciousness and Cognition*, *31*, 148–159. doi: 10.1016/j.concog.2014.11.002

Hurlburt, R. T., Heavey, C. L., & Bensaheb, A. (2009). Sensory awareness. *Journal of Consciousness Studies*, *16*(10–12), 231–251.

Hurlburt, R. T., Heavey, C. L., & Kelsey, J. M. (2013). Toward a phenomenology of inner speaking. *Consciousness and Cognition*, *22*(4), 1477–1494. doi: 10.1016/j.concog.2013.10.003

Hurlburt, R. T., & Schwitzgebel, E. (2007). *Describing inner experience?: Proponent meets skeptic*. Cambridge, Mass: MIT Press.

Hurlburt, R. T., & Schwitzgebel, E. (2011). Methodological pluralism, armchair introspection, and DES as the epistemic tribunal. *Journal of Consciousness Studies*, *18*(1), 253–273.

Kihlstrom, J. F., Eich, E., Sandbrand, D., & Tobias, B. A. (2000). Emotion and memory: Implications for self-report. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report* (pp. 81–99). Mahwah, NJ: Lawrence Erlbaum Associates.

Klinger, E., & Cox, W. M. (1987). Dimensions of thought flow in everyday life. *Imagination, Cognition and Personality*, *7*(2), 105–128. doi: 10.2190/7K24-G343-MTQW-115V

Kühn, S., Fernyhough, C., Alderson-Day, B., & Hurlburt, R. T. (2014). Inner experience in the scanner: Can high fidelity apprehensions of inner experience be integrated with fMRI? *Frontiers in Psychology*, *5*. doi: 10.3389/fpsyg.2014.01393

LaFrance, M., & Banaji, M. (1992). Toward a reconsideration of the gender–emotion relationship. In M. S. Clark (Ed.), *Emotion and social behavior* (pp. 178–201). Newbury Park, CA: Sage.

Lane, R. D., & Schwartz, G. E. (1987). Levels of emotional awareness: A cognitive-developmental theory and its application to psychopathology. *American Journal of Psychiatry*, *144*(2), 133–143. doi: 10.1176/ajp.144.2.133

Lenzner, T., & Menold, N. (2016). Question wording. *GESIS Survey Guidelines*. doi: 10.15465/gesis-sg_en_017

Locke, E. A. (2009). It's time we brought introspection out of the closet. *Perspectives on Psychological Science*, *4*(1), 24–25. doi: 10.1111/j.1745-6924.2009.01090.x

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General*, *145*(2), 200–219. doi: 10.1037/a0039923

Monson, C. K., & Hurlburt, R. T. (1993). A comment to suspend the introspection controversy: Introspecting subjects did agree about "Imageless Thought." In *Sampling inner experience in disturbed affect* (pp. 15–26). New York: Plenum Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. doi: 10.1037/0033-295X.84.3.231

Picker, C. J. (2017). *Perceptual differentiation of feelings model: Development of individual differences in emotion differentiation* (Qualifying Paper). University of Nevada, Las Vegas, Las Vegas, NV.

Robinson, M. D., & Clore, G. L. (2002a). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. doi: 10.1037/0033-2909.128.6.934

Robinson, M. D., & Clore, G. L. (2002b). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, *83*(1), 198–215. doi: 10.1037//0022-3514.83.1.198

Rush, J., & Hofer, S. M. (2014). Differences in within- and between-person factor structure of positive and negative affect: Analysis of two intensive measurement studies using multilevel structural equation modeling. *Psychological Assessment*, *26*(2), 462–473. doi: 10.1037/a0035666

Saunders, D. M., Fisher, W. A., Hewitt, E. C., & Clayton, J. P. (1985). A method for empirically assessing volunteer selection effects: Recruitment procedures and responses to erotica. *Journal of Personality and Social Psychology*, *49*(6), 1703–1712. doi: 10.1037/0022-3514.49.6.1703

Seager, W. (2002). Emotional introspection. *Consciousness and Cognition*, *11*(4), 666–687. doi: 10.1016/S1053-8100(02)00027-2

Siewert, C. (2011). Socratic introspection and the abundance of experience. *Journal of Consciousness Studies*, *18*(1), 63–91.

