COGNITION AND EMOTION, 2001, *15* (3), 333–363

Mental representations of affect knowledge

Lisa Feldman Barrett
Boston College, USA

Thyra Fossum
The Pennsylvania State University, USA

The circumplex structure derived from similarity ratings of affect words is assumed to be a conceptual representation of affect anchored in semantic knowledge. Recently, it been suggested that this structure is not based on semantic knowledge at all, but may instead reflect a type of episodic knowledge: The degree to which emotions covary in everyday life. In two experience-sampling studies, we compared the semantic and the episodic hypotheses by comparing participants' similarity ratings to the observed covariations in their own affective experience computed from their momentary reports. In Study 2, participants also provided estimates of the degree to which their emotions covaried. Evidence from both studies indicate that similarity judgements are related both to semantic and episodic information, indicating that a pure episodic account of similarity ratings, and the mental representation of affect that they reflect, is untenable.

Much of the psychological knowledge that we have generated is based on what people tell us in response to the questions we ask. No where is this more true than in research on emotion. We ask participants to judge the similarity of emotion words, the emotional content contained in facial expressions, pictures, or movies, or their own feeling states. Based on what participants can tell us, we infer something about the nature of emotion knowledge, emotional experience, or both, depending on what we think the questions assess. This last inferential step is often where the process of knowledge accumulation breaks down, because researchers do not always agree about which information participants rely upon when making judgements of emotion. The purpose of this paper is to

Correspondence should be addressed to Lisa Feldman Barrett at the Department of Psychology, 427 McGuinn Building, Boston College, Chestnut Hill MA 02167, USA; barretli@bc.edu

Many thanks to Michael Ross, Jim Russell, Paula Niedenthal, Ulrich Schimmack, and Rainer Reisenzein for their comments on early drafts of this manuscript. Also thanks to Daniel Schacter for his input on the semantic-episodic knowledge distinction. Preparation of this article was facilitated by a grant from the National Science Foundation to Lisa Feldman Barrett (SBR 9727896).

http://www.tandf.co.uk/journals/pp/02699931.html DOI:10.1080/02699930042000196

address one such disagreement regarding similarity ratings and the representation of emotion that they index.

The affective circumplex

A large body of research has accumulated to suggest that people hold a general mental representation of affect in the form of a circular structure, or circumplex. In abstract terms, a circumplex is an empirically derived dimensional structure that represents the conceptual, mental structure of a group of stimuli (Guttman, 1954, 1966). A circumplex structure is typically derived from a dimensional analysis (e.g., factor analysis or multidimensional scaling) of proximity ratings (e.g., similarity ratings) for a set of stimuli (e.g., affect terms). The proximity judgements are typically taken to be the index of this mental structure (e.g., Shepard, 1962, 1974, 1980; Tversky, 1977), such that the dimensions resulting from their analysis are thought to represent the underlying attributes or properties of that mental structure.

Dimensional analyses of similarity ratings indicate that valence and arousal serve as the primary dimensions of the affect circumplex (Russell & Feldman Barrett, 1999). Valence refers to the hedonic quality or pleasantness of an affective experience, and arousal refers to subjective feelings of activation associated with an affective experience. All affective stimuli (i.e., emotions such as anger, sadness, and fear, as well as nonemotional affective states like fatigue, sleepiness, and placidity) can be defined as combinations of these two independent dimensions. Emotion terms array along valence and arousal dimensions because these two properties represent core affective features of emotion concepts. In addition to these basic components, emotion representations contain other elements of emotion knowledge that differentiate discrete emotions like fear, anger, and sadness. Thus, the valence and arousal dimensions represent necessary, but not sufficient, components of emotion concepts. Additional knowledge would, for example, differentiate among discrete emotion concepts like fear and anger. In essence, the valence/arousal circumplex is like a very primitive representation or cognitive map of affective space.

This cognitive map is considered essentially nomothetic (i.e., it is thought to apply equally to all persons within a culture). It is highly robust and emerges whenever individuals label or communicate their own or others' affective experiences. The valence/arousal circumplex structure has been identified in proximity ratings of emotion terms (e.g., Block, 1957; Bush, 1973; Feldman, 1995a, Russell, 1980), in perceptions of facially expressed emotion (e.g., Abelson & Sermat, 1962; Adolphs, Tranel, Damasio, & Damasio, 1994; Cliff & Young, 1968; Dittmann, 1972; Fillenbaum & Rapaport, 1971; Green & Cliff, 1975; Russell & Bullock, 1985; Russell, Lewicka, & Niit, 1989; Schlosberg, 1952, 1954; Shepard, 1962), as well as in self- reported affective states

(Almagor & Ben-Porath, 1989; Feldman, 1995a; Feldman Barrett, 1998; Feldman Barrett, & Russell, 1998; Reisenzein, 1994; Russell, 1980).

The assumption, until recently, has been that the mental representation of affect represented by similarity ratings and depicted by the circumplex is anchored by information stored in semantic memory. On this view, valence and arousal reflect semantic components of emotion concepts. Semantic memory refers to a corpus of impersonal, conceptual knowledge shared by members of the same culture (Tulving, 1972). This view of the circumplex derives from three sources. First, the circumplex is consistent with the semantic differential work by Osgood (Osgood, Suci, & Tannenbaum, 1957) who demonstrated that there are three major components of meaning in natural language (evaluation, activity, and potency). Second, evidence indicates that direct semantic ratings of affect terms are highly correlated (*r*s greater than .93) with dimension coordinates (Feldman, 1995a). Third, valence/arousal circumplexes derived from similarity ratings of affective language are highly replicable across individuals, word sets, and have been found across many cultures (Russell, 1983, 1991; Russell et al., 1989), suggesting that valence and arousal represent all-purpose semantic knowledge that individuals use when generating some judgement about an affective stimulus.

Recently, Schimmack, and Reisenzein (1997) have challenged the idea that similarity ratings of affect words, and the mental representation that they index, are anchored by people's abstract, semantic knowledge (also see Conway & Bekerian, 1987). Instead, they suggest that the similarity ratings reflect a type of knowledge stored in episodic memory: The degree to which emotions covary in a person's everyday life. Episodic memory is defined as past events that are definable with respect to time and place (Tulving, 1972). They are typically, but not exclusively personal in nature (Maguire & Mummery, 1999; Nadel & Moscovitch, 1998). Unlike semantic knowledge which represents conceptual information about things, episodic knowledge is about specific events (Nadel & Moscovitch, 1998) that typically require conscious recollection (Tulving, 1985). According to the episodic hypothesis, similarity judgements are computed, not by accessing abstract semantic information, but by comparing exemplar-based representations of how often affective experiences occur together. From this perspective, respondents rate the similarity of two emotions on the basis of their beliefs about how often the two were actually experienced together. Affective experiences that co-occur frequently will be judged as similar, whereas those that co-occur infrequently will be judged as dissimilar. In this scenario, the valence and arousal dimensions that derive from analyses of similarity ratings represent summaries of how often emotional states co-occur, rather than semantic information *per se*. Although Schimmack and Reisenzein (1997) allow for the existence of semantic representations of emotion concepts, they argue that similarity judgements are not based on such knowledge. On the episodic view, then, the valence and arousal dimensions that derive from dimensional analyses of similarity ratings reflect the co-occurrences of emotional states.

The issue of whether similarity ratings are semantically or episodically derived is crucial to understanding the role of the circumplex in emotion theory. The original theory associated with the circumplex is that it represents general, conceptual knowledge about affect used to generate a conscious representation of feeling states, be they affective or emotional (Russell & Feldman Barrett, 1999). In contrast, an episodic view suggests that the mental representation is the result of conscious representations of feelings states. Furthermore, on the semantic view, it has been suggested that variations in the way that semantic knowledge is applied in the generation process result in individual differences in the co-occurrence of conscious emotional states (Feldman, 1995a; Feldman Barrett, 1998). That is, the co-occurrence is produced by the way in which the mental representation is accessed and applied, resulting in episodic experiences that differ across individuals. In contrast, an episodic view suggests that the mental representation itself reflects the specific co-occurrences of emotion that are stored in episodic memory for a particular person.

An episodic interpretation of circumplex structure?

For an episodic hypothesis to be tenable, accurate episodic information must be contained in the circumplex representation. Similarity ratings must be based, at least in part, on actual experienced emotion covariations, and not some general theory of co-occurrence. According to Reisenzein and Schimmack (1999), similarity ratings are computed based on respondents' experience of which emotions tend to co-occur (or closely follow one another in time). Specifically, they suggest that participants implicitly generate covariation estimates when judging the similarity between emotions, such that these estimates should correspond closely to the correlation between emotion experienced during emotion episodes. As a result, the affect covariation information represented in episodic memory that is used to generate similarity judgements ''should be a fairly accurate reflection of the actual co-occurances of affects'' that people encounter in their lives (p. 552). Indeed, evidence suggests that episodic representations like emotion covariation estimates are, at least in part, constructed from the information provided in context-specific, everyday experience (McKenzie, 1994; for reviews see Alloy & Tabachnik, 1984; Crocker, 1981). The purpose of the present article was to examine the degree to which mental representations of emotion contain actual covariation information (i.e., the degree to which emotions covaried in everyday life).

