

PARTICIPANT TRAINING AND ITS EFFECT ON ACTUAL RETROSPECTIVE
TIMEFRAMES

Shannah R. Sphar

A Thesis Submitted to the
University of North Carolina at Wilmington in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts

Department of Psychology
University of North Carolina at Wilmington

2008

Approved by

Advisory Committee

Dr. Richard Ogle

Dr. James Johnson

Dr. Len Lecci
Chair

Accepted by

Dean, Graduate School

TABLE OF CONTENTS

ABSTRACT	iv
ACKNOWLEDGEMENTS	v
DEDICATION	vi
LIST OF TABLES	vii
LIST OF FIGURES	viii
INTRODUCTION	1
Conceptualizing different timeframes.....	2
The use of timeframes in the literature	4
Participant training to improve convergence with expected timeframes	8
Other research confounds that could be attributed to variability in mood timeframes.....	11
The current study	15
METHOD	20
Participants.....	20
Materials	21
Procedure	23
RESULTS	24
Examining the effects of participant training.....	24
The effects of participant training on variability	26
Examining the association between NA and PA	27
Examining the effects of different affect terms on the correlation between PA and NA ..	30
DISCUSSION	31
Participant training’s effect on actual retrospective timeframes.....	31

Effects of participant training on variability	33
Effects of participant training on the dimensionality of affect	35
Decision to use the Mann-Whitney Rank-sum test.....	37
Limitations and future research	38
LITERATURE CITED	40
APPENDIX.....	43

ABSTRACT

When rating moods (e.g., How do you feel “at this moment”), individuals employ lengthy timeframes that do not converge with the expected timeframe (Lecci & Wirth, 2006). Participants (N = 1,096) were used to validate a method referred to as participant training that increases concordance between the expected and actual amount of time sampled in a commonly employed mood assessment instrument (the PANAS) as well as terms used in mood related research. Results indicate that exposure to other time frames can help to reduce the variability in “moment” and “year” ratings and increase variability for “in general” as well as result in greater concordance between expected and actual timeframes employed by participants. Furthermore, the study examines the effects of the variability in actual retrospective timeframes on the longstanding debate on the dimensionality of affect (e.g., Watson, 1988; Diener & Emmons, 1985; Warr, Barter, & Brownbridge, 1983; Russell & Carroll, 1999, etc.). Participant training does not effect the correlation between positive and negative affect, however, the terms themselves have a significant impact on the correlation. Implications for these findings are discussed.

ACKNOWLEDGEMENTS

I would first like to thank Dr. Lecci, whose patience, support, and encouragement were unwavering. His belief in me and this project gave me the strength to continue when I thought it was impossible. I would also like to thank the lab for all of their input and help throughout the past 3 years. Their assistance both in the creation and implementation phases of this project were invaluable. Finally, I would like to thank my committee members whose suggestions and interest greatly contributed to the final product.

DEDICATION

I dedicate this thesis to my cousin; Raylene Amber Michael, whose life on earth was far too short, but whose memory will live in my heart for eternity. I miss her every second of every day.

LIST OF TABLES

Table	Page
1. Definitions of the 3 conceptual timeframes and examples.....	4
2. Descriptive timeframe differences in instructional sets.....	5
3. A summary of descriptive timeframes used in mood assessment research.....	6
4. Interpretations of correlation coefficients by various researchers.....	20
5. Descriptive statistics for “moment”.....	25
6. Descriptive statistics for “month”.....	25
7. Descriptive statistics for “year”.....	25
8. Descriptive statistics for “in general”.....	26
9. Comparisons of correlations between PA and NA for the words used by Diener and Emmons (1985) and the words used in the PANAS (Watson, Tellegen, & Clark, 1988).....	31

LIST OF FIGURES

Figure	Page
1. A bubble plot of the variances (logged seconds) for each instructional set and order of presentation. Bubble size and shade reflect the proportion of variance observed compared to the Right Now by First Condition (Reproduced with permission from Lecci & Wirth, 2006).....	10

INTRODUCTION

Mood assessment has been prominent in psychology for many years. To get a sense of approximately how prominent mood assessments are within the field, a PsycINFO search of peer reviewed journals was performed on the following terms, in the abstract, title, and keyword sections: mood rating, mood assessment, mood measurement, mood evaluation, emotional rating, affect measurement, and affect evaluation. These searches resulted in 220 articles published in the past 5 years alone. These 220 articles are published in journals, including those specifically directed at issues relating to affect (e.g., *Cognition & Emotion*, *Journal of Affective Disorders*, *Emotion*, etc.), as well as in journals in applied areas of research (e.g., *International Journal of Sports & Exercise Psychology*, *Addiction*, *Journal of Sleep and Sleep Disorders*, *Journal of Psychosomatic Obstetrics & Gynecology*, etc.). It is important to note that this search is most likely a significant underestimation of the number of articles that involve mood assessment, despite the possibility for redundant results under these terms. Articles that involve specific mood assessments (i.e., depression, anxiety, sadness, etc.) or specific scales used to assess mood (e.g., the BDI, BAI, CES-D, PANAS, etc.) would probably not be included when searching for broad spectrum terms such as mood rating, despite the fact that they do assess mood. Moreover, this search would not include articles that do not have these words in the abstract, title or keywords, which could quite easily be the case for many articles that do assess mood, but where this does not reflect the primary research question.

A second PsycINFO search found that the most used mood assessment instruments over the past 5 years include the Beck Depression Inventory (BDI) (Beck & Streer, 1987), Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988), Beck Anxiety Inventory (BAI) (Beck & Streer, 1988), and the Multiple Affect Adjective Check List Revised

(MAACL-R) (Zuckerman & Lubin, 1965). This was determined by first searching the terms mood measurement and generating a list of the instruments that were cited most often in the PsycINFO database. The list was compiled and then each of the instruments received its own search. The four above-mentioned instruments resulted in the most hits in the abstract, title, and keyword over the past 5 years.

Although these mood measures are well validated and standardized, such that their acceptance for use has become somewhat routine in psychology, the primary thesis of this paper is that there is a fundamental flaw in at least one of the underlying assumptions for these measures. Specifically, each of these mood assessment instruments involves instructions that reference an expected timeframe. However, despite the widespread use of such measures and the implications for many sub-disciplines both within and outside of psychology, surprisingly little work has been devoted to exploring the instruments themselves and the manner in which they capture the moods they reportedly address. This is essentially an issue of examining the construct validity of these measures, and it is this line of research that has been largely neglected (for an exception see the published debates on the bipolarity of affect; e.g., Russell & Carroll, 1999; Watson, 1988; Watson & Clark, 1997; Diener & Emmons, 1985; Warr, Barter, & Brownbridge, 1983; Terry, Stevens & Lane, 2005, and the issue of multi-method assessments to separate random and nonrandom error in mood measurement; e.g., Green, Goldman, & Salovey, 1993; Russell & Carroll, 1999). The focus of this study, and the fundamental aspect of this literature that has virtually been ignored to date, is the ability of respondents to understand and effectively apply the expected timeframes referenced in the initial instructions for these mood measures.

Conceptualizing different timeframes

An important distinction needs to be made in order to understand the 3 types of timeframes that are discussed from here on out. The 3 different timeframes are the descriptive timeframe, the expected timeframe, and the actual retrospective timeframe. A descriptive timeframe consists of descriptive words used to characterize a timeframe in either a mood assessment measure or in research. For example, “How do you feel at this moment?” in this question “moment” would be the descriptive timeframe.

An expected timeframe is the amount of time one would expect the participant to take into account when assessing his/her mood. For example, if an individual is asked how he/she is feeling “today,” this might intuitively reflect a period of approximately 12 to 24 hours from which the person would sample their affective experience and use that information to assess his/her mood. It has been shown, however, that individuals vary greatly in the amount of time that is actually taken into account (i.e., sampled) when mood assessment measures are given (Lecci & Wirth, 2006). Therefore, an actual retrospective timeframe can be defined as the amount of time taken into account when a participant is asked to describe how he/she is feeling. For example, when asked about “today”, the range of time taken into account has been shown to be approximately 3 weeks (Lecci & Wirth, 2006). Table 1 summarizes these definitions and examples.

Table 1. *Definitions of the 3 conceptual timeframes and examples.*

<u>Descriptive</u>	<u>Expected</u>	<u>Actual Retrospective</u>
<ul style="list-style-type: none"> ■ The descriptive words used to characterize a timeframe in either a mood assessment measure or in research. ■ e.g., How do you feel <u>at this moment</u>? 	<ul style="list-style-type: none"> ■ The amount of time one would expect participants to take into account if they did a literal interpretation ■ e.g., When asked about “today”, intuitively/literally it would be equivalent to <u>24 hours</u> 	<ul style="list-style-type: none"> ■ The amount of time participants actually take into account. ■ e.g., When asked about “today”, the range of time taken into account has been shown to be <u>approximately 3 weeks</u> (Lecci & Wirth, 2006)

The use of timeframes in the literature

In order to address this issue, one can begin by examining the most frequently used instruments and identify their descriptive language. Table 2 summarizes the different descriptive timeframes used in each of 4 commonly used mood assessment instruments. It should be noted, that the differences in the descriptive timeframes are not incidental, and indeed have important theoretical and practical implications. For example, the PANAS employs the term “moment,” and this is thought to assess mood, whereas the term “in general” is thought to assess personality (Watson, Clark, & Tellegen, 1988). The MAACL-R makes a similar distinction in terminology to differentiate mood, which presumably samples a relatively short timeframe, from personality, which presumably samples a relatively long timeframe (Zuckerman & Lubin, 1965). This illustrates that the descriptive temporal language is a critical component of these measures and the constructs they purport to assess.

One can also examine the temporal language employed in mood-related research that does not involve well established mood assessment instruments. Table 3 summarizes the descriptive timeframes and there is clearly considerable overlap between this language and that

employed in the most commonly used instruments. The consistency of this language means that we have created a culture of intuitive understanding of what these terms mean, even though there is little or no data to explicitly study the meaning of that language.