Skinner, B. F. (1974). *About behaviorism*. New York: Vintage.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. doi: 10.1037/pspp0000096

Strassberg, D. S., & Lowe, K. (1995). Volunteer bias in sexuality research. *Archives of Sexual Behavior*, *24*(4), 369–382. doi: 10.1007/BF01541853

Tellegen, A., Watson, D., & Clark, L. A. (1999). On the dimensional and hierarchical structure of affect. *Psychological Science*, *10*(4), 297–303. doi: 10.1111/1467-9280.00157

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report* (pp. 29–47). Mahwah, NJ: Lawrence Erlbaum Associates.

Trivedi, N., & Sabini, J. (1998). Volunteer bias, sexuality, and personality. *Archives of Sexual Behavior*, *27*(2), 181–195.

Turner, D. T., II. (2015). *The experience of participating in DES* (Unpublished doctoral dissertation). University of Nevada, Las Vegas, Las Vegas, NV.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. doi: 10.1037/0022-3514.54.6.1063

Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1990). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. *Advances in Experimental Social Psychology*, *22*, 287–343.

Wolchik, S. A., Braver, S. L., & Jensen, K. (1985). Volunteer bias in erotica research: Effects of intrusiveness of measure and sexual background. *Archives of Sexual Behavior*, *14*(2), 93–107. doi: 10.1007/BF01541656

CURRICULUM VITAE

Leiszle R. Lapping-Carr

(née Leiszle Rae Ziemba)
E-mail: leiszle.lappingcarr@gmail.com

EDUCATION

Doctor of Philosophy: Clinical Psychology **In Progress**
University of Nevada, Las Vegas (APA Accredited)
Dissertation: *Multimethod Investigation of Pristine Inner Experience*
GPA: 4.0
Advisor: Christopher L. Heavey, Ph.D.

Master of Arts: Clinical Psychology **2016**
University of Nevada, Las Vegas (APA Accredited)
Thesis: *Inner Experience while Reading an Erotic Short Story*
GPA: 4.0
Advisor: Christopher L. Heavey, Ph.D.

Bachelor of Arts: Psychology, Religious Studies **2009**
Brown University
Honors Thesis: *Investigation of the Relationship between Spirituality and Sex Behaviors in College Students*
GPA: 3.7
Advisor: John Wincze, Ph.D.

SCHOLARSHIPS, GRANTS, AND AWARDS

2018 to 2019
UNLV Graduate College Summer Doctoral Research Fellowship $7,000

2017 to 2018 **$29,000**
President's UNLV Foundation Graduate Research Fellowship $25,000
UNLV Graduate College Summer Session Grant $2,000
UNLV Graduate Access Grant $2,000

2016 to 2017 **$1,267.50**
UNLV GPSA Service Award $300
UNLV GPSA Annual Research Forum, 2nd Place $125
UNLV GPSA Travel Grant $700
Society for the Scientific Study of Sexuality Student Ambassador Award $142.50

2015 to 2016 **$2,515**
UNLV Graduate College Summer Session Grant $2,000
UNLV GPSA Travel Grant $515

2014 to 2015 **$750**
UNLV GPSA Travel Grant $750

2013 to 2014 **$500**
UNLV GPSA Travel Grant $500

RESEARCH EXPERIENCE

Graduate Research Assistant 2013 - 2018
Descriptive Experience Sampling Lab
University of Nevada, Las Vegas
Supervisor: Christopher L. Heavey, Ph.D.
Research projects in this lab center around exploring pristine inner experience in a variety of populations and tasks. I have been actively involved in projects focused on inner speaking, inner experience while reading fiction, inner experience while reading erotica, and inner experience of individuals diagnosed with depression. This lab also examines and evaluates psychological research methods that purport to measure pristine inner experience.

Specific Responsibilities: Develop research protocols, administer and consent participants, administer questionnaires, conduct qualitative interviews, analyze data, and co-authoring manuscripts. Also, recruit, train, and over-see undergraduate research assistants in participant recruitment, consent, and study protocols.

Graduate Research Assistant 2017-2018
Interactive Measurement Group
University of Nevada, Las Vegas
Supervisor: Kimberly A. Barchard, Ph.D.
Research projects in this lab focus on the development and validation of questionnaires. Currently, we are developing a measure to allow humans to assess the social appropriateness of robotic movement, a collaborative project with the University of Nevada, Reno Department of Computer Science and funded by NASA.