OVERVIEW OF THE PRESENT STUDIES

This article contains two studies that assessed the relationship between proximity judgements of affect terms and actual, observed covariations in experienced affective states. Using an experience-sampling method, participants in two studies made ratings of their affective state three times a day over an extended observation period. These data allowed us to compute observed co-occurrences in everyday experience. Observed co-occurrences were then compared to participants' proximity ratings of emotion. Two types of proximity ratings were examined.

First, we compared observed emotion covariations with similarity judgements (Studies 1 and 2). To date, no study has empirically compared covariations in idiographically derived emotion ratings from real life experience with similarity judgements. The two studies that have examined the correspondence idiographically either did not report a direct empirical comparison of the two (Zevon & Tellegen, 1982), or had participants rate how they believed they would feel in a series of emotion eliciting situations (Reisenzein & Schimmack, 1999, Study 4). Such hypothetical responses tap propositional knowledge that may or may not be an accurate index of actual emotion covariation in daily life (for a discussion of bias in hypothetical responses, see Greenwald & Banaji, 1995; Ross, 1989).

Second, we compared observed emotion covariations with emotion covariation estimates (in the form of conditional probability judgements; Study 2 only). According to previous research, conditional probability judgements are constructed, in part, from the information provided in context-specific, everyday experience (McKenzie, 1994; for reviews see Alloy & Tabachnik, 1984; Crocker, 1981). If participants generate accurate covariation estimates, as suggested by Reisenzein and Schimmack (1999), then there should be a strong correspondence between their observed emotion covariations and their conditional probability judgements (i.e., how often they believe two emotions co-occur in their experience). To date, no study has empirically compared covariations in idiographically derived emotion ratings from real life experience with estimated covariation judgements. One study has investigated the nomothetic correspondence in cross-sectionally derived ratings (Schimmack & Hartmann, 1997), and a second study compared conditional probability judgements to idiographically derived responses to hypothetical situations (Reisenzein & Schimmack, 1999, Study 4).

STUDY 1: RE-ANALYSIS OF FELDMAN (1995A)

Study 1 examined whether accurate covariation information is contained in similarity judgements by reanalysing a dataset containing experience-sampling ratings of 16 affective experiences (Feldman, 1995a). In this study, participants provide detailed, quantitative descriptions of their affective experiences as soon as they occurred across a 90-day period, allowing an estimate of observed emotion co-occurrence. Our aim was to determine whether the covariations that we observed in participants' affect ratings were related to their similarity ratings of those same affect terms. If similarity judgements of emotion words reflect episodic knowledge of emotion covariation, then similarity judgements should strongly correspond to the observed co-occurrence of affective experiences.

Methods

At the beginning of the study, participants provided similarity ratings of 16 affect terms. The 16 terms were chosen to represent the circumplex space to ensure that all of its octants were equally represented (*enthusiastic*, *peppy*, *happy*, *satisfied*, *calm*, *relaxed*, *quiet*, *still*, *sleepy*, *sluggish*, *sad*, *disappointed*, *nervous*, *afraid*, *surprised*, *aroused*). Each term served either as the referent or as the subject in a given pairing, resulting in 120 judgements. Participants were asked to rate the similarity of the word pairs according to the meanings of the words (1 = extremely dissimilar, through 4 = unrelated, 7 = extremely similar). Terms appeared equally spaced throughout the measure (see Davison, 1983). The adjective pairs were presented in a single random order.

Participants then rated their momentary affective state using a series of 88 terms that included those target 16 terms. Participants indicated on a 7-point scale the extent to which each of these 88 mood adjectives described their emotional state (0 = not at all, 3 = a moderate amount, 6 = a great deal) by responding to the following instruction: ''Indicate to what extent you feel this way right now, that is, at the present moment.'' Participants completed a questionnaire in the morning (7 a.m. to 12 p.m.), afternoon (12 p.m. to 5 p.m.), and evening (5 p.m. to 12 a.m.) every day for 90 consecutive days. They returned completed forms on Monday, Wednesday, and Friday of each week. The experimenters contacted participants within 48 hours if they failed to return a questionnaire, and they interviewed participants three times during the study to ensure compliance with the research procedures. Full details about the procedure can be found in Feldman (1995a). Data were available for the 16 participants (of the 24 in the original study) who completed both the similarity judgements and the momentary emotion ratings.

Results

Intercorrelations between the 16 circumplex markers across the period of observation were calculated for each participant. These 120 intercorrelations represented the measured covariation between self-reported affective states for each participant.[1] Each participant's observed covariations were compared to his/her similarity judgements across the 120 word pairs. One correlation coefficient was generated for each participant. The resulting correlations ranged

[1] All correlation coefficients were subjected to a Fisher's r-to-z transformation before being used in additional analyses.

from $r = .31$ to $r = .72$, with a mean of $r = .55$ and a standard deviation of 0.17. This finding has two interpretations. It may indicate the correlation was .55 for every participant and that there was substantial sampling error in estimating the correlation. Alternatively, it may indicate that on average, experienced covariations were related to the similarity ratings, but the true strength of this relationship was stronger for some participants than for others. Although sampling error surely existed in this dataset, we suspect that the latter possibility is more likely.

To demonstrate that true strength of the similarity-observed covariation relationship did not fluctuate randomly but was systematically stronger from some individuals than for others, we conducted three additional analyses. First, we demonstrated that there were large individual differences in emotion covariation, yet respondents' similarity ratings strongly agreed with one another. An affective structure was constructed for each participant by submitting his/her observed covariations to a principle axis factor analysis, followed by plotting of the factor loadings (Feldman (1995a). The resulting structures varied from circular to elliptical. Two of the affective structures reported in Feldman (1995a) are reproduced in Figure 1a to illustrate this variation. Such variation indicates that the pattern of emotion covariation differed across individuals (for a full explanation, see Feldman, 1995a). The affective circumplex, derived from an individual differences multidimensional scaling (MDS) of the similarity ratings (using INDSCAL), presented in Figure 1b, was a circular structure that was highly consistent across participants. The small variations in the MDS solution across participants were not significantly related to variations in observed emotion covariations (Feldman, 1995a). Thus, the results suggested that the similarity ratings produced a circular structure with very little variation across participants, but the covariations in the self-report data produced structures of varying shape across participants.

Second, we investigated whether the strength of the similarity-observed covariation relationship changed systematically with the pattern of observed covariations in experienced affect. Specifically, we hypothesised that as the observed covariations produced a self-report structure approaching a circular shape, the more the strength of the relationship between the observed covariations and the similarity ratings would increase. A circularity index for each participant's self-report structure was established by computing a ratio of the percentage of variance accounted for by the two factors anchoring each structure. The higher the index, the more circular the structure; a value of one indicated a perfectly circular structure. The lower the index, the more elliptical the structure. This index was then compared to the similarity-covariation correlations. As predicted, the magnitude of the similarity-covariation correspondence was influenced by the degree to which respondents' self-report structure were circular, ($r = .42$, $p < .05$). Thus, the correspondence between

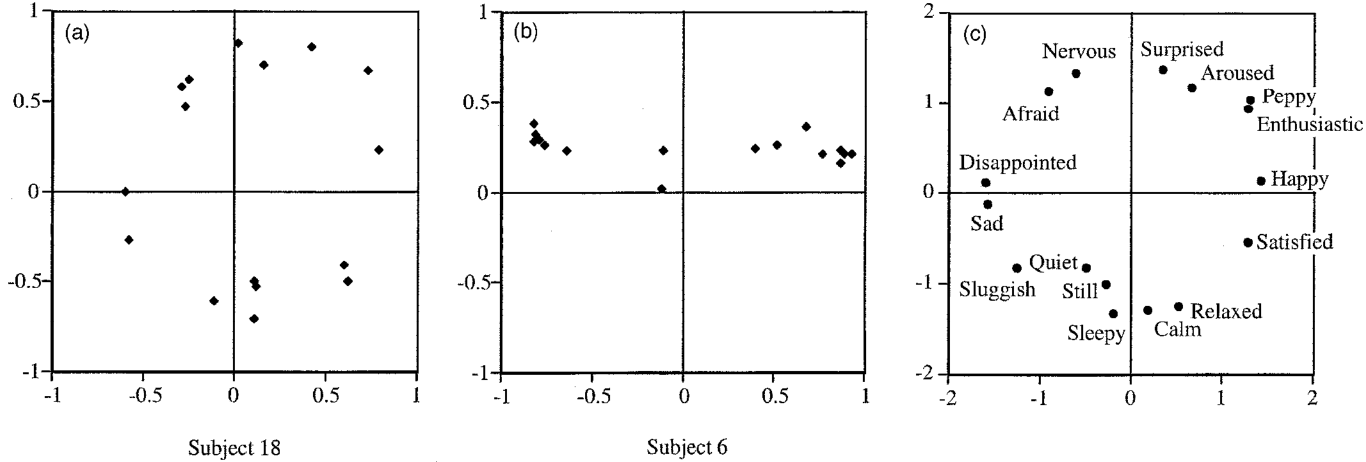

Figure 1. (a) and (b) Variations in emotion structure. Figure modified from Feldman (1995a). (c) The circumplex structure of affect derived from similarity judgements. Valence is the horizontal axis and arousal is the vertical axis. Figure taken from Feldman Barrett (1998).

participants' observed emotion covariations and their similarity ratings increased as the self-report structures came to resemble the circular structure of the similarity ratings.