One additional problem that exacerbates the lack of standardized actual retrospective timeframes is that there is a problem in the consistency of the language used to assess and describe presumably similar mood states (Lecci & Wirth, 2006). An example of this would be that many methods of measurement use different wording when describing what appear to be similar retrospective timeframes used in mood assessment (see, for example, Tables 1 & 2). As an illustration, the descriptive timeframe of “at this moment” (Watson, Clark, & Tellegen, 1988; Feldman Barrett, 1997) and “right now” (Terry, Stevens, & Lane, 2005; Rasmussen, Jeffrey, Willingham, & Glover 1994) appear to be targeting the same expected timeframe, but they do so with different language. What is unclear is whether these linguistic differences have any consequences with respect to the timeframes that they elicit from respondents. Given the variability in how individuals interpret the same descriptive timeframes, it appears to introduce unnecessary and unknown variance to employ different terminology.

Table 2. *Descriptive timeframe differences in instructional sets.*

Instrument	Descriptive Timeframe
BDI (Beck & Streer, 1987)	“past week, including today”
PANAS (Watson, Clark, & Tellegen, 1988)	“moment”, “today”, “past few days”, “week”, “past few weeks”, “year”, “general”
BAI (Beck & Streer, 1988)	“past week, including today”
MAACL-R (Zuckerman & Lubin, 1965)	“now-today”, “in general”

Table 3. *A summary of descriptive timeframes used in mood assessment research.*

Literature	Descriptive Timeframe
Thomas & Diener 1990	“currently” “for the entire day”
Winkielman, Knauper, & Schwarz 1998	“typical week” “typical year”
Terry, Stevens, & Lane 2005	“right now” “over the past week including today”
Rasmussen, Jeffrey, Willingham, & Glover 1994	“right now” “over time”
Feldman Barrett 1997	“at the moment of rating” “past 3 months in general”

The lack of standardization in the terminology used within these measures, especially when describing transient states, is highly problematic because it does not ensure that subjects are referring to the same actual retrospective timeframe. The results of a study by Lecci and Wirth (2006) suggest that many of the timeframes for phrases such as “right now” and “at this moment” are interpreted by subjects to be longer than what one might intuitively assume. If participants do not interpret all of these descriptive timeframes in the same manner, then this undermines standardization, which, in turn, unnecessarily compromises both reliability and validity.

The validity of retrospective ratings hinges on the assumption that an individual can accurately assess the moment (Russell & Carroll, 1999). However, there is evidence that this assumption may not be met with respect to retrospective timeframes (e.g., Lecci & Wirth, 2006; Watson, Clark, & Tellegen, 1988). Recent literature has found that even the same terminology generates great amounts of variability when asking about actual retrospective timeframes (Lecci & Wirth, 2006). This undermines both internal and external validity. A simple example would be if two participants were given a mood induction. Both participants are asked how he/she feels “at this moment”. The first participant interprets “at this moment” to mean the past week, whereas,

the second participant interprets this timeframe to mean the past 30 seconds. As a result, the first participant is not likely to change their mood rating to the same degree after the mood induction regardless of whether or not the mood induction worked, given the actual retrospective timeframe they employed. The second participant, however, could change their mood rating dependent upon whether or not the mood induction worked. In this example, the researcher might think that the mood induction did not work on the first participant, but did on the second if they do not take into account the differences in actual retrospective timeframes. Despite the proliferation of timeframe terminology, little research has systematically examined whether such differences in terminology have functional consequences for the assessments.

In a recent study, participants were asked how they felt at this moment, right now, today, and in general (Lecci & Wirth, 2006). After making their mood ratings the participants were asked how much time they took into account when making their assessments. A pilot study found that on average 3.35 days were considered when asked about “right now”. A second study, with a similar method as the pilot study, revealed an average timeframe of 2 and ¾ weeks when respondents were asked to rate their mood based on the descriptive timeframe “at this moment” (Lecci & Wirth, 2006). The results also indicated that the shortest range was 3 weeks when participants were asked about “today”. Interestingly, the participants’ rating of timeframe when “today” was used, most closely approached 24 hours. In other words, the data collected more closely matched the expected timeframe when employing the term “today” (Lecci & Wirth, 2006). Of course, this could be a direct function of the fact that the term “today” is one of the easiest to match to an expected timeframe. Obviously, there could likewise be a greater match with intuitive timeframes for phrases such as “the past year”, however there is no data published to support that “the past year” more closely corresponds to the expected timeframe relative to

other phrases. Also noteworthy is that the phrases most commonly used in the literature, such as, “at this moment”, “currently”, “right now”, and “in general” are more obscure and produce greater variability in actual retrospective timeframes, than “today” which is rarely used in the mood assessment literature. Lecci & Wirth (2006) found that the descriptive timeframe accounted for 74% of the variance in the retrospective timeframes (i.e., time taken into account when making the assessment) in a within subject one-way ANOVA, and that the instructional set had a significant effect.

Other indirect evidence that actual retrospective timeframes do not match expected timeframes has been found in experiments of test-retest reliability of instructional sets (e.g., Watson et al., 1988). If participants are using actual retrospective timeframes that are congruent with the expected timeframes, then the test-retest reliability should increase as the descriptive timeframe increases, because as the descriptive timeframe increases one would expect to be assessing less transient mood states (i.e., assessing personality instead). However, Watson et al. (1988) found the only significant differences between test-retest reliability coefficients to be between “in general” and the terms “today,” “past few days,” and “past week”. This means that there was not a significant difference between the test-retest reliability coefficients for “in general” and “at this moment.” Of course, these two terms should intuitively result in the greatest difference because “in general” is supposed to be assessing personality and “at this moment” is supposed to be measuring the mood state an individual is currently experiencing.

Participant training to improve convergence with expected timeframes

In order to obtain accurate results, it is important to develop mood assessments that have participants rating a more specific timeframe. Lecci and Wirth (2006), in an attempt to refine the retrospective timeframe, found a reduction in variability of descriptive timeframes by presenting

them in the context of other descriptive timeframes. In the study, they counterbalanced the order of 4 different timeframes: right now, at this moment, today, and in general. The results revealed that the actual retrospective timeframes most closely corresponded with the expected timeframe when it was presented to the participant in the third or fourth ordinal position. For example, “at this moment” had the least variability across subjects when it was preceded by the other three timeframes (Lecci & Wirth, 2006). Figure 1 is a variance bubble plot that shows the variability of the four timeframes across the four ordinal positions. As can be seen by the figure, the terms “in general” and “at this moment” have the greatest reduction in variability, while the variability for “right now” is fairly consistent across all four ordinal positions (Lecci & Wirth, 2006).

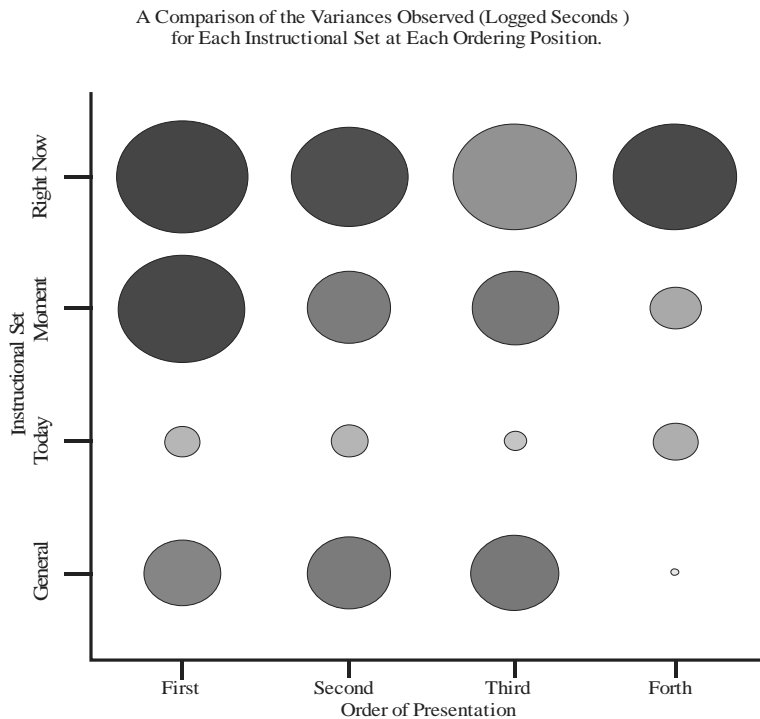


Figure 1. A bubble plot of the variances (logged seconds) for each instructional set and order of presentation. Bubble size and shade reflect the proportion of variance observed compared to the Right Now by First condition. (Reproduced with permission from Lecci & Wirth, 2006).

These results emerged regardless of what timeframes were in the first three ordinal positions, meaning that the order did not have to go from longest timeframe to shortest timeframe or vice versa. This suggests that being exposed to multiple timeframes could, by comparison, help the participant more closely match the expected timeframes to the actual retrospective timeframes, or it may bring more attention to the descriptive timeframes being used. In either case, this reduction in variability can be considered as a type of participant training. The mere exposure to different descriptive timeframes causes them to more closely match their ratings to the expected timeframes. Interestingly, what was ostensibly a form of participant training has been used in other studies that were not explicitly looking at the accuracy of actual retrospective timeframes (e.g., Winkielman, Knauper, & Schwarz, 1998; Green, Goldman, & Salovey, 1993; Diener & Emmons, 1985). This unintentional participant training occurs when the method involves participants being exposed to multiple descriptive timeframes. Exposure to multiple descriptive timeframes can have at least two consequences. First, the participant is exposed to the mood rating method more than once, thus leading to rehearsal effects. Rehearsal effects can make the participant more aware of the descriptive timeframes they are assessing. That is, simple repeated exposure to the descriptive timeframe may make one more aware of the timeframe, which in turn, could make the participant more exact (temporally speaking) in their assessment. An alternative explanation is that when the participants are exposed to multiple descriptive timeframes they can comparatively evaluate the timeframes, which could also increase specificity. These effects can be seen by comparing within (which involves rehearsal and comparative evaluations) and between subject methodologies in the literature (which does not involve rehearsal or comparative evaluations).