Specific Responsibilities: Conduct literature review, assist in developing constructs and writing items, assist in implementing validation study using online community members, data analysis, and co-authoring manuscripts.

Graduate Research Assistant 2014 - 2015
Psychophysiology of Emotion and Personality Lab
University of Nevada, Las Vegas
Supervisors: Taylor Oliver, Ph.D. & Stephen Benning, Ph.D.
Research projects in this lab focused on the role of personality in emotional reactivity, including psychophysiological measures such as electroencephalography (EEG), skin conductance, and heart rate. Studies examined how individuals with different clinical presentations (e.g., depression) reacted differently to emotional stimuli as a function of their personality. Also, studies examined the psychophysiological response to sexual stimuli.

Specific Responsibilities: consenting participants into the study, setting up the EEG equipment, attaching electrodes and properly applying EEG cap to participants, checking for adequate signal, running participants through trials, train and oversee undergraduate research assistants in the above tasks, co-authoring manuscripts, assist with literature reviews, assist with data analysis, and facilitating collaboration between the two PIs.

Research Assistant, Grant Writing Support 2013
Department of Psychiatry and Behavioral Sciences
University of Washington
Supervisor: Lydia Chwastiak, MD, MPH
Projects under this supervisor were investigating improvements to diabetes treatment by using a multidisciplinary team based at a community mental health center. This team approach addresses key patient-, provider- and systems-level factors which contribute to poor quality of care for a clinically complex population.

Specific Responsibilities: Review and edit grant proposals, assist in developing study protocols and literature review, and develop on-line and paper questionnaire packets.

Research Assistant 2011 – 2013
Department of Psychiatry and Behavioral Sciences
University of Washington
Supervisor: Douglas Zatzick, MD
Projects under this supervisor focused on implementation and dissemination studies of alcohol screening and brief intervention, interventions for PTSD and suicide in medical centers. The intervention contains an organizational development component to assist trauma centers in building capacity for mental health service delivery.

Specific Responsibilities: Identify potential participants using hospital electronic medical record, develop automated list production to assist in participant identification. Recruit and conduct semi-structured interviews with patients admitted to the hospital for a trauma related injury or suicide attempt. Conduct follow-up assessments with participants. Test website for online administration of questionnaires during development. Co-author manuscripts.

Research Assistant 2010 - 2012
Department of Psychiatry and Behavioral Sciences
University of Washington
Supervisor: Wayne Katon, MD
Projects under this supervisor consisted of an implementation and dissemination study of a depression care management program for women attending OB-GYN clinics in the University of Washington health care system. The depression intervention consisting of enhanced education, engagement, and depression care management (with a choice of antidepressant medication monitoring and/or provision of brief psychotherapy).

Specific Responsibilities: Identify potential participants using hospital electronic medical record, recruit participants and conduct structured interviews to patients who screen positive for depression in an OB-GYN clinic. Conduct follow-up assessments over the phone.

Undergraduate Research Assistant 2008 - 2009
Department of Psychiatry and Human Behavior
Brown University
Supervisor: Willoughby Britton, PhD
Projects under this supervisor focused on the impact of meditation practices on attention, emotional well-being, and school likeability in 6th graders as well as attention, emotion-regulation, sleep, and sexual function in undergraduates.

Specific Responsibilities: Administer neuropsychological assessments (Conner Continuous Performance Test and Trail Making A and B) to participants, data entry, and data analysis.