Finally, we directly tested whether the similarity of two emotions would increase with the number of episodes in which they co-occurred, and decrease with the number of episodes in which each emotion occurred alone. The average correlations between pairs of pleasant (*enthusiastic*, *happy*, and *calm*) and unpleasant (*nervous*, *sad*, and *sluggish*) affects were calculated across participants. Participants were divided into groups on the basis of whether their pairwise correlation fell above or below the median for that pair. This procedure was performed separately for each affect pair. *t*-tests were then calculated to determine whether participants with a large covariation between two emotions judged those states to be more similar than did those with a small covariation. The results, presented in Table 1, do not support the hypothesis that the covariation between two affective experiences influenced their judged similarity. The analyses were replicated with another set of pleasant (*peppy*, *satisfied*, and *relaxed*) and unpleasant (*afraid*, *disappointed*, *sleepy*) emotions: participants with large covariations did not differ from those with small covariations in their similarity judgments of the pairs of emotion words.

TABLE 1
Judged similarity of emotion pairs by participants high or low in actual emotion covariation

	Amount of covariation			
	Small	*Large*	*t*	*p*
First set of emotion pairs				
Enthusiastic/happy	6.6	5.5	1.59	n.s.
Enthusiastic/calm	1.9	3.0	1.84	0.09
Calm/happy	5.0	4.8	0.39	n.s.
Nervous/sad	4.6	3.4	1.71	0.10
Sluggish/nervous	3.1	2.8	0.49	n.s.
Sluggish/sad	5.6	5.0	0.71	n.s.
Second set of emotion pairs				
Satisfied/peppy	4.9	5.4	0.75	n.s.
Peppy/relaxed	1.9	2.3	0.62	n.s.
Satisfied/relaxed	5.6	5.9	0.72	n.s.
Afraid/disappointed	4.1	2.66	1.74	0.10
Sleepy/afraid	2.4	2.4	0.00	n.s.
Sleepy/disappointed	5.4	5.1	0.33	n.s.

Note: $df = 14$. Individuals were divided into groups using a median split procedure.

Discussion

Taken together, the results of Study 1 cast doubt on the hypothesis that similarity ratings are computed based on respondents' experience of which emotions tend to co-occur. First, individuals varied from one another in the co-occurrence of their affective experiences, but did not vary substantially in their similarity ratings. Of course it is possible that the variation observed in the co-occurrences are due to random error, but additional evidence suggests that they are meaningfully related to external variables, and therefore cannot be considered random fluctuations (Feldman Barrett & Gross, in press). Second, judgements of emotion similarity were not uniformly related to observed co-occurrences in experienced affect. This relationship was stronger for individuals who displayed a circular affective structure, and weaker for those with a more elliptical structure. Finally, individuals who displayed large correlations between specific pairs of emotions did not differ in their similarity judgements of those emotion pairs from individuals who displayed smaller pairwise correlations.

STUDY 2

Study 2 replicated and extended the findings from Study 1. Once again, an idiographic method was utilised to examine the degree to which mental representations of emotion contained accurate episodic information. In Study 2, mental representations were indexed both by similarity ratings and by conditional probability ratings of affect word pairs. Observed co-occurrences (derived from experience-sampling ratings of emotion three times per day over 60 days) were then compared to emotion covariation estimates (in the form of conditional probability judgements) and to similarity ratings of affect terms, taken both before and after the experience-sampling procedure. Using these data, we examined three hypotheses.

First, we sought to determine whether the pre-sampling proximity ratings (i.e., the similarity and conditional probability ratings) were related to the covariance information (the predictive correspondence hypothesis). Second, because our momentary rating procedure may have cued individuals to recall their actual experience more accurately (Linton, 1986), we investigated whether people would adjust or calibrate their proximity ratings to their actual affective experiences after reporting on their experience over a two month period (the calibration hypothesis). Third, we explored the possibility that both types of proximity ratings contained semantic knowledge (the semantic hypothesis).

The predictive correspondence, calibration, and semantic hypotheses were examined using path analyses (Pedhauzer, 1982). In traditional path analysis, the zero-order correlations between variables are decomposed in the manner specified by the hypotheses being tested and path coefficents are estimated using standardised regression weights obtained from a series of regression analyses. One path model was estimated for the similarity ratings, and another for the

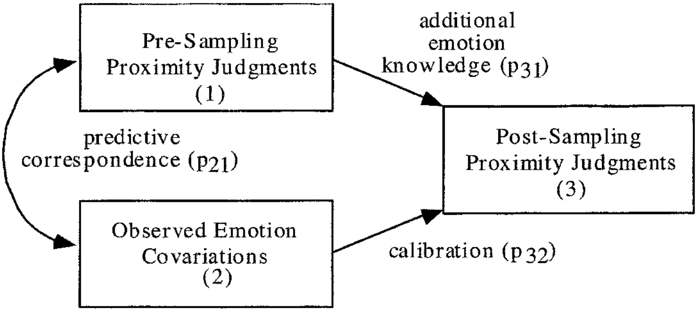

Figure 2. Framework for estimating the predictive accuracy and calibration hypotheses.

conditional probability judgements. Figure 2 portrays the ways in which the observed emotion covariations were related to proximity (i.e., similarity or conditional probability) judgements taken both before and after the experience-sampling portion of the study. Path p_{21} represented *the predictive correspondence effect* (i.e., the degree to which accurate covariance information and in pre-experience-sampling ratings were related); this path reflected the assumption that the profile of emotion covariations is stable for one individual across time. Analyses of two experience-sampling datasets (Feldman, 1995a; Feldman Barrett, 1998) support this assumption.

Path p_{32} represented *the calibration effect* (i.e., the degree to which individuals' calibrated their judgments to the co-occurrence of affective experiences that they documented during the 60 days of experience-sampling procedure); this path reflected the degree to which similarity or conditional probability ratings were sensitive to actual covariation experiences, over and above the influence of the pre-sampling ratings.

Path p_{31} represented *the additional emotion knowledge effect* (i.e., the degree to which the pre- and post-sampling ratings shared variance that was not an accurate reflection of the observed covariations). The direct effect of the pre-sampling similarity or conditional probability ratings on the post-sampling ratings, once the influence of the observed covariances was controlled, was evidence that additional emotion knowledge (whether rooted in semantic information, in erroneous beliefs about covariation, or in some other source of information) was contained in the proximity ratings. That is, the path represents the emotion knowledge not based on immediate experiences of emotion covariations that participants are bringing to bear when making similarity ratings, or when making estimates of emotion covariation.

We explored the extent to which this additional knowledge contained semantic information. According to a large body of previous research, affect terms can be defined on the basis of the valence and the level of arousal that they denote (for a review, see Russell, 1980). Although valence and arousal do not constitute all the semantic information associated with affect terms, they make up a large proportion of it (see Russell & Feldman Barrett, 1999). To test whether the additional knowledge contained any semantic information, we compared the size of path p_{31} in Figure 2 with the size of that path when semantic ratings were included in the models predicting the post-sampling judgements. That is, we evaluated whether the unique relationship between pre- and post-sampling proximity ratings (after removing the effect of the observed covariations) was mediated by semantic knowledge. The direct semantic ratings were obtained from an independent sample of participants. If semantic knowledge is a component of path p_{31}, then the size of that path would be significantly reduced when the semantic ratings were added as predictors to the regression model that estimated the post-sampling proximity ratings (Kenny, Kashy, & Bolger, 1997).

Method and participants

Participants were 42 students in the Department of Psychology at the Pennsylvania State University. All participants received extra credit for their participation and had an opportunity to partake in a cash lottery.

The original sample contained 64 participants. Because of the time-consuming nature of this study, the drop-out rate among participants was moderately high. Thirteen participants (20% of the original sample) dropped out of the study of their own accord, and three participants (5% of the original sample) were excluded from the sample because they missed more than 10% of their required momentary emotion ratings. Data from 6 more participants were deleted because they contained a large number of retrospective ratings (see below).

Measures

Momentary affect measure. The measure of momentary emotional experiences used in this study included 88 affect terms that anchored all octants of the affective circumplex. Participants indicated on a 7-point scale the extent to which each of these 88 affect adjectives described their emotional state (0 = not at all, 3 = a moderate amount, 6 = a great deal) by responding to the following instruction, ''Indicate to what extent you feel this way right now, that is, at the present moment''. Sixteen affect-related terms were sampled to represent the circumplex space (*excited*, *lively*, *cheerful*, *pleased*, *calm*, *relaxed*, *idle*, *still*, *dulled*, *bored*, *unhappy*, *disappointed*, *nervous*, *fearful*, *alert*, *aroused*). The 16 emotion terms were selected on the basis of their frequency of use (Francis & Kuçera, 1982). Each term was both in common use and was comparable in frequency to all the other terms selected. Very uncommon emotion terms (e.g., euphoric) and extremely common terms (e.g., happy) were not chosen.

The intercorrelations between the 16 circumplex markers were calculated for each participant across the experience-sampling period. These 120 intercorrelations represented the observed covariation between self-reported affective states for each participant; all correlations were submitted to a Fisher r-to-z transformation before use in further analyses.

Similarity measure. Participants rated the similarity of all possible pairs of the 16 circumplex markers (120 word pairs). Each term served both as the referent and as the subject in each pairing, resulting in 240 judgements. Participants were asked to rate the similarity of the word pairs according to the meanings of the words (1 = extremely dissimilar, through 4 = unrelated, 7 = extremely similar). Terms appeared equally spaced throughout the measure (see Davison, 1983). The adjective pairs were presented in a single random order. The 240 similarity ratings were reduced to 120 ratings by averaging the ratings across the corresponding word pairs. This procedure was followed for the ratings taken both before and after the momentary sampling procedure.