The use of participant training to explicitly improve a participant's ability to have actual retrospective timeframes more closely match expected timeframes could increase reliability and validity in the assessment of mood states. Currently this type of participant training is not a standard in mood assessment, and variability in the interpretation of timeframes could be affecting other inconsistencies in the empirical literature.

Until recently, there was little if any literature devoted to examining the amount of time taken into account when individuals are asked to assess their mood. Previous literature has however, addressed the discrepancies in on-line and retrospective mood assessments. Thomas and Diener (1990) concluded that an individual's retrospective reports of affective states tend not to be extremely accurate, meaning that when the participants are asked to generate a summative rating of their on-line emotional ratings, they are not consistent with the mathematical summing of the retrospective ratings. For example, if a participant is asked to make hourly mood ratings, then to rate his/her average mood for the day, it does not equal the mathematical average of all the hourly ratings. These inconsistencies could be due to inaccurate memories (e.g., Feldman Barrett, 1997) or it could be that the participants are not taking into account the actual retrospective timeframes that correspond to the expected timeframes.

Other research confounds that could be attributed to variability in mood timeframes

Another area of the mood literature that could be affected by actual retrospective timeframes is the inconsistencies surrounding the dimensionality of affect. There has long been a debate in the mood literature revolving around the dimensionality of positive affect (PA) and negative affect (NA) (e.g., Watson, 1988; Diener & Emmons, 1985; Warr, Barter, & Brownbridge, 1983; Russell & Carroll, 1999, etc.). Specifically, there are opposing views with respect to the independence versus the bipolarity of positive and negative affect. Researchers

who endorse the independence of PA and NA believe that a person can feel varying degrees of both positive and negative affect at the same moment (Watson, 1988; Diener & Emmons, 1985; Diener & Iran-Nejad, 1986). In other words, the experience of PA and NA and the subsequent rating of PA and NA are independent of each other (Diener & Emmons, 1985; Bradburn, 1969). The other view of dimensionality is that PA and NA are on a bipolar continuum (i.e., negatively correlated), which means that a person can feel either positive or negative affect, but can not simultaneously experience both moods (Warr et al., 1983; Russell & Carroll, 1999; Green, Goldman, & Salovey, 1993).

The debate over these two models has been an extensive one, partly due to the inconsistent results within the literature (e.g., Diener & Emmons, 1985; Watson, 1988; Watson & Clark, 1997). One of the more specific areas of debate arises from differing results with regards to the measurement of time (Watson & Clark, 1997).

Diener and Emmons (1985) found that shorter descriptive timeframes (e.g., “at this moment”) are negatively correlated (bipolar), but as the descriptive timeframes become what would intuitively be defined as longer periods of time (e.g., “during the last year”), the scores approached independence (i.e., they were less negatively correlated or independent). Diener et al. (1985) observed these results across 4 studies involving the descriptive timeframes “past year”, “past month”, “past 3 weeks”, and daily “moment” ratings.

In all four studies, Diener & Emmons (1985) found significant within subject results. In other words, there were differences in independence and bipolarity as a function of descriptive timeframe within a single subject’s rating. They did not, however, find significant results between subjects. One possible explanation for this could be due to the variability found in the time people take into account when assessing descriptive timeframes (Lecci & Wirth, 2006).

This variability creates noise in the data for between subject analyses (i.e., extensive variability exists in the retrospective timeframes employed as well as divergence from expected timeframes). This noise could be contaminating the between subject analyses, although it should have a significantly smaller effect on the within subject analysis. There are several reasons behind why the variability has an effect between subjects, but not within, and these reasons will be discussed later.

In an attempt to replicate the findings of Diener and Emmons (1985), Watson (1988) asked participants to rate their mood over differing descriptive timeframes. While Diener and Emmons (1985) asked each participant to rate over the “past year” and “past month”, Watson (1988) had participants rate how they felt over one of the following timeframes: “moment”, “today”, “past few days”, “past few weeks”, “past year”, and “general”. Using these descriptive timeframes instead of “past month” and “past year”, did not yield significant differences as a function of timeframe. Watson (1988) found similar negative correlations across the differing descriptive timeframes, thus concluded that there was not a systematic timeframe effect. The problem with this conclusion is that the method for Watson (1988) and Diener and Emmons (1985) differ on a number of aspects.

Diener and Emmons (1985) asked each participant, in study 2 to assess their mood over the “past month” and the “past year”. Using this method they found significant within subject differences in mood ratings with respect to descriptive timeframes, but non significant between subject differences. In studies 3-5, Diener and Emmons (1985) asked participants to complete mood assessments daily, momentary, and at three week intervals. In order to analyze the data, they aggregated across momentary assessments to create longer timeframes. Using the aggregation technique, they found that as the aggregated timeframe increased, PA and NA

became less negatively correlated, moving towards independence. In each study (3-5), they found significant within subject results, but not between subject results.

Watson (1988) asked each participant to assess their mood using one of six different descriptive timeframes (moment, today, past few days, past few weeks, past year, general). He then compared between subject differences across descriptive timeframes and did not find significant differences. Also, he did not aggregate across moments to increase the retrospective timeframe; he used descriptive timeframes that were assumed to be increasingly longer than the other descriptive timeframes (Watson, 1988).

Actual retrospective timeframes can help account for the inconsistencies between these two studies. First, Lecci and Wirth (2006) found a reduction in variability when an individual encounters multiple descriptive terms. Their results revealed that the time taken into account more closely approaches the expected timeframe when the descriptive terms appear in the third or fourth ordinal position (i.e., the participant is essentially undergoing training). In other words, there is less variability in time taken into account after the individual has been exposed to other descriptive timeframes, presumably due to a comparative and/or rehearsal effect. These findings give one possible explanation for the results found by Diener and Emmons (1985). In their studies, each individual encountered multiple descriptive timeframes, which would suggest that the time taken into account was more congruent with the expected timeframes, thus revealing significant within subject findings.

The amount of time an individual takes into account when making retrospective mood ratings can also explain the non significant between-subject findings. The variability in the amount of time taken into account between subjects can create noise that will contaminate the results of a between subjects analysis. This noise can also be the reason that Watson (1988) did

not find significant between subject differences, because his subjects were not exposed to multiple descriptive timeframes (see Lecci & Wirth, 2006).

The differing results that emerged due to aggregation versus non-aggregation can also be partially explained by the variability in actual retrospective timeframes. It is important to note that there are two types of aggregation. The first is aggregation by the subject (i.e., subject-driven aggregation). An example of this method of aggregation would be if participants were asked to fill out momentary ratings through the day, then at the end of the day, they are asked to assess how they felt “today”. This type of aggregation is presumably asking the participant to take the momentary ratings and put them together into a single summative rating. This type of aggregation is not problematic because the participants are able to interpret the descriptive timeframe of “today”, which, as can be seen by prior research, does not necessarily add up each moment to produce a sum of 24 hours. Aggregation by the researcher (researcher-driven aggregation), however, is when the participant is asked to make momentary ratings and then the researcher adds the ratings together to make longer expected timeframes (e.g., Diener & Emmons, 1985). By aggregating across momentary ratings to create longer timeframes, expected timeframes are forced, which, given the results of Lecci & Wirth (2006), may not be an accurate reflection of the subject-driven aggregation. The inaccuracy of aggregation can also be seen when we compare Watson (1988) and Diener and Emmons (1985) because when participants’ momentary ratings are aggregated, as in Diener and Emmons, PA and NA become less negatively correlated as the descriptive timeframe increases. However, when Watson (1988) increases the descriptive timeframe by asking about intuitively longer timeframes (i.e. past year) the negative correlation between PA and NA does not significantly differ.

The current study

Given the possible compromises to reliability and validity of mood assessments due to the variability of actual retrospective timeframes, it is important to illustrate experimentally that timeframes can account for some of the inconsistencies in the literature. Ultimately, the goal would be to eventually (1) standardize the language (i.e., the descriptive timeframe), (2) develop an instructional method that would reduce variability in the interpretations of the language, and/or (3) quantify the variance in ratings due to this variability in interpreting timeframes so as to factor this variability out when examining other issues. The current study examines the concept of participant training, and its ability to decrease variability due to actual retrospective timeframes. There is evidence that participant training, although not explicitly referred to as such in the current literature, is contributing to some of the inconsistencies within mood assessment (e.g., Lecci & Wirth, 2006; Diener & Emmons, 1985; Watson, 1988, etc.). For example, Lecci & Wirth (2006) were not intentionally employing participant training. Instead, they were counterbalancing the descriptive timeframes to check for order effects and found that there was evidence for a reduction in variability due to the ordinal position of the timeframe for the terms “at this moment”, “today”, and “in general”, with the greatest reduction in variability being for “at this moment” and “in general.” These results suggest that a participant training effect may not occur equally for all phrases, as terms such as “today” may already provide sufficient temporal information so as to minimize variability; at least to some degree (see Lecci & Wirth, 2006).

The terms “at this moment” and “in general” are also the phrases that represent the shortest and longest timeframes from the PANAS (Watson, Clark, & Tellegen, 1988). Therefore, these terms have considerable ecological validity with respect to one of the most commonly used measures in the field.

In the current study, participants were exposed to participant training. Each group of participants was exposed to one of four possible orders, with each phrase of interest (i.e., the past month, the past year, in general, and at this moment) occurring in both the first and fourth ordinal position. These four phrases represent the two phrases (in general, and at this moment) that had the greatest reduction in variability in Lecci & Wirth (2006) and the two phrases (the past month and the past year) used by Diener & Emmons (1985). Importantly, the results of Lecci and Wirth (2006) indicate that it is not necessarily a true order effect per se, such that one would need to investigate all possible orders. Rather, the critical issue appears to be the presence or absence of comparative temporal evaluations. Thus, only the first ordinal position (no comparison is possible) and the last (after three other assessments have been completed, thereby necessitating a temporal comparison) ordinal positions were considered.