PEER REVIEWED PUBLICATIONS

6. Heavey, C. L., Moynihan, S. A., Brouwers, V. P., **Lapping-Carr, L.**, Krumm, A. E., Kelsey, J. M., Turner, D. K., and Hurlburt, R. (2019). Measuring the frequency of inner-experience characteristics by self-report: The Nevada Inner Experience Questionnaire. *Frontiers in Psychology, 9*. doi: 10.3389/fpsyg.2018.02615
5. Brouwers, V. P., Heavey, C. L., **Lapping-Carr, L**., Moynihan, S., Kelsey, J. & Hurlburt, R. T. (2018) Pristine inner experience while silent reading: It's *not* silent speaking of the text. *Journal of Consciousness Science*, *25*(3-4), 29-54.
4. **Lapping-Carr, L.** & Heavey, C. L. (2017) Pristine inner experience and descriptive experience sampling: Implications for psychology. *Frontiers in Psychology, 8*. doi: 10.3389/fpsyg.2017.02170
3. Heavey, C. L., Lefforge, N. L., **Lapping-Carr, L**., & Hurlburt, R. T. (2017) Mixed emotions: Toward a phenomenology of blended and multiple feelings. *Emotion Review, 9,* 105-110. doi: 10.1177/1754073916639661
2. O'Connor, S. S., Comotois, K. A., Wang, J., Russo, J., Peterson, R., **Lapping-Carr, L**.,& Zatzick, D. (2015) The development and implementation of a brief intervention for medically admitted suicide attempt survivors. *General Hospital Psychiatry, 37*(5), 427-433. doi:10.1016/j.genhosppsych.2015.05.001
1. O'Connor, S. S., Dinsio, K., Wang, J., Russo, J., Rivara, F. P., Love, J., McFadden, C., **Lapping-Carr, L.**, Peterson, R., & Zatzick, D. F. (2014) Correlates of suicidal ideation in physically injured trauma survivors. *Suicide and Life-Threatening Behavior, 44*(5) 473-485. doi: 10.1111/sltb.12085

MANUSCRIPTS UNDER REVIEW

2. **Lapping-Carr, L.** & Heavey, C. L. Introspection as a tool in emotion research. (Manuscript submitted to *Emotion Review*).
1. **Lapping-Carr, L.**, Heavey, C. L., Brouwers, V. P., & Hurlburt, R. T. Pristine inner experience while reading erotica. (Manuscript submitted to Special Issue: Sexuality and Individual Differences, in *Personality and Individual Differences*).

MANUSCRIPTS IN PREPARATION

2. Barchard, K. A., **Lapping-Carr, L.**, Westfall, R. S., Banisetty, S. B., & Feil-Seifer, D. Measuring the perceived social intelligence of robots.
1. **Lapping-Carr, L.**, Krumm, A., Kelsey, J., Brouwers, V., Moynihan, S., Turner, D., Heavey, C. L., and Hurlburt, R. T. The frequency of self-talk as measured by questionnaire and descriptive experience sampling.

PEER REVIEWED TEACHING RESOURCES

11. Barchard, K. A., & **Lapping-Carr, L**. (2016). Introduction to SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience [a peer-reviewed website of resources for teaching research and statistics in psychology, supported by a grant from the Association for Psychological Science]. Retrieved from TeachPsychScience.org
10. Barchard, K. A., & **Lapping-Carr, L**. (2016). Graphs for SPSS: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
9. Barchard, K. A., & **Lapping-Carr, L**. (2016). The Syntax Window for SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org

8. Barchard, K. A., & **Lapping-Carr, L**. (2016). Calculating Statistics for Subgroups for SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
7. Barchard, K. A., & **Lapping-Carr, L**. (2016). Recoding for SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
6. Barchard, K. A., & **Lapping-Carr, L**. (2016). Inferences about the Mean of a Single Sample in SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
5. Barchard, K. A., & **Lapping-Carr, L**. (2016). Comparing Independent Samples in SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
4. Barchard, K. A., & **Lapping-Carr, L**. (2016). Comparing Dependent Samples in SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
3. Barchard, K. A., & **Lapping-Carr, L**. (2016). Correlation in SPSS 22.0: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
2. Barchard, K. A., & **Lapping-Carr, L**. (2016). Creating a Conference Poster on Sex Differences using SPSS 22.0 and PowerPoint: Assignment and Grading Rubric. TeachPsychScience. Retrieved from TeachPsychScience.org
1. Barchard, K. A., & **Lapping-Carr, L.** (2016). Power for the Single-Sample t-test: Assignment and Grading Rubric. *TeachPsychScience*. Retrieved from TeachPsychScience.org