Conditional probability measure. Participants rated the conditional probability of all possible pairs of the 16 circumplex markers, again resulting in 240 judgements. Participants were asked to estimate how frequently they experienced each emotion in the presence of every other emotion (1 = never, 3 = sometimes, 6 = always). Terms appeared equally spaced throughout the measure and the adjective pairs were presented in a single random order (although this order was different than that used for the similarity ratings). The conditional probability ratings served as estimates of participants' beliefs concerning the co-occurrence of their emotional experiences. The conditional probability ratings for corresponding word pairs were averaged, producing 120 ratings for the 120 word pairs. This procedure was followed for the ratings taken both before and after the momentary sampling procedure.

Procedure

At the beginning of the study, participants completed a battery of questionnaires including the similarity measure and the conditional probability measure. Participants were then presented with instructions for the momentary emotion ratings. Participants completed an emotion questionnaire in the morning (7 a.m. to 12 p.m.), afternoon (12 p.m. to 5 p.m.), and evening (5 p.m. to 12 a.m.) every day for 60 consecutive days. They returned completed forms on Monday, Wednesday, and Friday of each week. The experimenters contacted participants within 48 hours if they failed to return a questionnaire, and they interviewed participants three times during the study to ensure compliance with the research procedures.

After participants completed the study, the experimenters explained the purpose of the study and then asked a number of questions regarding participation. Subjects estimated the percentage of time that they used recall to complete their questionnaires. Six participants (approximately 9% of the total sample) reported doing so more than 25% of the time and were deleted from the final sample. None of the 42 remaining participants missed more than 5% of the observations, and the average observations missed was 1%. Some participants completed emotion ratings on more days than were required, and these observations were included in the study. The number of usable observations ranged from 171 to 206, with a mean of 181 and a standard deviation of 5.57. No participant reported awareness of the hypotheses under investigation. When we asked participants to describe their reactions to the momentary emotion measurements, participants reported that they found the experience to be mildly to moderately time-consuming, but not stressful. No participants reported that their participation in the study was significantly disruptive.

Semantic data

An independent sample of 65 undergraduate psychology students at the Pennsylvania State University judged the type of valence and the amount of arousal denoted by each of the 88 affect terms included on the momentary emotion measure. Participants were asked to rate the pleasantness or unpleasantness of each emotional state on a 7-point scale (1 = extremely unpleasant, 4 = neutral, 7 = extremely pleasant), and to rate the amount of bodily activation associated with each emotional state (1 = extremely low key, 4 = neutral, 7 = extremely keyed up). The mean rating for each word was computed. A valence-based proximity matrix was calculated by taking the absolute value of the difference between the mean valence ratings for all 120 pairs of emotion words. The smaller the absolute value between two means, the more similar those terms in the valence they denoted. The arousal-based proximity matrix was calculated analogously using the mean arousal ratings.

Results

The hypotheses to be tested using path analysis reduced down to one statistical question: What were the average relationships between the observed covariations, the conditional probability ratings or the similarity ratings, and the direct semantic ratings? We compared the various ratings across the 120 word pairs within each individual and then averaged these relationships across individuals. We expected individual differences in the sizes of these relationships, but we were primarily interested in the averaged estimates in this report.

We used hierarchical linear modelling (HLM; Bryk & Raudenbush, 1987, 1992; Bryk, Raudenbush, Seltzer, & Congdon, 1989) to compute path coefficients. Conceptually, HLM analyses calculate one regression equation at the level of the individual (e.g., regressing post-sampling similarity ratings on the pre-sampling ratings for each individual across 120 pairs of words) and then finds the average of that estimate across individuals. The variance across the individual estimates is also computed. Practically, HLM models within-subject and between-subject variation simultaneously, thus allowing us to model each source variation while taking the statistical characteristics of the other level into account. As a result, the person-level coefficients are treated as random effects when computing the average estimate. Individuals who produce more reliable estimates are weighted more in the computation of the average estimate, and in doing so, HLM controls for reliability differences across individuals, thereby protecting against spurious findings due to differences in measurement error across individuals. A more detailed explanation of the statistical analyses is presented in the Appendix.[2]

The paths presented in Figure 2 were estimated using two HLM regression analyses for the similarity ratings, and two for the conditional probability ratings: one to model the relationship between pre-sampling ratings and observed emotion covariations (predictive correspondence path p_{21}), and a second that regressed the post-sampling ratings on both the pre-sampling ratings (additional emotion knowledge path p_{31}) and the observed covariations (calibration path p_{32}).

Predictive correspondence

On average, participants' pre-sampling proximity ratings reflected some consistent information in the actual co-occurrence of emotional states. The estimates for the predictive correspondence path, p_{21}, are presented in the first two data columns of Table 2. The average predicted correspondence effect for the similarity ratings was $p_{21} = .57$, $t = 16.01$, $p < .001$. The magnitude of this finding is almost identical to that observed in Study 1. The average predicted correspondence effect for the pre-sampling conditional probability ratings was

[2]Using a conventional regression approach to analyse the present dataset, we would run a series of regression analyses for each participant, comparing the relevant variable across the 120 pairs of emotion words. The result would be one set of regression coefficients for each participant and would represent an attempt to model the within-subject variance in the emotion ratings. The regression coefficients from these within-subject anlayses would be considered the outcome variables and averaged across individuals to obtain mean estimates of the path coefficients. A version of this analysis strategy was used in the second study reported by Schimmack and Reisenzein (1997). This two-step ordinary least-squares analytic procedure can be less than optimal, however, because it does not take into account the estimation error in the within-subject regression coefficients (Kenny et al., 1997).

TABLE 2
Predictive accuracy and calibration of proximity judgements

	Predictive correspondence		*Calibration*	
Ratings	*b*	*B*	*b*	*B*
Similarity	0.11	0.57	0.99	0.21
Conditional probability	0.15	0.58	1.12	0.28

$p_{21} = .58$, $t = 13.34$, $p < .001$. These findings indicated that the indices of emotion representation taken before the experience-sampling portion of the study contained some covariance information; approximately 33% of the variance in the pre-sampling proximity ratings was accounted for by the observed covariations. Participants varied significantly from one another in the size of their unstandardised paths, SD = .04, χ^2 (41, $N = 42$) = 372.02, $p < .0001$, and SD = 0.06, χ^2 (41, $N = 42$) = 550.13, $p < .001$, suggesting the possibility that the similarity ratings and covariation estimates made by some participants contained more covariation information than did those made by other participants.

Calibration

On average, participants modestly calibrated their mental representations to their actual experience. The estimates for the calibration path, p_{32}, are presented in the third and fourth data columns of Table 2. The average calibration effect for the similarity ratings was $p_{32} = .21$, $t = 7.78$, $p < .001$, and for the conditional probability ratings was $p_{32} = .28$, $t = 11.09, p < .0001$. Although modest, the effects indicated that on average, participants' judgements did become somewhat more accurate after the momentary sampling procedure. Participants varied significantly from one another in the size of their unstandardised paths, SD = 0.73, χ^2 (41, $N = 42$) = 283.62, $p < .0001$, and SD = 0.56, χ^2 (41, $N = 42$) = 209.97, $p < .001$, suggesting the possibility that some participants calibrated more so than did others.

Additional emotion knowledge: A semantic hypothesis

The results presented thus far demonstrated that, on average, judgements representing what people know about emotion, be they similarity ratings or covariance estimates, do contain some accurate covariance information. In addition, individuals did modestly calibrate these judgments to new information about their experience when cued over a two month period. The results presented in Table 3 indicate, however, that indices representing emotion structure also contained other sources of knowledge. The average effect representing

TABLE 3
Additional emotion knowledge contained in proximity judgements

	Additional knowledge		Semantic knowledge				Additional knowledge controlling for semantic information	
			Valence		Arousal			
Ratings	*b*	*B*	*b*	*B*	*b*	*B*	*b*	*B*
Similarity	0.61	0.66	0.11	0.07	0.36	0.21	0.45	0.48
Conditional probability	0.59	0.58	0.15	0.11	0.18	0.12	0.46	0.45

Note: All coefficients were significant at $p < .001$.

additional emotion knowledge contained in the similarity ratings was $p_{31} = .66$, $t = 17.78$, $p < .0001$, and in the conditional probability ratings was $p_{31} = 0.58$, $t = 15.40$, $p < .001$. These findings indicate that participants' pre- and post-sampling ratings shared considerable variance that was independent of the covariations evident in their self-reports of experience. Participants varied significantly from one another in the size of their unstandardised paths, SD = 0.21, χ^2 (41, $N = 42$) = 438.14, $p < .0001$, and SD = 0.24, χ^2 (41, $N = 42$) = 413.24, $p < .001$, suggesting the possibility that the ratings made by some participants contained more noncovariation based emotion knowledge than did those made by other participants.