Accordingly, it was hypothesized that: 1. Variability in mood ratings due to actual retrospective timeframes will decrease following participant training, 2. The terms “year” and “month” will exhibit less reduction in variability relative to “in general” and “at this moment,” as the former terms already have a clearer intuitive timeframe (i.e., year = approximately 365 days, and month = approximately 30 days) and 3. Mean¹ actual retrospective timeframes will more closely match the expected timeframe when the ratings occur in the fourth ordinal position and this effect should be strongest for the terms with less obvious expected timeframes (i.e., “in general,” and to a less extent “at this moment” relative to “month” and “year”). That is, participant training will have a stronger effect on the terms “in general” and “at this moment” as

¹ Past research (Lecci & Wirth, 2006) revealed skewed data when examining actual retrospective timeframes. For this reason, the distribution will be examined and median values may be used if they prove to be a more meaningful representation of the collected data.

compared to “month” and “year.” The rationale for this hypothesis is that terms with more obvious expected timeframes should be less influenced by the participant training.

An important implication of this research is the effect of actual retrospective timeframe variability on the debate concerning the bipolarity of affect. For that reason this study will also attempt to use actual retrospective timeframes to account for some those inconsistencies. Diener and Emmons (1985) found that PA and NA are less negatively correlated as the descriptive timeframe increases when examining within subject data, but not between subject data. The current study will use their affect terms and descriptive timeframes in an attempt expand their findings to between subject data. In order to determine if actual retrospective timeframes account for the inconsistencies surrounding the debate over the dimensionality of affect, the current study will employ participant training as a method to reduce noise in the between subject data.

Green et al. (1993) propose that it is important to separate random and nonrandom error when obtaining correlations of PA and NA using a single method approach. When this error is accounted for in multi-method assessments they conclude that a largely bipolar model emerges (Green et al., 1993). However, the majority of the mood assessment literature does not employ this technique, and because the current study is attempting to account for some of the differences between studies using single method assessments, the biasing factor of combining random and nonrandom error will not be examined.

In a review of the literature on the dimensionality of affect, coefficients in the range of -.01 to -.23 have been interpreted as providing evidence for the independence of PA and NA with an average correlation coefficient of -.11 (Warr et al., 1983; Tellegen et al., 1994; Green et al., 1993; Diener & Emmons, 1985; Diener & Iran-Nejad, 1986; Watson & Clark, 1997). Coefficients ranging from -.26 to -.85 have been interpreted as evidence for the bipolarity of

affect, with an average of $-.50$ (Warr et al., 1983; Tellegen et al., 1994; Green et al., 1993; Diener & Emmons, 1985; Diener & Iran-Nejad, 1986; Watson & Clark, 1997). Table 4 summarizes the coefficients obtained by different researchers and their interpretation of each coefficient. It is likely that the results of this study will fall somewhere between the averages ($-.11$ to $-.50$) before participant training. It is hypothesized that after employing participant training the coefficients will move significantly in one direction or the other. If longer descriptive timeframes reveal coefficients moving towards a zero correlation and shorter descriptive timeframes reveal stronger negative correlations, then that will be evidence for the model proposed by Diener & Emmons (1985) in which affect is independent as the timeframe increases. However, if the coefficients for all descriptive timeframes become more negatively correlated, then there will be more evidence to support the bipolarity of affect (Russell & Carroll, 1999, Green et al., 1993, etc.). Regardless of which direction the coefficients go, it is hypothesized that it will be significantly different from the observations taken before participant training, as the primary function of participant training is to decrease noise (error) in the data.

Table 4. *Interpretations of correlation coefficients by various researchers.*

Literature	Bipolar	Independent
Warr et al., 1983	-.54 -.26	-.01 -.12
Tellegen et al., 1994	-.47 -.61	—
Green et al., 1993*	-.53 -.56 -.61 -.53	—
Diener & Emmons, 1985	-.42 -.57 -.31 -.85	-.10 -.10 -.15
Diener & Iran-Nejad, 1986	-.38 -.39	-.05 -.11
Watson & Clark, 1997	—	-.05 -.13 -.06 -.15 -.17 -.23 -.14 -.13
Average	-.50	-.11

*For the purpose of standardization across studies, we used the raw values, not the latent corrected values from Green et al., 1993.

METHOD

Participants

The participants for this study were undergraduate students who received partial class credit for their participation. Study 5 of Diener and Emmons (1985) revealed that college participants and participants collected from the local community produced similar findings with regards to correlations between PA and NA. Therefore, this study focused only on college-aged participants.

An a priori power calculation indicated that, given a small effect size (Eta values of approximately .2 to .3), each of the 4 groups will need to contain 274 participants, totally 1096 students, in order to achieve an adequate degree of power (i.e., power of .80; Cohen, 1992). Participants were recruited in introductory level classes during the last 5-10 minutes of class and asked to fill out a questionnaire about their mood.

Materials

Four different questionnaires were used (see appendix). Each questionnaire asks participants to rate the extent to which they have experienced 32 mood related words over differing descriptive timeframes. The first 20 words come from the PANAS short version and an additional 12 words come from the studies done by Diener and Emmons (1985). The instructions and the order of the first 20 words for each questionnaire have been taken from the PANAS (Watson, Clark, & Tellegen, 1988). Diener and Emmons (1985) did not specify the order in which their words were presented. Therefore, the word order for this set of words was randomized. The 5-point Likert scale used in the PANAS was also used for both sets of words, despite the fact that Diener and Emmons (1985) used a 7-point Likert scale. Prior research on Likert scales suggests that after a 3-point scale there are no significant differences with respect to “stability, predictive validity, or concurrent validity” in scales containing up to 19 items (Matell & Jacoby, 1972). One point of concern in the literature is that 3 and 5-point Likert scales show a greater percentage of “uncertain” responses. For this reason, the current study did not employ a scale containing an “uncertain” response (Matell & Jacoby, 1972). Instead, the middle rating was labeled as “moderately.” The term “moderately” is important in the mood literature because it can represent an actual state. It is not an entirely different rating as would be the case if the middle option were “uncertain”, rather it is a further quantification of the construct.

The current study also employed a unipolar rating scale, and this is consistent with the PANAS (Watson, Clark, & Tellegen, 1988). It is important to use a unipolar rating scale because a unipolar scale will be able to assess both bipolarity and independence, whereas a bipolar scale forces bipolarity rather than testing for it (Watson & Tellegen, 1999; Russell & Carroll, 1999). This type of response format is also important because as Comrey (1988) explains it is crucial for the participants to find a “response alternative that is reasonably appropriate”. The use of a bipolar scale would force participants to be low on PA if they rated high on NA, which might not be the appropriate response.

Using PA and NA terms from both the PANAS and Diener and Emmons (1985) will speak to both research and the most frequently used scale to assess mood. It has also been suggested that there are differences in correlations between PA and NA regarding each set of words, with the PANAS revealing smaller correlation coefficients (closer to zero) as compared to other measures (e.g., Diener & Emmons, 1985; Watson, 1988).

Each word was evaluated on a 5-point Likert scale, with 1 denoting “very slightly” to 5 denoting “extremely.” The Likert scale and instructions are that of the PANAS with the exception of the descriptive timeframes. Each questionnaire contains 4 different descriptive timeframes. The order of the descriptive timeframes according to the questionnaire was as follows:

Questionnaire one; 1. the past month, 2. in general, 3. the past year, 4. at this moment;

Questionnaire two; 1. the past year, 2. the past month, 3. at this moment, 4. in general;

Questionnaire three; 1. in general, 2. at this moment, 3. the past month, 4. the past year;

Questionnaire four; 1. at this moment, 2. the past year, 3. in general, 4. the past month.

After completing the mood ratings for the first descriptive timeframe (no prior participant training) participants were asked “How much of your life did you take into consideration when making your mood ratings?” (Lecci & Wirth, 2006). Participants were then asked to report this using the same format as Lecci and Wirth (2006), which is to first report a numerical value and then to circle the correct timeframe (e.g., seconds, minutes, hours, etc.) This question was then asked after the mood ratings for the fourth descriptive timeframe (the most participant training).

Procedure

Each participant was randomly assigned to one of four groups. The groups correspond to the 4 different questionnaires. The questionnaires were passed out in introductory classes in the last 5-10 minutes of class. Each participant filled out the form and returned it to the research assistant. During data collection, research assistants watched as participants filled out the questionnaires. If a participant was seen looking through the entire packet before filling it out, the researcher was notified and that questionnaire was not used. The importance of not using these questionnaires was an attempt to get an accurate rating in the first ordinal position (i.e., before participant training). If the participant looked at the entire questionnaire before completing it participant training would have occurred before the first rating, thus contaminating the results. Contamination was also controlled for by manually checking each questionnaire before data entry. Questionnaires that were not completely filled out were excluded, as well as those who did not fill them out correctly. The questionnaires were set up so that each of the following phrases were in the first ordinal position and the fourth ordinal position; the past month, the past year, at this moment, and in general. The ratings made after the first and fourth descriptive timeframe of actual time taken into account (i.e., actual retrospective timeframe) were first converted to seconds. The conversion to seconds facilitated a comparison with data

collected by Lecci and Wirth (2006). The Mann-Whitney statistic was then used to determine the effects of participant training. With this test, it can be determined if two distributions are significantly different. The Mann-Whitney is a non-parametric test denoted by U , where U is the number of pairs (X,Y) , where $X < Y$ (Higgins, 2004). Due to the large numbers involved a Z score is reported instead of a U value, for ease of comprehension.

RESULTS

Examining the effects of participant training

Due to the highly positively skewed distributions that were obtained from the data, Mann-Whitney rank sum tests were used to determine if there is a significant difference in the median actual retrospective timeframe as a function of ordinal position of descriptive timeframes.