PSYCHOLOGICAL TESTS

2. Barchard**,** K. A., **Lapping-Carr, L.**, Westfall, R. S., Banisetty, S. B., & Feil-Seifer, D. (2018). *Perceived Social Intelligence (PSI) Scales test manual.* Unpublished psychological test and test manual. Observer report of 20 aspects of social intelligence of robots, with four items per scale. Available form KimBarchard, kim.barchard@unlv.edu and David Feil-Seifer, dave@cse.unr.edu
1. Heavey, C. L., Moynihan, S. A., Brouwers, V. P., **Lapping-Carr, L.**, Krumm, A. E., Kelsey, J. M., Turner, D. K., and Hurlburt, R. (2019). Measuring the frequency of inner-experience characteristics by self-report: The Nevada Inner Experience Questionnaire. *Frontiers in Psychology, 9*. doi: 10.3389/fpsyg.2018.02615

CONFERENCE PRESENTATIONS

17. Barchard, K. A., **Lapping-Carr, L.**, Westfall, R. S., & Feil-Seifer, D. (2019, February). *Perceived social intelligence of robots.* Poster to be presented at the Society for Personality and Social Psychology, Portland, Oregon.
16. Barchard, K. A., **Lapping-Carr, L.**, Westfall, R. S., Banisetty, S. B., & Feil-Seifer, D. (2018, October). *Measuring 20 aspects of the perceived social intelligence of robots.* Poster presented at the International Conference on Intelligent Robots, Madrid, Spain.
15. Krumm, A. E., **Lapping-Carr, L.**, Kaneshiro, C., Moynihan, S., Hurlburt, R. T., & Heavey, C. L. (2018, May). *Subjective Experience is Not All the Same: Private Phenomena vs. Inferred States.* Poster presented at Association for Psychological Science, San Francisco, CA.
14. Kaneshiro, C., Moynihan, S., Krumm, A. E., **Lapping-Carr, L.**, Heavey, C. L., & Hurlburt, R. T. (2018, May). *Can First-Person Methods Reliably Apprehend Inner Experience? Lessons from Eyewitness Testimony.* Poster presented at Association for Psychological Science, San Francisco, CA.
13. Heavey, C. L., **Lapping-Carr, L.** (2017, April). Exploring naturally occurring feelings with descriptive experience sampling. In in K. A. Barchard (Chair), *Innovative Methods in Emotion Research.* Symposium conducted at the meeting of the Western Psychological Association, Sacramento, California