Next, we explored the hypothesis that at least part of this knowledge was semantic by observing whether a significant portion of the variance shared by the pre- and post-sampling proximity ratings would be accounted for by the semantic ratings of the affect terms. HLM regression analyses were conducted to estimate the size of the p_{31} paths as described above, with the sole exception that the semantic proximity matrices were added as predictors to the regression analyses. Path $p_{31'}$ represented the relationship between the pre- and post-sampling proximity ratings, after controlling for the effects of the observed covariations, and after the mediating effects of the semantic ratings were removed. One HLM analysis was conducted for the post-sampling similarity ratings, and one for the post-sampling covariation estimates. The size of resulting additional knowledge paths are presented in the right half of Table 3. Both valence-based and arousal-based semantic matrices were significantly related to the proximity ratings, indicating that they may be potential mediators of the additional knowledge effect.

To examine whether the semantic proximity matrices significantly accounted for any portion of the additional emotion knowledge effect, we compared the standardised regression coefficients from the original analysis as portrayed in Figure 1 ($p_{31} = .66$ and $p_{31} = .58$) with those from the models where the mediating effects of the semantic ratings were included ($p_{31'} = 0.48$ and

$p_{31'} = .45$; Baron & Kenny, 1986; Kenny et al., 1997). The additional emotion knowledge contained in the similarity ratings was significantly reduced by both the valence-based and arousal-based semantic information, $z = 3.92$, $p < .001$ and $z = 8.43$, $p < .0001$; this was also true for the covariance estimates, $z = 6.58$, $p < .0001$ and $z = 8.49$, $p < .0001$. These findings suggest that semantic knowledge was contained in both the similarity ratings, and in the covariance estimates, over and above any contribution of the observed covariances. A significant portion of the emotion knowledge contained in those judgements was left unexplained by the semantic information, however, indicating that additional sources of knowledge must be influencing the proximity judgements.

Subsidiary analyses

Additional analyses were conducted to further examine the relationships between the proximity judgements and semantic knowledge. First, as proximity ratings, both the similarity ratings and the conditional probability estimates are indices of the mental representation of emotion. Previous research has shown a strong correspondence between the two (Schimmack & Reisenzein, 1997). Because they both contain semantic emotion knowledge, it is possible that the strong association between the two may be due, in part, to this semantic knowledge.

We first replicated the finding that the two types of proximity ratings were highly related. Both the pre-sampling similarity ratings and the pre-sampling conditional probability judgements were subjected to a weighted Euclidean multidimensional scaling analysis.[3] Plots of the Stress values by the number of dimensions for the MDS solution revealed a clear elbows at the two dimensional solution for both analyses, suggesting the suitability of the two dimensional MDS solution for both the similarity and pre-sampling conditional probability ratings, $Stress = .16$, $RSQ = .86$, and $Stress = .17$, $RSQ = .83$, respectively.[4] The MDS solutions for the similarity ratings and the conditional probability judgements are presented in Figures 3a and b, respectively. Both solutions produced a circular structure anchored by two dimensions. The congruence coefficient for the two solutions (Davison, 1983) was .98, indicating an excellent match between the two configurations. The congruence coefficient for the two post-sampling solutions was .97. The similarity ratings and the conditional probability judgements produced almost identical structures.

[3] The primary approach to ties (allowing data to become untied) was used in the analysis because it typically results in a better fit to the data (Davison, 1983, p. 86).

[4] Although Kruskal and Wish (1978, p. 56) caution against accepting solutions with a Stress value above .10, the ‘‘elbow’’ in the plot clearly appeared for the two-dimensional solution. Furthermore, the two-dimensional solution was superior to other solutions with additional dimensions on the basis of the relative interpretability of the various solutions (Davison, 1983).

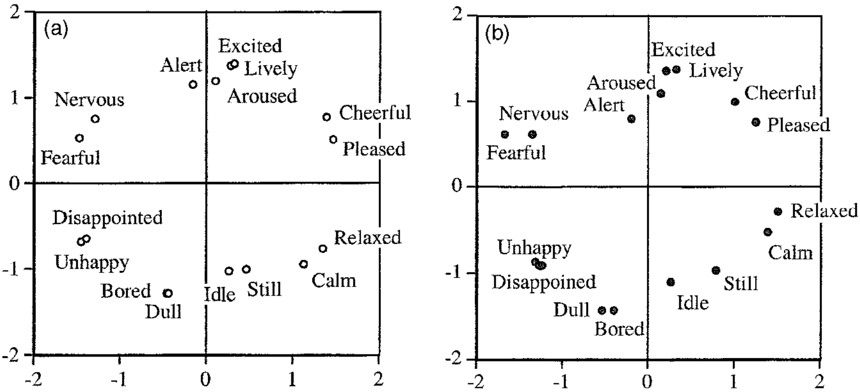

Figure 3. (a) The circumplex structure of affect derived from the pre-sampling similarity judgements. Valence is the horizontal axis and arousal is the vertical axis. (b) The circumplex structure of affect derived from the pre-sampling conditional probability judgements. Valence is the horizontal axis and arousal is the vertical axis.

The strong correspondence between the similarity and conditional probability judgements was confirmed by additional HLM analyses. Using the pre-sampling ratings, we regressed the similarity ratings onto the conditional probability ratings. On average, the similarity ratings were substantially related to the conditional probability judgements, ($B=.71$, $t=23.50$, $p<.001$), although this relationship was stronger for some individuals than for others, SD = 0.20, χ^2 (41, $N=42$) = 338.38, $p<.001$.

Further HLM analyses indicated that the correspondence between the two types of proximity judgements was in part due to semantic knowledge that the two shared. The similarity ratings continued to be related to participants' conditional probability estimates even after the influence of the observed covariations in emotional experiences was controlled ($B=.55$, $t=16.16$, $p<.001$), although, once again, this relationship was stronger for some than for others, SD = 0.20, χ^2 (41, $N=42$) = 301.06, $p<.001$. A path coefficient of .55 indicates that additional emotion knowledge contained in the covariation estimates, over and above that abstracted from actual experience, was also contained in the similarity ratings. When the variance due to direct semantic ratings were partialled out, the relationship between the similarity ratings and the covariation estimates decreased significantly to $B=.32$, $t=12.35$, $p<.001$, although once again, the magnitude of this effect varied significantly across individuals, SD = 0.18, χ^2 (41, $N=42$) = 154.92, $p<.001$). The observed decrease was due both to the valence-based and to the arousal-based semantic information ($z=4.91$, $p<.0001$ and $z=13.25$, $p<.0001$). Analyses of the post-sampling judgements produce exactly the same findings. These findings indicate that the similarity and conditional probability ratings shared significant content that was semantically based, although the strength of their correspondence was not completely accounted for by the semantic ratings.

GENERAL DISCUSSION

The circumplex model of affect has a long history in the psychological literature (Russell, 1980; Schlosberg, 1941, 1952, 1954; Woodworth, 1938; Wundt, 1912/1924). For some time, it was assumed that general semantic knowledge about pleasure/displeasure and activation was represented by this circumplex structure. This assumption was challenged by Schimmack and Reisenzein (1997) who suggested that the similarity ratings that produce the structure contain episodic, but not semantic, knowledge of emotion.

Semantic or episodic?

Taken together, the findings from Studies 1 and 2 clearly indicate that a pure episodic account of similarity judgements like that presented by Schimmack and Reisenzein (1997) is not supported. Using an episodic or exemplar theory of knowledge, they argued that every experienced co-occurrence leaves a unique memory trace, and that, on presentation of a cue (i.e., a request to compute a similarity judgement), all stored episodic traces (e.g., Semon, 1909/1923) or some assemblage of exemplars (Kahneman & Miller, 1986) are activated, and somehow combined to result in a proximity judgement. The most activated traces connect the cue to stored knowledge and a response is generated. The results of the two studies presented here are not consistent with this view, however.

A conservative view would interpret the results as more consistent with mixed semantic-episodic theory. From this standpoint, mental representations of affect are composed of a mixture of detail (specific memories of associated co-occurrence experiences) and more abstract, conceptual knowledge bound together in representations (e.g., Conway, Gardens, Perfect, Anderson, & Cohen, 1997), such that both types of knowledge are implicated in similarity computations. Clearly, similarity ratings and co-occurrences of actual affective experience are related. Both studies indicated that about one-third of the variance in similarity ratings was accounted for by observed emotion covariations. Furthermore, participants modestly calibrated their proximity judgements to their actual experience after they had been cued about their affective experience for a 60-day period, indicating that they can take advantage of episodically generated experience to compute both similarity and conditional probability judgements. In addition to covariance information, proximity ratings also contained other sources of knowledge, some of which appeared to be semantic. Ratings of valence- and arousal-based semantic information obtained from a set of independent judges contributed unique variance to both sets of proximity ratings, indicating that the circumplex (which derives from those ratings), does indeed contain semantic information.