The Mann-Whitney rank sum test comparing the actual retrospective timeframe for “moment” is significant, $z = -1.843$, $p = .03$, with moment in the fourth position (MR = 262.05) being smaller than in the first position (MR = 286.95). The actual retrospective timeframe for “past month” is significantly larger in the fourth position (MR = 287.60) than the first position (MR = 261.40), $z = -2.098$, $p = .036$. For “past year”, the actual retrospective timeframe is significantly larger in the fourth position (MR = 302.97) than the first position (MR = 246.03), $z = -4.781$, $p = .000$. Finally, the actual retrospective timeframe for “in general” is significantly larger in the fourth position (MR = 335.09) than in the first position (MR = 213.91), $z = -8.966$, $p = .000$. For “moment” and “in general”, a one tail test was used because it was hypothesized that the actual retrospective timeframes for “moment” would decrease after participant training, while the actual retrospective timeframes for “in general” would increase. Two tailed tests were used for both “month” and “year” because it was not hypothesized which direction the actual

retrospective timeframes would go after participant training. The mean ranks in these analyses are somewhat more ambiguous and difficult to interpret due to the skewed nature of the data and presence of extreme outliers. In order to clarify some of the results, tables 5-8 summarize the median, mode, N at the mode, the frequency %, and the cumulative % for each descriptive timeframe in the first and fourth ordinal position.

Table 5. Descriptive statistics for “moment”.

Moment	First	Fourth
Median	60 minutes	42.5 minutes
Mode	60 minutes	60 minutes
N at Mode	23	29
%	8.4	10.6
Cumulative %	51.1	60.9

Table 6. Descriptive statistics for “month”.

Month	First	Fourth
Median	30 days	30 days
Mode	30 days	30 days
N at Mode	136	153
%	49.6	55.8
Cumulative %	88.0	84.3

Table 7. Descriptive statistics for “year”.

Year	First	Fourth
Median	1 year	1 year
Mode	1 year	1 year
N at Mode	134	199
%	48.9	72.6
Cumulative %	93.1	93.4

Table 8. Descriptive statistics for “in general”.

In General	First	Fourth
Median	51 days	2 years
Mode	1 year	2 years
N at Mode	26	28
%	9.5	10.2
Cumulative %	83.6	51.5

The effects of participant training on variability

The first two hypotheses state that 1. Variability in mood ratings due to actual retrospective timeframes will decrease following participant training and 2. The terms “year” and “month” will exhibit less reduction in variability relative to “in general” and “at this moment,” as the former terms already have a clearer intuitive timeframe (i.e., year = approximately 365 days, and month = approximately 30 days). Due to the non-normal distribution of the data, the interquartile range (IQR) was used to test these hypotheses. The data is presented in the smallest unit of analysis provided by the participants (i.e., seconds), and all values were converted to be on the same scale.

For “moment” in the first position IQR = 172,680, whereas after participant training the IQR reduces by 86,400 for an IQR = 86,280 in the fourth position. In the case of “month” the IQR does not change as a function of participant training. For “month” in the first and fourth position the IQR = 777,600. For “year” the IQR reduces by 10,000,000 from the first position (IQR = 10,000,000) to the fourth position (IQR = 0). Finally, for “in general” the IQR actually increases by 162,828,800 from the first position (IQR = 29,395,200) to the fourth position (IQR = 192,224,000).

The results of Levene’s homogeneity of variance test (which includes all the data) resulted in 3 of the 4 descriptive timeframes having significant differences in variability by

comparing the first and the fourth position. More specifically, variance for “month” was significantly different in the fourth position than the first, $F = 23.3, p = .000$. The variance for “year” was also significantly different as a function of ordinal position, $F = 5.5, p = .02$. Finally, the variance for “year” was significantly different from the first to the fourth position, $F = 133.3, p = .000$. Although the variance for “moment” did not change significantly as a function of descriptive timeframe, it was approaching significance, $F = 2.4, p = .12$.

Due to the non-normal distribution of the data, a second significance test was performed to test changes in variability. A method proposed by Shoemaker (1995) examines the variability for non-normal data by looking at the interquartile range. Although, this is a very conservative method, as it eliminates 50 percent of the data (25 percent on each side). However, similar results emerged as when using all the data. “Year” ($F = 10,000,000, p = .000$) and “general” ($F = 6.53, p = .01$) resulted in significant differences in variance as a function of ordinal position. “Moment” also resulted in similar findings as the less conservative test, results in a change in variance that approaches significance, $F = 2.0, p = .16$. The only major difference between the two tests was in regards to “month”. Using Shoemaker’s method (1995), the variance for month was not significantly different as a function of ordinal position, $F = 1.00, p = .31$.

Examining the association between NA and PA

In order to determine if actual retrospective timeframes are contributing to the inconsistencies surrounding the debate over the dimensionality of affect, a Fisher’s z transformation of independent samples was conducted using the mean correlation coefficients calculated from previous literature (Warr et al., 1983; Tellegen et al., 1994; Green et al., 1993; Diener & Emmons, 1985; Diener & Iran-Nejad, 1986; Watson & Clark, 1997). These analyses determined if the correlation in the first position was considered independent or bipolar as

compared to previous literature. In some cases the correlation coefficient for the current data was significantly different from both the mean independent correlation coefficient and the mean bipolar correlation coefficient from previous literature. In this situation, the correlation coefficient obtained from the current data was determined bipolar or independent based on the ranges for bipolarity (i.e., $-.26$ to $-.85$; $M = -.50$) and independence (i.e., $-.01$ to $-.23$; $M = -.11$) obtained from the literature. It is important to note that the range for bipolarity is nearly three times larger than that of independence. For most of the analyses this is irrelevant because comparisons are being made to the mean. However, in the case that the mean does not statistically fit into one category or the other the range is used to determine if the correlation is bipolar or independent. It is in this case that it becomes important to note the differences in the ranges. There is an increased likelihood that a correlation will be determined bipolar when using the ranges, which could mask a true shift in correlations to emerge from the first the fourth position. Similarly, a Fisher's z transformation is used to compare the correlation coefficient of PA and NA from the first position to the fourth position. Both of these analyses were done using the words from the PANAS and the words used by Diener and Emmons (1985).

Starting with the words from the PANAS, for "moment" in the first position, the correlation between PA and NA is independent ($r = -.19$) because it is significantly different from the hypothesized value of for bipolarity (i.e., $r = -.5$), $z = 5.87$, $p = .000$. "Moment" in the first position is not significantly different from the hypothesized value for independence ($r = -.11$), $z = -1.35$, $p = .18$. For "moment", there is a significant difference between the first position ($r = -.19$) and the fourth position ($r = -.007$), $z = 2.23$, $p = .025$, with PA and NA becoming more independent in the fourth position.

For “past month” in the first position, the correlation coefficient ($r = .016$) is significantly different from both the hypothesized value for bipolarity and independence. Therefore, as mentioned earlier it was determined that this value can be considered independent, based on previous literature (i.e., the value does not differ significantly from zero, therefore it is functionally independent). For “past month”, there is a significant difference between the first position ($r = .016$) and the fourth position ($r = -.678$), $z = -8.11$, $p = .000$, with PA and NA becoming more bipolar in the fourth position.

When “past year” is in the first position, the correlation coefficient ($r = -.13$) is independent because it is significantly different from the hypothesized value for bipolarity, $z = 6.86$, $p = .000$ and is not significantly different from the hypothesized value for independence, $z = -.345$, $p = .73$. For “past year”, there is a significant difference between the first position ($r = -.134$) and the fourth position ($r = -.073$), $z = 2.33$, $p = .02$, with PA and NA becoming more independent in the fourth position.

When the term “In general” is in the first position, the correlation coefficient ($r = .01$) is independent because it is significantly different from the hypothesized value for bipolarity, $z = 9.21$, $p = .000$ and is not significantly different from the hypothesized value for independence, $z = 1.97$, $p = .05$. For “in general”, there is not a significant difference between the first position ($r = .01$) and the fourth position ($r = -.08$), $z = -1.05$, $p = .29$.

When examining the words used by Diener and Emmons (1985), the results are different. For “moment” in the first position, the correlation between PA and NA is bipolar because it is not significantly different from the hypothesized value for bipolarity ($r = -.5$), $z = -.23$, $p = .82$ and it is significantly different from the hypothesized value for independence ($r = -.11$), $z = -7.45$, $p = .000$. For “moment”, there is no significant difference between the first position ($r = -$

.51) and the fourth position ($r = -.375$), $z = 1.89$, $p = .06$. The value does approach significance, with the data indicating that the correlation coefficients for PA and NA are becoming less bipolar (more independent) in the fourth position.

When “Past month” is in the first position, the correlation coefficient ($r = -.3$) is significantly different from both the hypothesized value for bipolarity and independence. Therefore, as mentioned earlier it was determined that this value can be considered bipolar, as the value falls within the range for bipolarity based on previous literature and it does differ significantly from zero. For “past month”, there is not a significant difference between the first position ($r = -.3$) and the fourth position ($r = -.4$), $z = -1.33$, $p = .18$.

When “Past year” is in the first position, the correlation coefficient ($r = -.34$) is significantly different from both the hypothesized value for bipolarity and independence. Therefore, as with “past month”, it was determined to be bipolar based on the range of values found in the previous literature. For “past year”, there is a not significant difference between the first position ($r = -.34$) and the fourth position ($r = -.238$), $z = 1.27$, $p = .2$

For “in general” in the first position, the correlation coefficient ($r = -.32$) is significantly different from both the hypothesized value for bipolarity and independence. Therefore, it too was determined to be bipolar by using the ranges obtained from previous literature. For “in general”, there is not a significant difference between the first position ($r = -.32$) and the fourth position ($r = -.36$), $z = -.06$, $p = .95$.

Examining the effects of different affect terms on the correlation between PA and NA

While doing Fisher’s z transformations to determine if participant training had an effect on the debate surrounding the bipolarity of affect it became apparent that the correlation between PA and NA obtained by the words used by Diener and Emmons (1985) and those used in the

PANAS (Watson, Clark, & Tellegen, 1988) were very different. A Fisher’s z transformation was used in order to determine if the different sets of affect terms were significantly different. Table 9 summarizes these correlations and the level at which they are significantly different. Each descriptive timeframe in both ordinal positions revealed a significant difference between the Diener and Emmons (1985) words and the PANAS (Watson, Clark, & Tellegen, 1988) words.