12. **Lapping-Carr, L.** & Heavey, C. L. (2016, November). *Inner experience while reading an erotic short story*. Poster presented at the 2016 Annual Meeting of Annual Meeting of The Society for the Scientific Study for Sexuality, Pheonix, AZ.
11. **Lapping-Carr, L.**, Barchard, K. A., Bartlett, A. B., Quiroz, J., & Martinex, N. (2016, February). *Teaching undergraduates to use SPSS: Assignments and rubrics*. Poster presented at the 19th Annual American Association of Behavioral and Social Sciences Conference, Las Vegas, NV.
10. **Lapping-Carr, L.**, Oliver, T. L., Duckro, A., & Benning, S. D. (2015, September). *Central and reflexive measures of reactivity to human and primate erotica.* Poster presented at the 55th Annual Meeting of the Society for Psychophysiological Research, Seattle, WA.
9. Heavey, C. L., Lefforge, N. L., **Lapping-Carr, L.**, & Gunter, J. D. (2015, May). *Dysphoria is subtly present in the momentary experience of depressed individuals.* Poster presented at the Association for Psychological Science 27th Annual Convention, New York, NY.
8. **Lapping-Carr, L.**, Brouwers, V., Hurlburt, R. T., & Heavey, C. L. (2015, April). Erotica and inner experience. In C. L. Heavey (Chair), *The inner experience of inner speaking, reading fiction, and reading erotica: What's really going on?* Symposium conducted at the Western Psychological Association 95th Annual Convention, Las Vegas, NV.
7. Moynihan, S. A., Turner, D., Brouwers, V., Kelsey, J., **Lapping-Carr, L.**, Heavey, C. L., & Hurlburt, R. T. (2015, April). Validating the Nevada Inner Experience Questionnaire through an inner speaking perspective. In C. L. Heavey (Chair), *The inner experience of inner speaking, reading fiction, and reading erotica: What's really going on?* Symposium conducted at the Western Psychological Association 95th Annual Convention, Las Vegas, NV.
6. Brouwers, V., Kelsey, J., Turner, D., **Lapping-Carr, L.**, Moynihan, S. A., Heavey, C. L., & Hurlburt, R. T. (2015, April). Inner experience while reading classical fiction. In C. L. Heavey (Chair), *The inner experience of inner speaking, reading fiction, and reading erotica: What's really going on?* Symposium conducted at the Western Psychological Association 95th Annual Convention, Las Vegas, NV.
5. Kelsey, J., Brouwers, V., Turner, D., **Lapping-Carr, L.**, Moynihan, S. A., Hurlburt, R. T., & Heavey, C. L. (2015, April). Inner experience and self ratings of self-talk. In C. L. Heavey (Chair), *The inner experience of inner speaking, reading fiction, and reading erotica: What's really going on?* Symposium conducted at the Western Psychological Association 95th Annual Convention, Las Vegas, NV.
4. Lefforge, N., **Lapping-Carr, L**. (2014, April). *Beep here now: Descriptive experience sampling provides a structured path towards mindfulness.* Paper presented at the Toward a Science of Consciousness conference, Tucson, AZ.
3. O'Connor, S., Zatzick, D., Love, J., McFadden, C., **Ziemba, L.**, Peterson, R., Ali, N. (2013, April). *Pilot study of a brief intervention delivered to patients admitted to an acute care medical/surgical hospital setting following a suicide attempt.* Paper presented at the American Association of Suicidology 46th Annual Conference, Austin, TX.
2. Love, J., **Ziemba, L.**, Wang, J., Russo, J., Zatzick, D. (2012, November). *Using nationwide U.S. surveys to evaluate policy guided implementation strategies and PTSD in acute care medical settings*. Poster presented at the International Society for Traumatic Stress Studies 28th Annual Meeting, Los Angeles, CA.
1. Silverstein, R. G., **Ziemba, L.**, Devlin, K., Anderson, C., Rothamel, K., Coleman, C., Lepp, N., Britton, W. (2009, March). *Get out of your mind and into your body: The role of mindfulness in the treatment of female sexual dysfunction.* Poster presented at Brown University's Alpert Medical School Department of Psychiatry and Human Behavior Research Day, Providence, RI.

SUPERVISED CLINICAL EXPERIENCE

Clinical Psychology Intern 2018-present
Northwestern Medicine (Chicago, IL)

Supervisors: Mark Reinecke, Ph.D, Richard Carroll, Ph.D., John Stutesman, Psy.D., Ellen Astrachan-Fletcher, Ph.D.
Description: Combination of individual, group, couples, and family therapy at multiple locations, including an outpatient community mental health clinic and an intensive outpatient mood, anxiety, and trauma program. One full day of didactic training a week with four days spent seeing patients and doing administrative work. Supervision in CBT, DBT, psychodynamic, systems orientations and sex therapy.

Doctoral Practicum Clinician 2016-2018
Family and Child Treatment (FACT; Las Vegas, NV)
Supervisor: John Matthias, Ph.D.
Description: FACT is a non-profit community mental health clinic providing individual and group services for child and adult victims of sexual abuse, sex trafficking, and domestic violence as well as psychosexual risk evaluations, individual, and group services to sex offenders.

Doctoral Practicum Student 2017-present
Nevada Relationship Center
Supervisor: Scott Wolfe, Ph.D.
Description: The Nevada Relationship Center provide couples therapy and psycho-education workshops to Las Vegas couples presenting with relational distress using Gottman Couples Therapy principles. Conduct Gottman Couples Therapy training seminars to clinicians.

Supervisor in Training Summer 2016, 2017
Doctoral Practicum Student 2014-2015, Summer 2016, Fall 2017
The PRACTICE: A UNLV Community Mental Health Clinic
Supervisors: Michelle G. Paul, Ph.D.; Noelle Lefforge, Ph.D.
Description: The PRACTICE is an interdisciplinary, university-housed training clinic providing evidence-based sliding-scale individual, family, and group therapy and psychodiagnostic assessment to children and adults in the Las Vegas metropolitan area and tele-mental health services to rural Nevada. Received training in individual, group, and couples counseling, tele-mental health services, and psychodiagnostic assessment as well as providing supervised peer-supervision.