The view that mental representations of affect are composed of both semantic and episodic knowledge is broadly consistent with current thinking on the nature of memory. Although memory was traditionally classified exclusively into episodic or semantic components (Tulving, 1972), most researchers now assume that the two are inextricably linked. Some believe that memory for specific events and for general facts are subsumed in one general, declarative memory system (e.g., Echienbaum, 1997; Squire & Knowlton, 1995). Others believe that semantic and episodic memory are organised in a hierarchical way, such that episodic memory is a specific subsystem of semantic memory (Tulving, 1983, 1984) and depends on semantic memory for its integrity (Tulving, 1985; for a review see Barba, Parlato, Jobert, Samson, & Pappata, 1998). Recent neuroscience evidence basically supports this view (e.g., Barba et al., 1998; Maguire & Mummery, 1999; Wiggs, Weisberg, & Martin, 1998).[5]

Episodic memory is dependent on semantic knowledge such that the retrieval of episodic knowledge is supported by semantic memory (Nadel & Moscovich, 1998). Various parts of an experience are structurally disaggregated in the brain for purposes of storage, and must be re-aggregated when attempting to retrieve an episode from memory (Nadel & Moscovich, 1998; Schacter, 1996). Retrieving episodes or exemplars from memory is necessarily a reconstructive act, and one that can be influenced by semantic knowledge. Thus, it very well may be, as Schimmack and Reisenzein (1997) claim, that similarity ratings are based on covariation estimates. These estimates may be generated, however, on the basis of semantic knowledge (Borkenau & Ostendorf, 1987; D'Andrade, 1974; Schweder, 1975, 1977; Schweder & D'Andrade, 1979). This suggestion is supported by the finding that conditional probability judgements for affect word pairs, which were used as estimates of episodic knowledge of co-occurrence, were, in fact, related to the semantic information about affect obtained from a separate set of judges.

Another reason to suspect that the conditional probability ratings were substantially influenced by semantic knowledge has to do with how the questions were asked. Participants in the current studies, as well as those in Schimmack and Reisenzein (1997), were to make judgements of covariation without reference to context (i.e., time or place), leading to the likelihood that these aggregated episodic ensembles constitute a semanticised version of episodic knowledge (Nadel & Moscovitch, 1998). Although participants may have had to retrieve specific instances or episodes, those episodes were summarised such

[5] Cortical and limbic networks implicated in semantic and episodic memory show both common and unique regions although the exact pattern of neurological findings is still a matter of considerable debate (for reviews, see Barba et al., 1998).

that the memory lost its connection to time and place, both of which are necessary components of a pure episodic memory.

Although episodic knowledge can be dependent on semantic memory, it is also the case that episodic memory helps to consolidate semantic knowledge (Nadel & Moscovitch, 1998), although little is known about how this transition occurs (Conway et al., 1997). It is generally assumed that semantic knowledge derives from a series of episodes in the course of which semantic structure is extracted from experience (McClelland McNaughton, & O'Reilly, 1995; Nadel & Moscovitch, 1998). The knowledge acquired in any given episode goes beyond the specifics of that episode to include broader knowledge gained as a result of having had that experience.

Nowhere is this interplay between semantic and episodic memory more likely than in the case of developing emotion knowledge. As children, we derive our knowledge of emotion and emotion terms both from instruction about emotion concepts and from our actual experience. Emotion concepts probably begin as fairly stereotyped scripts or schemas (e.g., Fehr & Russell, 1984; Shaver, Schwartz, Kirson, & O'Connor, 1987) as children are socialised to learn the semantic, interpersonal, and behavioural elements associated with specific emotion labels (Harris, 1993). Children as young as two readily label their emotional experiences (Bretherton, McNew, & Beeghly-Smith, 1981). They rapidly learn the type of psychological events and abstract situations that are associated with particular emotion labels (e.g., fear, sadness, happiness, anger, guilt, and so forth; e.g., Harris, Olthof, Meerum, Terwogt, & Hardman, 1987), and they are also aware of the typical actions and expressions that are supposed to accompany a particular emotional state (Trabasso, Stein, & Johnson, 1981). These concepts then become developed and elaborated with accumulating episodic experience. Individuals incorporate elements of their actual experience into their semantic representations of emotion concepts with repeated use, such that their resulting semantic representations become more idiographically tailored regarding the possible class of objects that can cause an emotional response, the relational contexts associated with the response, and the behavioural repertoire that exists for dealing with the response and the larger situation.

As a result, emotion terms may come to be defined over time, in part, as a function of our beliefs about how often distinct emotional experiences occur together (i.e., episodic influencing semantic). In addition, however, our consciously accessible emotional experience is defined, in part, from our semantic knowledge of emotion language. That is, we likely rely on semantic information when applying emotion labels to our emotional experiences, which are in turn encoded in episodic memory (i.e., semantic influencing episodic). Thus, we derive our knowledge of emotion, in part, from our actual experience and over time; conversely, we likely also rely on semantic information when applying emotion labels to our immediate experiences are then available to be encoded in episodic memory.

The question, then, is not whether there is episodic influence on mental representations of emotion: Of course there is. The question, rather, is whether the general valence and arousal information depicted in the circumplex represents episodic knowledge. If it does, then would expect to see the considerable idiographic variability in the similarity ratings that we see in the self-report based structures. The fact that similarity ratings seem not to vary substantially across individuals, however, is inconsistent with a strong episodic view. Rather, the evidence seems more consistent with the view that similarity ratings are primarily computed using some general knowledge that all individuals share, that is abstracted from specific emotional experiences. Episodic variation in emotion representations may be present in other, more subtle aspects of emotion concepts.

So an empirical resolution to the semantic-episodic question seems unsatisfying because it appears as if everything is related to everything else. But perhaps this is about as accurate as we can be, given the procedures that we have used to test these competing hypotheses. It is probably impossible to precisely estimate semantic and episodic components in similarity judgements or the mental representation that they index. Behavioural studies such those reported here, and those by Schimmack and Reisenzein (1997), cannot precisely distinguish between episodic and semantic sources of knowledge in an experimental task. Given the interrelationship between semantic and episodic memory, it is probably virtually impossible to devise behavioural tasks that are pure measures of one or the other (cf. Barba et al., 1998). Despite this ambiguity, the two studies reported here make several things clear. First, a pure episodic view of similarity ratings is not tenable. Second, the interpretation of the findings likely provide an underestimate of the semantic contribution because semantic knowledge may be involved in computing covariation estimates as it plays a role in consciously representing our affective experience (and the co-occurrences in those experiences) in episodic memory. Thus, the strong relationship between similarity ratings and observed emotion covariations may result from their shared semantic components, rather than from the fact that similarity computations are rooted in episodic experience.

Additional issues

The results raised two additional issues that might benefit from further study. First, further study is needed to identify the specific facets of emotion knowledge that contribute to mental representations of emotion, over and above semantic and accurate covariation information. The additional knowledge contained in the similarity and conditional probability ratings was partially semantic, but a substantial portion of those ratings remained unexplained. Perhaps such knowledge is based on dominance (Russell & Mehrabian, 1977), affiliation (Russell, 1991), or social desirability (Feldman Barrett, 1996). Or perhaps it is related to beliefs about the situational determinants of emotions (Conway & Bekerian, 1987; Feldman Barrett, Lane, Sechrest, & Schwartz, 2000), although it is unclear whether these beliefs accurately represent the true influences on how individuals use emotion knowledge to consciously label their emotional states (Russell, 1987; Weiss & Brown, 1977, as cited in Nisbett & Ross, 1980; Wilson, Laser, & Stone, 1982). Unfortunately, we did not assess any of these facets in the present study.

Second, there was considerable variation in the relationships between the observed covariances, the proximity ratings, and the semantic components. This variation might possibly reflect sampling error or other types of artifactual variance (e.g., restriction in range for some individuals), but it might also suggest that the configuration of episodic and semantic components in proximity ratings may substantively differ across individuals. There is also the possibility of individual differences in the organisation of emotion knowledge across different people, such that some individuals may have more overlap between their semantic and episodic knowledge of emotion terms than do others. Although not a primary concern in the present study, these findings are certainly worthy of serious consideration in future research on the organisation of emotion knowledge.

Potential limitations

Of course, it is possible to argue with our interpretation of the results. For example, it might be argued that momentary affective experiences were assessed only three times a day in the present study and may not provided enough temporal resolution to assess moment-to-moment emotion covariations accurately. As a result, the role of actual covariation information may be larger or smaller than estimated in the present study. Alternatively, it could also be argued that participants' assessments of emotion were randomly sampled from their daily lives and so they likely constitute a reasonable estimate of emotion co-occurrences.

A more serious criticism of the present study is that we cannot claim support for any particular directional hypotheses because the data are only correlational. As recommended by regression experts (Pedhauzer, 1982), the zero-order correlations between variables were decomposed into their various components via regression analyses according to the hypotheses being tested. It is possible to use other theories to decompose the relationships between the measured variables in ways that are different from those presented here, however. Yet even if paths are reversed and alternative causal models estimated, with the result that the zero-order correlations are parsed somewhat differently, the findings send approximately the same message: similarity ratings and covariation estimates are associated with covariation information, semantic knowledge, as well as additional, as yet unspecified knowledge about emotion.

Another potential criticism of our findings is that they are inconsistent with Schimmack and Reisenzein (1997), who reported two studies that they consider strong support for the episodic model of similarity ratings. The strength of their position comes from the finding that their participants made asymmetrical similarity judgements when emotion terms were presented as both the subject and the referent of the comparison. According to Schimmack and Reisenzein (1997), a semantic theory assumes that similarity judgements are symmetrical (i.e., *happiness* and *calm* will be judged to be equally similar regardless of which emotion serves as the subject of the comparison and which serves as the referent). An episodic theory predicts that similarity judgments can be asymmetrical, however (i.e., the similarity ratings of *happiness* and *calm* will vary depending on which emotion serves as the subject of the comparison). Distinctive features of a concept decrease similarity more if the concept serves as the subject of the comparison than if it serves as the referent (Tversky, 1977). Schimmack and Reisenzein (1997) assumed that an emotion's distinctiveness increased: (1) with the number of episodes in which it occurs alone (as opposed to occurring with another emotion); or (2) as the frequency of the emotion increases. Thus, they predicted, and found, that an emotion with a lower conditional probability or higher frequency would be judged by respondents as less similar if it served as the subject of the similarity comparison rather than as the referent. Furthermore, they found that asymmetrical similarity judgements were related to beliefs about the frequency and conditional probability of occurrence of emotions, but were unrelated to the semantic properties of the emotion concepts. In addition, respondents judged the similarity of emotion pairs more rapidly when a frequent emotion was the referent of the comparison, which was interpreted as support for the hypothesis that frequency of experience (or distinctiveness) of an emotion affects similarity judgements depending on the word's position in the comparison. Support for the covariation hypothesis was most robust when respondents were judging emotion terms denoting unpleasant emotions. In fact, they only used unpleasant emotion terms in the second study they reported.