Table 9. Comparisons of correlations between PA and NA for the words used by Diener and Emmons (1985) and the words used in the PANAS (Watson, Tellegen, & Clark, 1988)

Descriptive Timeframe	Ordinal Position	Diener and Emmons Correlations	PANAS Correlations	Difference
Moment	First	-.51	-.19	-.32***
	Fourth	-.38	-.01	-.37***
Past Month	First	-.30	.02	-.32***
	Fourth	-.40	-.06	-.34***
Past Year	First	-.34	-.13	-.21*
	Fourth	-.24	.07	-.31**
In General	First	-.32	.01	-.33***
	Fourth	-.36	-.08	-.28**

Note: All of the correlations for the words used by Diener and Emmons (1985) are significantly different from the PANAS correlations based on Fisher’s z transformations. * $p < .01$, ** $p < .001$, *** $p < .0001$ level.

DISCUSSION

Participant training’s effect on actual retrospective timeframes

With adequate power, the Mann-Whitney rank-sum test revealed that all 4 analyses resulted in a significant difference in actual retrospective timeframe ratings depending on the position of each timeframe. More specifically, after participant training (i.e., the fourth ordinal position), the terms “moment”, “month”, “year” and “in general” more closely matched the expected timeframe. With the exception of “moment” whose actual retrospective timeframe decreased (as per the hypothesis), the other three descriptive timeframes (i.e., “month”, “year”,

and “in general”) all increased in magnitude from the first to the fourth ordinal position. These findings support the hypothesis that participant training results in actual timeframes more closely matching expected timeframes. The support for this hypothesis speaks to the reliability and validity of mood assessment instruments and mood related research. Studies that essentially employ participant training by having participants rate numerous constructs in succession are unwittingly benefiting by reducing the variability surrounding actual retrospective timeframes which in turn reduces the noise, thereby yielding more accurate results.

Actual retrospective timeframes are a very important component to mood related research. As demonstrated by this study and previous research (Lecci & Wirth, 2006), the amount of variability can be clouding effects that would otherwise surface. It is now important to reevaluate research that may not have gained significant results when involving actual retrospective timeframes. Concern should also be raised when evaluating any research that involves researcher-driven aggregation (e.g., Diener & Emmons, 1985). As mentioned before, researcher-driven aggregation is when the researcher asks participants to make a mood rating over a specific descriptive timeframe and then the researcher does a mathematical summing in order to obtain longer timeframes. Researcher-driven aggregation is very problematic, given that when asked about a descriptive timeframe with no other timeframes to reference, the actual retrospective timeframe does not temporally match the expected timeframe. Prior to this study, it was difficult to evaluate from the literature what results were due to researcher-driven aggregation and which were due to the use of different terms, because Diener and Emmons (1985) not only used different terms, but also employed researcher-driven aggregation. In the present study, it is possible to tease apart what is due to aggregation and what is due to the use of different terms, and clearly the terminology differences are responsible for at least some of the

observed difference. It is therefore critical that mood assessment measures and mood related research standardize the language being used for descriptive timeframes. Given the variability in the results when using identical terms, the use of non-identical (i.e., similar) terms can only create more variability and therefore more contamination of results.

Effects of participant training on variability

The first two hypotheses deal with a decrease in variability after participant training is employed. Due to the distribution of the data, examining the variance does not accurately depict the variability because of outliers and skewness. Therefore, the IQR was used as the measure of variability. The IQR looks at the spread of the middle 50% of the data. In order to test if there is a reduction in variability, the IQR was calculated for every descriptive timeframe in both the first and the fourth position. Using the IQR as a measure of variance, 3 of the 4 descriptive timeframes had a change in variability. “Month” and “year” were predicted to have less of a reduction in variability relative to “moment” and “in general” because the former have clearer expected timeframes. “Month” was the only timeframe to have kept the same IQR and although it was predicted to have less reduction, the fact that there was zero reduction (or increase for that matter) ultimately supports the claim that “month” has a clearer expected timeframe. In this case, participant training did not reduce variability, but a review of Table 6 makes it apparent by looking at the mode that participant training did allow participants to more accurately match their actual retrospective timeframes to the expected timeframes.

Two of the three descriptive timeframes did in fact have a reduction in variability of actual retrospective timeframes after participant training. The variability for “year” reduced after participant training from an IQR of 10,000,000 in the first position to that of zero in the fourth position, which supports the hypothesis stating that there will be a reduction in variability after

participant training, but it does not support the hypothesis saying that “year” will have less reduction than terms such as “moment” and “in general”. Although these results are not exactly congruent with what was predicted, there is a logical, albeit post hoc, explanation. It has been shown in previous literature (i.e., Lecci and Wirth, 2006) and in the current study that participants do not accurately assess descriptive timeframes. Even when the descriptive timeframes have an intuitively clearer expected timeframe. In other words, when using terms such as “month” and “year” there is still a great deal of variability that influences actual retrospective timeframes. As a result, the data may reflect a floor effect in terms of a minimum amount of variability that may be inevitable as individuals have different interpretations of the construct, rather than differences being due largely to error in how the constructs are being assessed.

The difference between how each of these terms (“year” and “month”) responded to participant training could even speak to the way the two terms are defined. Assuming a person can accurately define the terms “year” and “month” there are a number of ways “month” is commonly defined (e.g., 4 weeks, 30 days, 31 days, 28 days). However, when asking about a “year” there is less variation (e.g., 365 days, 52 weeks). A person could also define a year as 12 months, however, the variability in month is taken out of the equation in this research because month was standardized to 30 days. The variability that emerges even when it is assumed that participants know how to define descriptive timeframes “year” and “month” could partially explain why there was a reduction in variability for “year” but not for “month”.

The variability for “moment” was reduced by about half after participant training. Compared to “year” it seems as though this is not a great effect, however, it is unlikely the variability for terms such as “moment” and “in general” will ever reduce the IQR to zero in these

instances due to the ambiguity of the terms. Even if all the participants evaluated the terms accurately (in a temporal sense), there will still be some variability. For example, if a participant rates “moment” as the past 30 seconds most researchers would say that is an accurate evaluation, however a second participant could define “moment” as the past 2 minutes. It is likely that the researcher will also agree that this is an accurate definition. Therefore, there will always be some variability surrounding the actual retrospective timeframe “moment”, likewise with “in general”. In addition, it is important to note that examining the IQR in this instance as an absolute value rather than a proportion to the median creates a problem in terms of comparing the reductions. For example, a similar absolute IQR reduction for “year” will be much larger than that of “month” given that their medians vary.

Finally, participant training actually results in an increase in variability for the term “in general”. Although it wasn’t predicted as such, it is logical that “in general” would increase rather than decrease as the goal is to get the participant to take their entire life into consideration. Therefore, the age range for participants also plays a role in the variability of “in general.” If for instance, a group of 23-25 year olds were surveyed, participant training should change the spread of the data to closely match the age range.

Effects of participant training on the dimensionality of affect

The results of the Fisher’s z transformation did not have a systematic effect on the correlation coefficients between PA and NA. Some of the correlations from the first position were significantly different from the fourth position (see results). However, a consistent effect or trend did not emerge. For example, using the PANAS words, the correlation for “moment” became more independent from the first position to the fourth, while the correlation for “month” went from independent to bipolar from the first to the fourth position.

While running the Fisher's z transformation, it became very apparent that there is a distinct difference in the correlation coefficients obtained by the PANAS words as compared to the Diener and Emmons (1985) words (see table 9). These findings could speak to why there was not a consistent shift in the correlation between PA and NA as a function of the ordinal position. Using the same method and instructional set for both sets of words, the PANAS words seem to be inherently independent, while the Diener and Emmons (1985) are inherently bipolar, regardless of the descriptive timeframe that was used. Specifically, the words used by Diener and Emmons (1985) had a bipolar correlation in the first position and remained that way for the fourth position. Whereas, with the exception of "month" the correlation for the PANAS words started as independent and remained that way from the first to the fourth position. Thus, here again it appears that the over-riding effect is one of language/terminology, where the specific terms employed by the researchers essentially determined the outcome.

One possible explanation for "month" being the only descriptive timeframe that produced a significant shift could be a product of "month" being more specific temporally speaking than some of the other timeframes. However, it seems to follow that if month produced a significant shift from independence to bipolarity then "year" should result in the same effect, although this is not the case. Another explanation is that the effect of words used is too strong to override the current effect under investigation. The results reveal that there is clearly a word effect when examining the dimensionality of affect. The strength of this effect could be more powerful than that of participant training to the point that it suppresses any possible competing effects.

The significant differences in correlations that emerged as a function of the words being used speaks to the long running debate over the dimensionality of affect (e.g., Watson, 1988; Diener & Emmons, 1985; Warr, Barter, & Brownbridge, 1983; Russell & Carroll, 1999, etc.). It

seems apparent that the word differences are a core component of the differing results that continue to prolong the discussion. In other words, the affect terms are determining the correlation between PA and NA, instead of examining the actual relation between PA and NA. If this is the case, the words used in the PANAS and the words used by Diener and Emmons (1985) may be measuring different constructs. Therefore, it is inaccurate to compare studies that use different terms when debating the dimensionality of affect. The current study helps to disentangle the word effect from the method effect. Meaning that it is should now be possible to attribute any differences to construct differences in the words used rather than the results being masked due to actual retrospective timeframes. Any studies attempting to determine the different constructs should take actual retrospective timeframes into account so as not to contaminate the results.

Decision to use the Mann-Whitney Rank-sum test

Due to the shape of the distribution (i.e., non-normal with unequal variance), an important decision had to be made about how to handle the data for analyses. Previous literature on retrospective timeframes obtained similar distributions and were able to apply a transformation that resulted in a normalization of the data (Lecci & Wirth, 2006). In an attempt to employ a similar statistical methodology, a number of transformations were applied to the data but they did not result in a normal distribution.