Doctoral Practicum Student
UNLV Counseling and Psychological Services (CAPS) Fall 2015
Supervisors: Ann Allen, Ed.D.; Ashley Bock, Psy.D.; Shauna Landis, Ph.D.
Description: CAPS is a university counseling center, offering intakes to determine eligibility for services (and to make appropriate referrals), up to 12 free individual therapy sessions, group therapy, crisis intervention, and psychiatric medication management to university students.

TEACHING & MENTORSHIP

Graduate Student Mentor 2014-2018
University of Nevada, Las Vegas
Outreach Undergraduate Mentoring Program
Activities: Mentor undergraduate psychology students from under-represented backgrounds, such as ethnic minorities, LGBTQ students, students with disabilities; OUMP is designed to help

students adequately prepare and apply to psychology graduate programs. Connect students with appropriate faculty mentors, provide recommendations for research laboratories, assist in CV development, and guide career planning.

Graduate Student Instructor
University of Nevada, Las Vegas
Psychology 210: Introduction to Statistical Methods (2 sections) Spring 2016
Psychology 101: General Psychology (2 sections) Fall 2016

Graduate Teaching Assistant 2014 - 2016
University of Nevada, Las Vegas
Psychology 210: Introduction to Statistical Methods
Psychology 740: Psychometrics
Supervisor: Kimberly Barchard, Ph.D.
Activities: Guest lecture, manage online course management system, edit SPSS assignments, and create, use, and edit grading rubrics for weekly SPSS assignments.

Workshop Leader Spring 2016
University of Nevada, Las Vegas
Exploratory Factor Analysis (EFA)
Supervisor: Kimberly Barchard, Ph.D.
Activities: Adapt graduate lecture transcripts and SPSS lab assignments on EFA for undergraduate use. Provide instruction on the basics of EFA.

Graduate Teaching Assistant 2013 - 2014
University of Nevada, Las Vegas
Psychology 360: Foundations of Social Psychology
Psychology 412: Motivation and Emotion
Psychology 435: Personality
Supervisor: Jane Powell, Ph.D.
Activities: Proctor and grade exams

Undergraduate Teaching Assistant Spring 2009
Brown University
Behavior Modification
Supervisor: John Wincze, Ph.D.
Activities: Provide exam review, grade exams, advise students on term papers

Sex Education Classroom Coordinator Spring 2009
Brown University; The Metropolitan Regional Career and Technical Center High School
Brown's Swearer Center for Public Service
Peer HIV/AIDS and Sex Education (PHASE)
Supervisor: Cassidy Boesch, Gina Silverstein (student-led organization)
Activities: Provide near-peer sex education and facilitate discussions with approximately 10 high school students (ages 15-18). Customize PHASE curriculum to meet needs of students in the classroom.

Meiklejohn Peer Advisor 2006 - 2007
Brown University
Supervisor: Elizabeth Silbermann (student-led organization)

Activities: Provide mentorship to incoming students at Brown University. Resource for course selection, administrative and academic requirements, student organizations, and academic career planning.

LEADERSHIP & SERVICE

American Psychological Association for Graduate Students
State Advocacy Coordinator 2016 - 2018

Description: APAGS aims to engage students from psychology graduate schools in advocacy activities including social justice, graduate student rights, and professional development. APAGS also coordinates between the graduate student body, APA, and state psychological associations.

Activities: Specific responsibilities include recruiting graduate students statewide into advocacy positions, providing resources on how to engage students in advocacy activities, and relaying thoughts and concerns of students to APA and state psychological associations. During my time, also created and implemented a series of didactic trainings of interest to students to supplement graduate training (e.g., psychopharmacology), created and implemented social justice forums for students and faculty across health professions to discuss current events and the role of psychologists/mental health providers in addressing these issues, and created and implemented a state-wide mental health awareness event providing the undergraduate student body with resources about mental health treatment on campus.

Clinical Psychology Student Committee
Co-Chair 2017 - 2018
Cohort Representative 2016 - 2017

Description: The Clinical Student Committee is a student organization dedicated for advocacy and support for clinical psychology doctoral students.