We are less convinced that asymmetries are the definitive evidence that mental representations of emotion, as indexed by similarity ratings, are primarily episodic in nature. First, the asymmetries documented by Schimmack and Reisenzein (1997) could have been the result of response bias. Asymmetries in similarity judgements are generally weak (Holyoak & Gordon, 1983; Tversky & Gati, 1978) and if they exist at all they may be produced by response bias (Nosofsky, 1991). For example, Schimmack and Reisenzein (1997) reported that participants largely agreed with one another in their asymmetry of their similarity judgements. Given the individual differences in actual emotion covariations documented in the present study, as well as in other studies (e.g., Feldman Barrett, 1998), such agreement should have been highly unlikely.

Second, the asymmetries documented by Schimmack and Reisenzein (1997) may have been augmented by the experimental procedures. Participants were told to focus their attention on the subject of the comparison, were informed as to which emotion term would serve as the subject of the comparison, and were instructed to think about this emotion. In addition, the subject of the comparison was presented in large letters to increase its visual salience, and the presentation of the referent emotion was delayed to provide participants with more time to focus on the subject of the comparison. Any of these procedures could have produced more pronounced asymmetry effects, and so confidence in the existence of asymmetrical similarity judgements requires further investigation.

Finally, Schimmack and Reisenzein (1997) describe a process by which similarity ratings are generated that is inconsistent with their findings. Schimmack and Reisenzein (1997) argued that the episodic knowledge contained in similarity judgements is not prestored in memory, but is thought to be computed from exemplar knowledge that is accessible at the time when judgements are requested. That is, this information is derived from memory-based, rather than on-line, recall strategies (Hasties & Park, 1986; McConnell, Sherman, & Hamilton, 1994). If episodic knowledge is constructed at the time of recall, then the judgements made on the basis of that knowledge should be influenced by whatever is most accessible at that time. This influence might derive from heuristics, implicit theories, or goals and motivations that are present or accessible at the time when the judgements are made (for reviews see Fiske & Taylor, 1991; Greenwald & Banaji, 1995; Nisbett & Ross, 1980; Ross, 1989). Typically, we do not expect to see much consistency in memory-based judgements (McConnell et al., 1994). In this context, it is interesting to note that Schimmack and Reisenzein (1997) reported very high inter-rate agreement in their conditional probability ratings (above .90).

CONCLUSIONS

A sceptic might review the results of this study and conclude that they do not determine much, because everything that was measured was related to everything else. Another equally valid conclusion, however, is that mental representations of emotion, like most psychological phenomena, appear to be multiply determined. They contain accurate covariation information, semantic information that is unique from the information contained in actual emotion covariations, as well as additional sources knowledge yet to be specified. The similarity ratings contained as much additional knowledge as they did accurate covariation information. Even the conditional probability judgements, as estimates of covariation, contained a substantial additional knowledge component. Taken together, the study did not serve to conclusively determine the extent to which semantic knowledge is contained in similarity ratings, but it clearly indicated that the ratings are not primarily grounded in episodic knowledge.

Manuscript received 31 July 1998
Revised manuscript received 31 July 2000

REFERENCES

Abelson, R.P., & Sermat, V. (1962). Multidimensional scaling of facial expressions. *Journal of Experimental Psychology*, *63*, 546–554.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*, 669–672.

Alloy, L.B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, *91*, 112–149.

Almagor, M., & Ben-Porath, Y.S. (1989). The two-factor model of self-reported mood: A cross-cultural replication. *Journal of Personality Assessment*, *53*, 10–21.

Barba, G.D., Parlato, V., Jobert, A., Samson, Y., & Pappata, S. (1998). Cortical networks implicated in semantic and episodic memory: Common or unique? *Cortex*, *34*, 547–561.

Baron, R.M., & Kenny, D.A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Block, J. (1957). Studies in the phenomenology of emotions. *Journal of Abnormal and Social Psychology*, *54*, 358–363.

Borkenau, P., & Ostendorf, F. (1987). Retrospective estimates as act frequencies: How accurately do they reflect reality? *Journal of Personality and Social Psychology*, *52*, 626–638.

Bretherton, I., McNew, S., & Beeghly-Smith, M. (1981). Early person knowledge as expressed in gestural and verbal communication: When do infants acquire a theory of mind? In M. Lamb & L. Sherrod (Eds.), *Infant social cognition*. Hillsdale, NJ: Erlbaum.

Bryk, A.S., & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, *101*, 147–158.

Bryk, A.S., & Raudenbush, S.W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Bryk, A.S., Raudenbush, S.W., Seltzer, M., & Congdon, R.T. (1989). *An introduction to HLM: Computer program and user's guide*. Chicago: University of Chicago Press.

Bush, L.E. (1973). Individual differences multidimensional scaling of adjectives denoting feelings. *Journal of Personality and Social Psychology*, *25*, 50–57.

Cliff, N., & Young, F.W. (1968). On the relation between unidimensional judgments and multidimensional scaling. *Organizational Behavior and Human Performance*, *3*, 269–285.

Conway, M.A., & Bekerian, D.A. (1987). Situational knowledge and emotions. *Cognition and Emotion*, *1*, 145–191.

Conway, M.A., Gardiner, J.M., Perfect, T.J., Anderson, S.J., & Cohen, G.M. (1997). Changes in memory awareness during learning: The acquisition of knowledge by psychology undergraduates. *Journal of Experimental Psychology: General*, *126*, 393–413.

Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, *90*, 272–292.

D'Andrade, R.G. (1974). Memory and the assessment of behavior. In T. Blalock (Ed.), *Measurement in the social sciences* (pp. 159–186). Chicago, IL: Aldine-Atherton.

Davison, M.L. (1983). *Multidimensional scaling*. Toronto: Wiley.

Dittmann, A.T. (1972). *Interpersonal messages of emotion.* New York: Springer.
Echienbaum, H. (1997). How does the brain organize memories? *Science*, *277*, 330–332.
Fehr, B., & Russell, J.A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, *113*, 464–486.
Feldman, L.A. (1995a). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, *69*, 153–166.
Feldman, L.A. (1995b). Variations in the circumplex structure of emotion. *Personality and Social Psychology Bulletin*, *21*, 806–817.
Feldman Barrett, L. (1996). Hedonic tone, perceived arousal, and item desirability: Three components of self-reported mood. *Cognition and Emotion*, *10*, 47–68.
Feldman Barrett, L. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition and Emotion*, *12*, 579–599.
Feldman Barrett, L., & Gross, J. (in press). Emotion representation and regulation: A process model of emotional intelligence. In T. Mayne & G. Bonnano (Eds.), *Emotion: Current issues and future directions.* New York: Guilford Press.
Feldman Barrett, L., Lane, R., Sechrest, L., & Schwartz, G. (2000). Sex differences in emotional awareness. *Personality and Social Psychology Bulletin*, *26*, 1027–1035.
Feldman Barrett, L., & Russell, J.A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, *74*, 967–984.
Fillenbaum, S., & Rapaport, A. (1971). *Structures in the subjective lexicon.* New York: Academic Press.
Fiske, S.T., & Taylor, S.E. (1991). *Social cognition.* Reading, MA: Addison-Wesley.
Francis, W.N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon & grammar.* Boston: Houghton Mifflin.
Green, R.S., & Cliff, N. (1975). Multidimensional comparisons of structures of vocally and facially expressed emotion. *Perception and Psychophysics*, *17*, 429–438.
Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*, 4–27.
Guttman, L. (1954). A new approach to factor analysis: The radex. In P.F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258–348). New York: Columbia University Press.
Guttman, L. (1966). Order analysis of correlation matrices. In R.B. Cattell (Ed.), *Handbook of multivariate experimental psychology*. Chicago, IL: Rand-McNally.
Harris, P.L. (1993). Understanding emotion. In M. Lewis & J.M. Haviland (Eds.), *Handbook of emotions* (pp. 237–246). New York: Guilford Press.
Harris, P.L., Olthof, T., Meerum, Terwogt, M., & Hardman, C.E. (1987). Children's understanding of situations that provoke emotion. *International Journal of Behavioral Development*, *10*, 319–343.
Hasties, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review*, *93*, 258–268.
Holyoak, K.J., & Gordon, P.C. (1983). Social reference points. *Journal of Personality and Social Psychology*, *44*, 881–887.
Kahneman, D., & Miller, D.T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*, 136–153.
Kenny, D.A., Kashy, D.A., & Bolger, N. (1997). Quantitative methods in social psychology. In D. Gilbert, S. Fiske, & G. Lindsey (Eds.), *Handbook of social psychology* (4th ed.). New York: McGraw-Hill.
Kruskal, J.B. & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.
Linton, M. (1986). Ways of searching and the contents of memory. In D.C. Rubin (Ed.), *Autobiographical memory* (pp. 50–67). Cambridge, UK: Cambridge University Press.