Another technique that is commonly used in reaction time literature, which deals with similar distributions, is to trim the data (Ratcliff, 1993). This technique was also applied to the current data; however, this method was problematic for two reasons. First, after trimming 10% of the data, the distribution was still non-normal. Second, to maintain adequate power and sample representativeness, it is not ideal to exclude such a large number of subjects (i.e., over 100). It

was also important to avoid imposing arbitrary structure to the data. For example, categories could have been developed in order to facilitate nonparametric analyses. However, this would require categorizing continuous data, which is not ideal because the results would only be meaningful based on categorical decisions that were ultimately subjective.

The final decision was to use a non-parametric test (i.e., the Mann-Whitney) because of its ability to incorporate all of the data while imposing minimal structure (i.e., categories). Although, non-parametric tests, such as, the Mann-Whitney are not commonly used in psychology, it has been shown to have more power than that of the t -test when working with distributions that are non-normal (Blair & Higgins, 1980).

The use of the IQR in order to test a change in variability as a function of ordinal positions speaks to the validity of the study. Comparing results of the Levene's test to Shoemaker's test resulted in most of the findings staying consistent. It is important to point out that by using Shoemaker's approach (1995) 50 percent of the data (25 percent on each side) was cut from the analysis yet the effect was still strong enough to emerge. What is also important is that the current research was examining the variability, which should be very sensitive to eliminating that much data.

Limitations and future research

Taking into account that this is one of the initial studies dealing with actual retrospective timeframes, there are a few procedural changes that could have made the study more methodologically sound. First, the sample was mainly college students receiving extra credit for their participation. Although, Diener and Emmons (1985) found no significant difference between the college aged population and the general public population, it will be important to replicate these findings.

Second, as with Lecci and Wirth (2006), the directions for filling out the questionnaire led to misunderstandings. More specifically, a number of participants had to clarify how to respond when asked “How much of your life did you take into account when making your mood ratings?” These misunderstandings need to be minimized as much as possible in future research. Studies should be conducted that focus on finding an instructional set that is the least ambiguous.

Third, the decision to use the exact wording employed by Lecci and Wirth (2006) when asking about the actual retrospective timeframe could have inadvertently caused similar findings. In other words, it would be important to utilize a different instructional set and see if similar findings materialize. Future research should also attempt to confirm the effects of participant training on actual retrospective timeframes in a less obvious way. For example, one could study the effects of an already validated and reliable mood induction using participant training on the experimental group and comparing it to the control group (i.e., no participant training). This technique may eliminate some of the ambiguity that is involved in explicitly asking participants how much time they took into account. Finally, more research needs to be done on the dimensionality of affect to determine which words are involved in the construct of positive and negative affect and which terms measure a different construct.

Literature Cited

- Beck, A.T. & Steer, R.A. (1987). *Beck depression inventory manual*. San Antonio, TX: The Psychological Corporation.
- Beck, A.T. & Steer, R.A. (1988). *Beck anxiety inventory*. San Antonio, TX: The Psychological Corporation.
- Blair, R. C., & Higgins, J. J. (1980). A comparison of the power of the Wilcoxon's rank-sum test to that of Student's t statistic under various non-normal distributions. *Journal of Educational Statistics*, 5 (4), 309-335.
- Bradburn, N. M. (1969). *The structure of psychological well-being* Chicago: Aldine.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112 (1), 155-159.
- Comrey, A.L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, 56 (5), 754-761.
- Diener, E. & Emmons, R.A. (1985). The independence of positive and negative affect.
- Diener, E. & Iran-Nejad, A. (1986). The relationship in experience between various types of affect. *Journal of Personality and Social Psychology*, 50 (5), 1031-1038.
- Feldman Barrett, L. (1997). The relationship among momentary emotion experiences, personality descriptors, and retrospective ratings of emotion. *Personality and Social Psychology Bulletin*, 23 (10), 1100-1110.
- Green, D.P., Goldman, S.L., & Salovey, P. (1993). Measurement error masks bipolarity in affect ratings. *Journal of Personality and Social Psychology*, 64 (6), 1029-1041.
- Lecci, L. & Wirth, R.J. (2006). Methodological and practical issues in the experience, induction, and assessment of mood states. In A.V. Clark (Ed.) *Psychology of moods: New Research*

- (pp.37-55). New York: Nova Science Publishers, Inc.
- Matell, M.S. & Jacoby, J. (1972). Is there an optimal number of alternatives for likert-scale items: Effects of testing time and scale properties. *Journal of Applied Psychology*, *56* (6), 506-509.
- Rasmussen, P.R., Jeffrey, A.C., Willingham, J.K., & Glover, T.L. (1994). Implications of the true score model in assessment of mood state. *Journal of Social Behavior and Personality*, *9*, 107-118.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510-532.
- Russel, J.A. & Carroll, J.M. (1999). The phoenix of bipolarity: Reply to Watson and Tellegen (1999). *Psychological Bulletin*, *125* (5), 611-617.
- Terry, P.C., Stevens, M.J., & Lane, A.M. (2005). Influence of response time frame on mood assessment. *Anxiety, Stress, and Coping*, *18* (3), 279-285.
- Thomas, D.L. & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, *59* (2), 291-297.
- Warr, P., Barter, J., Brownbridge, G. (1983). On the independence of positive and negative affect. *Journal of Personality and Social Psychology*, *44* (3), 644-651.
- Watson, D. (1988). The vicissitudes of mood measurement: Effects of varying descriptors, time frames, and response formats on measure of positive and negative affect. *Journal of Personality and Social Psychology*, *55* (1), 128-141.
- Watson, D. & Clark, L.A. (1997). Measurement and mismeasurement of mood: Recurrent and emergent issues. *Journal of Personality Assessment*, *68* (2), 267-296.
- Watson, D., Clark, L. A., & Tellegen (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.
- Watson, D. & Tellegen, A. (1999). Issues in the dimensional structure of affect – Effects of descriptors, measurement error, and response formats: Comment on Russell and Carroll (1999). *Psychological Bulletin*, 125 (5), 601-610.
- Winkielman, P., Knauper, B., & Schwarz, N. (1998). Looking back at anger: Reference periods change the interpretation of emotion frequency questions. *Journal of Personality and Social Psychology*, 75 (3), 719-728.
- Zuckerman & Lubin (1965). *Manual for the MAACL-R: The multiple affect adjective check list revised*. San Diego, CA: EdITS.

Appendix A. Demographic page

Please fill out the following information about yourself.

_____ Age Gender: Male Female

Race: _____

Are you currently a psychology major?: Yes No

In the following pages you will be asked to make ratings about your mood. **Please pay close attention to the timeframes in bold!**

Appendix B

Questionnaire 1

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past month**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>
___ interested		___ irritable		
___ distressed		___ alert		
___ excited		___ ashamed		
___ upset		___ inspired		
___ strong		___ nervous		
___ guilty		___ determined		
___ scared		___ attentive		
___ hostile		___ jittery		
___ enthusiastic		___ active		
___ proud		___ afraid		
___ delighted		___ miserable		
___ annoyed		___ happy		
___ frustrated		___ content		
___ glad		___ gloomy		
___ satisfied		___ sad		
___ depressed		___ pleased		

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way **in general**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>
<input type="checkbox"/> interested		<input type="checkbox"/> irritable		
<input type="checkbox"/> distressed		<input type="checkbox"/> alert		
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed		
<input type="checkbox"/> upset		<input type="checkbox"/> inspired		
<input type="checkbox"/> strong		<input type="checkbox"/> nervous		
<input type="checkbox"/> guilty		<input type="checkbox"/> determined		
<input type="checkbox"/> scared		<input type="checkbox"/> attentive		
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery		
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active		
<input type="checkbox"/> proud		<input type="checkbox"/> afraid		
<input type="checkbox"/> delighted		<input type="checkbox"/> miserable		
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy		
<input type="checkbox"/> frustrated		<input type="checkbox"/> content		
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy		
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad		
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased		

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past year**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested		<input type="checkbox"/> irritable
<input type="checkbox"/> distressed		<input type="checkbox"/> alert
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed
<input type="checkbox"/> upset		<input type="checkbox"/> inspired
<input type="checkbox"/> strong		<input type="checkbox"/> nervous
<input type="checkbox"/> guilty		<input type="checkbox"/> determined
<input type="checkbox"/> scared		<input type="checkbox"/> attentive
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active
<input type="checkbox"/> proud		<input type="checkbox"/> afraid

<input type="checkbox"/> delighted		<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy
<input type="checkbox"/> frustrated		<input type="checkbox"/> content
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you feel **at this moment**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested	<input type="checkbox"/> irritable
<input type="checkbox"/> distressed	<input type="checkbox"/> alert
<input type="checkbox"/> excited	<input type="checkbox"/> ashamed
<input type="checkbox"/> upset	<input type="checkbox"/> inspired
<input type="checkbox"/> strong	<input type="checkbox"/> nervous
<input type="checkbox"/> guilty	<input type="checkbox"/> determined
<input type="checkbox"/> scared	<input type="checkbox"/> attentive
<input type="checkbox"/> hostile	<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic	<input type="checkbox"/> active
<input type="checkbox"/> proud	<input type="checkbox"/> afraid

<input type="checkbox"/> delighted	<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed	<input type="checkbox"/> happy
<input type="checkbox"/> frustrated	<input type="checkbox"/> content
<input type="checkbox"/> glad	<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied	<input type="checkbox"/> sad
<input type="checkbox"/> depressed	<input type="checkbox"/> pleased

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

Appendix C

Questionnaire 2

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past year**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>
___ interested				___ irritable
___ distressed				___ alert
___ excited				___ ashamed
___ upset				___ inspired
___ strong				___ nervous
___ guilty				___ determined
___ scared				___ attentive
___ hostile				___ jittery
___ enthusiastic				___ active
___ proud				___ afraid
___ delighted				___ miserable
___ annoyed				___ happy
___ frustrated				___ content
___ glad				___ gloomy
___ satisfied				___ sad
___ depressed				___ pleased