Activities: Attended monthly committee meetings; attended faculty meetings as a non-voting member; acted as liaison between faculty and students; managed committee funds; organized social events and faculty appreciation fundraiser; coordinated a pre-interview event and graduate student housing for program applicants.

Journal of Sex Research
Reviewer 2017 - present

Society for the Scientific Study of Sexuality 2016 - 2017
Student Ambassador

Association for Psychological Science Student Caucus 2015 - 2017
Student Grant Competition Reviewer

Nevada Psychological Association (NPA) 2014 - 2016
Campus Representative

Description: Serve as a liaison between graduate students at UNLV and the Nevada Psychological Association. Plan and host social events, including organizing a raffle to fundraise for NPA's PAC.

Sex.Religion@Brown 2009
Planning Committee

Description: Develop talking points and serve on the panel for a discussion on the impact of religion on Brown students' sexuality.

SPECIALIZED TRAINING

Integrating Behavioral Health into Primary Care (8 hours) — May 2017
Nevada Psychological Association
Jeffrey Reiter, Ph.D., ABPP

Workshop on Linear Regression using R (18 hours) — May 2017
University of Nevada, Las Vegas
Andrew Freeman, Ph.D.

Sex Therapy Skills for the Practicing Clinician (8 hours) — April 2017
Society for Sex Therapy and Research
Instruction: Kathryn Hall, Ph.D., Daniel Watter, Ph.D., Aline Zoldbrod, Ph.D., Dennis Kalogeropoulos, Ph.D., Sophie Bergeron, Ph.D., Natalie Rosen, Ph.D., Marta Meana, Ph.D., Michael Berry, Ph.D., Barry McCarthy, Ph.D.

Everything You Need to Know About Internet Sex Offenders (8 hours) — October 2016
UNLV Boyd School of Law
Suzonne Kline, Ph.D.; Eric Imhof, Ph.D.

The APA Saga of Torture Interrogation: Lessons for Psychological Ethics in Institutional Settings (8 hours) — September 2016
Nevada Psychological Association
Jean Maria Arrigo, Ph.D.**;** David DeBatto

Gottman Method Couples Therapy Training
Level 1 (12 hours) — September 2016
Level 2 (24 hours) — October 2017
Nevada Relationship Center
Instruction: Scott Wolfe, Ph.D, Certified Gottman Therapist and Trainer

Introduction to Couples Therapy: A Gottman Research Based Approach (6 hours) — October 2015
Nevada Psychological Association
Scott Wolfe, Ph.D, Certified Gottman Therapist and Trainer

Comprehensive Training in Dialectical Behavior Therapy
Part I: Theory, Structure, Targets and Treatment Strategies (24 hours) — February 2015
Part II: DBT Skills, Skill Training, and Coaching (24 hours) — April 2015
Nevada Psychological Association
Instruction: Alan Fruzzetti, Ph.D.

Hot Topics in Ethics and Risk Management in Psychological Practice (8 hours) — September 2014
Nevada Psychological Association
Instruction: Eric Harris, J.D., Ed.D

Motivational Interviewing and Engagement (8 hours) — January 2010
University of Washington
Instruction: Nancy Grote, PhD, MSW

Problem Solving Therapy (12 hours) — January 2010
University of Washington
Instruction: Evette Ludman, PhD; Anna LaRocco-Cockburn, MSW

Mindfulness Based Cognitive Therapy (60 hours) — Summer 2008
Brown University
Instruction/Supervision: Willoughby Britton, Ph.D.

PROFESSIONAL AFFILIATIONS

2018-present: Member, Illinois Psychological Association
2017-present: Student/ECP Committee Member, Society for Sex Therapy and Research
2016-present: Member, Society for the Scientific Study of Sexuality
2016-present: Member, Society for the Teaching of Psychology (APA Division 2)
2014-present: Member, American Psychological Association
2014-pesent: Member, American Association of Sexuality Educators Counselors and Therapists
2013-present: Member, Association for Psychological Science
2014-2018: Member, Nevada Psychological Association
2014-2018: Member, Western Psychological Association
2015-2016: Member, American Association of Behavioral and Social Sciences
2015-2016: Member, Society for Psychophysiological Research
2008-2009: Member, Contemplative Studies Initiative, Brown University

REFERENCES

Available upon request