Maguire, E.A., & Mummery, C.J. (1999). Differential modulation of a common memory retrieval network revealed by positron emission tomography. *Hippocampus*, *9*, 54–61.

McClelland J.L., McNaughton, B.L., & O'Reilly, R.C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failues of connectionist models of learning and memory. *Psychological Review*, *102*, 419–457.

McConnell, A.R., Sherman, S.J., & Hamilton, D.L. (1994). On-line and memory-based aspects of individual and group target judgments. *Journal of Personality and Social Psychology*, *67*, 173–185.

McKenzie, C.R.M. (1994). The accuracy of intuitive judgment strategies: Covariation assessment and bayesian inference. *Cognitive Psychology*, *26*, 209–239.

Nadel, L., & Moscovich, M. (1998). Hippocampal contributions to cortical plasticity. *Neuro-pharmacology*, *37*, 431–439.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.

Nosofsky, R.M. (1991). Stimulus bias, asymmetric similarity, and classification. *Cognitive Psychology*, *25*, 94–140.

Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Pedhauzer, E.J. (1982). *Multiple regression in behavioral science* (2nd ed.). New York: Holt, Rinehart, and Winston.

Reisenzein, R. (1994). Pleasure-activation theory and the intensity of emotions. *Journal of Personality and Social Psychology*, *67*, 525–539.

Reisenzein, R., & Schimmack, U. (1999). Similarity judgments and covariations of affects: Findings and implications for affect structure research. *Personality and Social Psychology Bulletin*, *25*, 539–555.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.

Russell, J.A. & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*, 273–294.

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.

Russell, J.A. (1983). Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality and Social Psychology*, *45*, 1281–1288.

Russell, J.A. (1987). Comments on articles by Fridja and by Conway and Bekerian. *Cognition and Emotion*, *1*, 193–197.

Russell, J.A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, *110*, 426–450.

Russell, J.A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expression: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, *38*, 1290–1298.

Russell, J.A., Lewicka, M., & Niit, T. (1989). A cross-cultural study of a circumplex model of affect. *Journal of Personality and Social Psychology*, *57*, 848–856.

Russell, J.A., & Feldman Barrett, L. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805–819.

Schacter, D.L. (1996). *Searching for memory: The brain, the mind, and the past*. Cambridge MA: Harvard University Press.

Schimmack, U., & Hartmann, K. (1997). Individual differences in the memory representation of emotional episodes: Exploring the cognitive processes in repression. *Journal of Personality and Social Psychology*, *73*, 1064–1079.

Schimmack, U., & Reisenzein, R. (1997). Cognitive processes involved in similarity judgments of emotion. *Journal of Personality and Social Psychology*, *73*, 645–661.

Schlosberg, H. (1941). A scale for the judgment of facial expressions. *Journal of Experimental Psychology*, *29*, 497–510.

Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology, 44*, 229–237.

Schlosberg, H. (1954). Three dimensions of emotion. *Psychological Review, 61*, 81–88.

Semon, R. (1923). *Mnemic psychology* (B. Duffy, Trans.). Concord, MA: George Allen & Unwin. (Original work published 1909)

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*, 1061–1086.

Shepard, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, *27*, 219–246.

Shepard, R.N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika*, *39*, 373–422.

Shepard, R.N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, *210*, 390–398.

Shweder, R.A. (1975). How relevant is an individual difference theory of personality? *Journal of Personality*, *43*, 455–484.

Shweder, R.A. (1977). Likeness and likelihood in everyday thought: Magical thinking in judgments about personality. *Current Anthropology*, *18*, 637–648.

Shweder, R.A. & D'Andrade, R.G. (1979). Accurate reflection or systematic distortion? A reply to Block, Weiss, and Thorne. *Journal of Personality and Social Psychology*, *37*, 1075–1084.

Squire, L.R., & Knowlton, B. (1995). Memory, hippocampus, and brain systems. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 825–837). Cambridge MA: MIT Press.

Trabasso, T., Stein, N., & Johnson, L.R. (1981). Children's knowledge of events: A causal analysis of story structure. In G. Bower (Ed.), *Learning and motivation* (Vol. 15, pp. 237–282). New York: Academic Press.

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Conaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, UK: Oxford University Press.

Tulving, E. (1984). Precis of elements of episodic memory. *Behavioral and Brain Science*, *2*, 223–268.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, *26*, 1–12.

Tversky, A. (1977). Features of similarity. *Psychological Review*, *76*, 31–48.

Tversky, A., & Gati, I. (1978). Studies in similarity. In E. Rosch & B.B. Lloyd (Eds.), *Cognition and categorization* (pp. 79–98). Hillsdale, NJ: Erlbaum.

Weiss, J., & Brown, P. (1977). *Self-insight error in the explanation of mood*. Unpublished manuscript, Harvard University.

Wiggs, C.L., Weisberg, J., & Martin, A. (1998). Neural correlated of semantic and episodic memory retrieval. *Neuropsychologia*, *37*, 103–118.

Wilson, T.D., Laser, P.S., & Stone, J.I. (1982). Judging predictors of one's own mood: accuracy and the use of shared theories. *Journal of Experimental Social Psychology*, *18*, 537–556.

Woodworth, R.S. (1938). *Experimental psychology*. New York: Holt.

Wundt, W. (1924). *An introduction to psychology* (R. Pintner, Trans.) London: Allen & Unwin. (Original work published 1912)

Zevon, M.A. & Tellegen, A. (1982). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, *43*,111–122.

APPENDIX A

All paths reported in this report were estimated using the HLM computer program (Bryk et al., 1989). For illustrative purposes, we will present a breakdown of the conditional probability rating analyses used to estimate the paths for Figure 1. Two hierarchical linear regressions were used to calculate the paths for this analysis. The within-subject level of the first hierarchical linear model estimated predictive correspondence hypothesis (p_{21} in Figure 1) using the model:

$$\mathrm{OC}_{ij} = b_0 + b_1 \mathrm{CPPRE}_{ij} + r_{ij} \quad (1)$$

where OC_{ij} is participant j's observed covariation for the ith pair of emotions, b_1 is the relationship between participant j's observed covariations and his/her pre-sampling conditional probability ratings (p_{21} in Figure 1), CPPRE_{ij} is participant j's pre-sampling estimate of the degree to which the ith pair of emotions are experienced together, b_0 is the value of the observed covariation in j when $\mathrm{CPPRE}_{ij} = 0$, and r_{ij} is a within-subject residual component. The between-subject level of the first hierarchical linear model allowed us to assess the average of the within-subject evidence for the predictive correspondence hypothesis, as well as between-subject variance by estimating the degree to which the within-subjects coefficients (b_0, b_1) for participants varied from one another.

Because we were not interested in intercept differences, we only present the modelling of the slopes (the b_1 coefficients), as follows:

$$b_{1j} = b_{10} + u_{1j} \quad (2)$$

where b_{1j} is the relationship between the observed covariation estimates and pre-sampling conditional probability ratings of participant j, b_{10} is the mean relationship between the two for all participants (i.e., the mean p_{21} value), and u_{1j} represents the random variation in this relationship (i.e., variation in values of p_{21}).

The within-subject level of the second hierarchical linear model estimated the calibration hypothesis (i.e., the direct effect of participants' actual emotion covariations on their post-sampling conditional probability ratings; p_{32} in Figure 1) and the additional knowledge path (p_{32} in Figure 1) using the model:

$$\mathrm{CPPOST}_{ij} = b_0 + b_1 \mathrm{OC}_{ij} + b_2 \mathrm{CPPRE}_{ij} + r_{ij} \quad (3)$$

where CPPOST_{ij} is participant j's post-sampling conditional probability estimate for the ith pair of affect words, b_0 is the mean estimate for participant j, b_1 is the direct effect of participant j's actual emotion covariations on his/her post-sampling conditional probability estimates (p_{32}; this effect represents the relationship between participant j's actual emotion covariations and post-sampling conditional probability estimates, controlling for his/her pre-sampling), OC_{ij} is participant j's correlation for the ith pair of emotions, b_2 is the effect of participant j's pre-sampling estimates on his/her post-sampling estimates (p_{31}; this effect represents the effect of participant j's pre-sampling beliefs about emotion covariation on his/her post-sampling beliefs, controlling for the actual covariation in his/her emotional experiences), CPPRE_{ij} is participant j's pre-sampling estimate of the degree to which the ith pair of emotions are experienced together, b_0 is the value of CPPOST when the predictions are zero, and r_{ij} is a within-subject residual component.

The between-subject level of the second hierarchical linear model allowed us to assess the average within-subject regression parameters of interest and the between-subject variance in those parameters, as follows:

$$b_{1j} = b_{10} + u_{1j} \quad (4)$$
$$b_{2j} = b_{20} + u_{2j} \quad (5)$$

where b_{10} is the mean calibration effect for all participants (i.e., the mean p_{32} value), b_{20} is the mean additional knowledge effect (i.e., the mean p_{31} value), u_{1j} represents the variation in the calibration effect (i.e., variation in values of p_{32}), and u_{2j} represents the variation in influence of additional knowledge on the post-sampling conditional probability ratings (i.e., variation in values of p_{31}).