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past month**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

___ interested

___ irritable

___ distressed

___ alert

___ excited

___ ashamed

___ upset

___ inspired

___ strong

___ nervous

___ guilty

___ determined

___ scared

___ attentive

___ hostile

___ jittery

___ enthusiastic

___ active

___ proud

___ afraid

___ delighted

___ miserable

___ annoyed

___ happy

___ frustrated

___ content

___ glad

___ gloomy

___ satisfied

___ sad

___ depressed

___ pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you feel **at this moment**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested	<input type="checkbox"/> irritable
<input type="checkbox"/> distressed	<input type="checkbox"/> alert
<input type="checkbox"/> excited	<input type="checkbox"/> ashamed
<input type="checkbox"/> upset	<input type="checkbox"/> inspired
<input type="checkbox"/> strong	<input type="checkbox"/> nervous
<input type="checkbox"/> guilty	<input type="checkbox"/> determined
<input type="checkbox"/> scared	<input type="checkbox"/> attentive
<input type="checkbox"/> hostile	<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic	<input type="checkbox"/> active
<input type="checkbox"/> proud	<input type="checkbox"/> afraid

<input type="checkbox"/> delighted	<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed	<input type="checkbox"/> happy
<input type="checkbox"/> frustrated	<input type="checkbox"/> content
<input type="checkbox"/> glad	<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied	<input type="checkbox"/> sad
<input type="checkbox"/> depressed	<input type="checkbox"/> pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way **in general**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested	<input type="checkbox"/> irritable
<input type="checkbox"/> distressed	<input type="checkbox"/> alert
<input type="checkbox"/> excited	<input type="checkbox"/> ashamed
<input type="checkbox"/> upset	<input type="checkbox"/> inspired
<input type="checkbox"/> strong	<input type="checkbox"/> nervous
<input type="checkbox"/> guilty	<input type="checkbox"/> determined
<input type="checkbox"/> scared	<input type="checkbox"/> attentive
<input type="checkbox"/> hostile	<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic	<input type="checkbox"/> active
<input type="checkbox"/> proud	<input type="checkbox"/> afraid

<input type="checkbox"/> delighted	<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed	<input type="checkbox"/> happy
<input type="checkbox"/> frustrated	<input type="checkbox"/> content
<input type="checkbox"/> glad	<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied	<input type="checkbox"/> sad
<input type="checkbox"/> depressed	<input type="checkbox"/> pleased

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

Appendix D
Questionnaire 3

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way **in general**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>
<input type="checkbox"/> interested		<input type="checkbox"/> irritable		
<input type="checkbox"/> distressed		<input type="checkbox"/> alert		
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed		
<input type="checkbox"/> upset		<input type="checkbox"/> inspired		
<input type="checkbox"/> strong		<input type="checkbox"/> nervous		
<input type="checkbox"/> guilty		<input type="checkbox"/> determined		
<input type="checkbox"/> scared		<input type="checkbox"/> attentive		
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery		
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active		
<input type="checkbox"/> proud		<input type="checkbox"/> afraid		
<input type="checkbox"/> delighted		<input type="checkbox"/> miserable		
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy		
<input type="checkbox"/> frustrated		<input type="checkbox"/> content		
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy		
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad		
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased		

How much of your life did you take into consideration when making your mood ratings?
 Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you feel **at this moment**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested	<input type="checkbox"/> irritable
<input type="checkbox"/> distressed	<input type="checkbox"/> alert
<input type="checkbox"/> excited	<input type="checkbox"/> ashamed
<input type="checkbox"/> upset	<input type="checkbox"/> inspired
<input type="checkbox"/> strong	<input type="checkbox"/> nervous
<input type="checkbox"/> guilty	<input type="checkbox"/> determined
<input type="checkbox"/> scared	<input type="checkbox"/> attentive
<input type="checkbox"/> hostile	<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic	<input type="checkbox"/> active
<input type="checkbox"/> proud	<input type="checkbox"/> afraid

<input type="checkbox"/> delighted	<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed	<input type="checkbox"/> happy
<input type="checkbox"/> frustrated	<input type="checkbox"/> content
<input type="checkbox"/> glad	<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied	<input type="checkbox"/> sad
<input type="checkbox"/> depressed	<input type="checkbox"/> pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past month**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested		<input type="checkbox"/> irritable
<input type="checkbox"/> distressed		<input type="checkbox"/> alert
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed
<input type="checkbox"/> upset		<input type="checkbox"/> inspired
<input type="checkbox"/> strong		<input type="checkbox"/> nervous
<input type="checkbox"/> guilty		<input type="checkbox"/> determined
<input type="checkbox"/> scared		<input type="checkbox"/> attentive
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active
<input type="checkbox"/> proud		<input type="checkbox"/> afraid

<input type="checkbox"/> delighted		<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy
<input type="checkbox"/> frustrated		<input type="checkbox"/> content
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past year**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested	<input type="checkbox"/> irritable
<input type="checkbox"/> distressed	<input type="checkbox"/> alert
<input type="checkbox"/> excited	<input type="checkbox"/> ashamed
<input type="checkbox"/> upset	<input type="checkbox"/> inspired
<input type="checkbox"/> strong	<input type="checkbox"/> nervous
<input type="checkbox"/> guilty	<input type="checkbox"/> determined
<input type="checkbox"/> scared	<input type="checkbox"/> attentive
<input type="checkbox"/> hostile	<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic	<input type="checkbox"/> active
<input type="checkbox"/> proud	<input type="checkbox"/> afraid

<input type="checkbox"/> delighted	<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed	<input type="checkbox"/> happy
<input type="checkbox"/> frustrated	<input type="checkbox"/> content
<input type="checkbox"/> glad	<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied	<input type="checkbox"/> sad
<input type="checkbox"/> depressed	<input type="checkbox"/> pleased

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

Appendix E

Questionnaire 4

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you feel **at this moment**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>
<input type="checkbox"/> interested		<input type="checkbox"/> irritable		
<input type="checkbox"/> distressed		<input type="checkbox"/> alert		
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed		
<input type="checkbox"/> upset		<input type="checkbox"/> inspired		
<input type="checkbox"/> strong		<input type="checkbox"/> nervous		
<input type="checkbox"/> guilty		<input type="checkbox"/> determined		
<input type="checkbox"/> scared		<input type="checkbox"/> attentive		
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery		
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active		
<input type="checkbox"/> proud		<input type="checkbox"/> afraid		
<input type="checkbox"/> delighted		<input type="checkbox"/> miserable		
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy		
<input type="checkbox"/> frustrated		<input type="checkbox"/> content		
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy		
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad		
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased		

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past year**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested		<input type="checkbox"/> irritable
<input type="checkbox"/> distressed		<input type="checkbox"/> alert
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed
<input type="checkbox"/> upset		<input type="checkbox"/> inspired
<input type="checkbox"/> strong		<input type="checkbox"/> nervous
<input type="checkbox"/> guilty		<input type="checkbox"/> determined
<input type="checkbox"/> scared		<input type="checkbox"/> attentive
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active
<input type="checkbox"/> proud		<input type="checkbox"/> afraid

<input type="checkbox"/> delighted		<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy
<input type="checkbox"/> frustrated		<input type="checkbox"/> content
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way **in general**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested		<input type="checkbox"/> irritable
<input type="checkbox"/> distressed		<input type="checkbox"/> alert
<input type="checkbox"/> excited		<input type="checkbox"/> ashamed
<input type="checkbox"/> upset		<input type="checkbox"/> inspired
<input type="checkbox"/> strong		<input type="checkbox"/> nervous
<input type="checkbox"/> guilty		<input type="checkbox"/> determined
<input type="checkbox"/> scared		<input type="checkbox"/> attentive
<input type="checkbox"/> hostile		<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic		<input type="checkbox"/> active
<input type="checkbox"/> proud		<input type="checkbox"/> afraid

<input type="checkbox"/> delighted		<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed		<input type="checkbox"/> happy
<input type="checkbox"/> frustrated		<input type="checkbox"/> content
<input type="checkbox"/> glad		<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied		<input type="checkbox"/> sad
<input type="checkbox"/> depressed		<input type="checkbox"/> pleased

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Use the following scale to record your answers. Indicate to what extent you felt this way over the **past month**.

1	2	3	4	5
<i>Very slightly or not at all</i>	<i>a little</i>	<i>moderately</i>	<i>quite a bit</i>	<i>extremely</i>

<input type="checkbox"/> interested	<input type="checkbox"/> irritable
<input type="checkbox"/> distressed	<input type="checkbox"/> alert
<input type="checkbox"/> excited	<input type="checkbox"/> ashamed
<input type="checkbox"/> upset	<input type="checkbox"/> inspired
<input type="checkbox"/> strong	<input type="checkbox"/> nervous
<input type="checkbox"/> guilty	<input type="checkbox"/> determined
<input type="checkbox"/> scared	<input type="checkbox"/> attentive
<input type="checkbox"/> hostile	<input type="checkbox"/> jittery
<input type="checkbox"/> enthusiastic	<input type="checkbox"/> active
<input type="checkbox"/> proud	<input type="checkbox"/> afraid

<input type="checkbox"/> delighted	<input type="checkbox"/> miserable
<input type="checkbox"/> annoyed	<input type="checkbox"/> happy
<input type="checkbox"/> frustrated	<input type="checkbox"/> content
<input type="checkbox"/> glad	<input type="checkbox"/> gloomy
<input type="checkbox"/> satisfied	<input type="checkbox"/> sad
<input type="checkbox"/> depressed	<input type="checkbox"/> pleased

How much of your life did you take into consideration when making your mood ratings? Please indicate a numerical value (e.g., 1, 2, 3, etc.) in the space provided and circle the appropriate timeframe.

_____ second(s) minute(s) hour(s) day(s) week(s) month(s) year(s)