Available online at www.sciencedirect.com

SCIENCE DIRECT®

Consciousness and Cognition 13 (2004) 657–690

Consciousness and Cognition

www.elsevier.com/locate/concog

Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention

Jonathan Smallwood,[a,*] John B. Davies,[b] Derek Heim,[b] Frances Finnigan,[a] Megan Sudberry,[b] Rory O'Connor,[c] and Marc Obonsawin[b]

[a] *The Psychology Department, Glasgow Caledonian University, Glasgow, UK*
[b] *Department of Psychology, The University of Strathclyde, Glasgow, UK*
[c] *Department of Psychology, The University of Stirling, Stirling, UK*

Received 1 June 2004
Available online 17 July 2004

Abstract

Three experiments investigated the relationship between subjective experience and attentional lapses during sustained attention. These experiments employed two measures of subjective experience (thought probes and questionnaires) to examine how differences in awareness correspond to variations in both task performance (reaction time and errors) and psycho-physiological measures (heart rate and galvanic skin response). This series of experiments examine these phenomena during the Sustained Attention to Response Task (SART, Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). The results suggest we can dissociate between two components of subjective experience during sustained attention: (A) task unrelated thought which corresponds to an absent minded disengagement from the task and (B) a pre-occupation with one's task performance that seems to be best conceptualised as a strategic attempt to deploy attentional resources in response to a perception of environmental demands which exceed ones ability to perform the task. The implications of these findings for our understanding of how awareness is maintained on task relevant material during periods of sustained attention are discussed.

Keywords: Attentional lapses; Action slips; Task unrelated thought; Subjective experience; Awareness; Sustained attention

* Corresponding author. Present address: Psychology Department, University of British Columbia, Vancouver, Canada.
E-mail address: jonsmallwood2004@yahoo.com (J. Smallwood).

1053-8100/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.concog.2004.06.003

1. Introduction

When engaged on a task, particularly one that is repetitive in nature, one's thinking often drifts from information readily observable in the current environment. The attentional shift that accompanies the processing of internally generated information represents a situation in which the content of conscious awareness becomes, to some extent, decoupled from the processing of 'external' perceptual information: a state under laboratory conditions which can be operationalised as task unrelated thought (TUT, Smallwood, Obonsawin, & Heim, 2003a; Smallwood, Baraciaia, Lowe, & Obonsawin, 2003b; Smallwood, Obonsawin, & Reid, 2003c) or zoning out (Schooler, 2002; Schooler, Reichle, & Halpern, in press). A closely related phenomenon in psychological research, the notion of attentional lapses/action slips (Reason & Lucas, 1984) reflect a situation when "an action is triggered inappropriately... is targeted at the wrong stimulus... or when a plan becomes derailed by distraction" (Manly, Robertson, Galloway, & Hawkins, 1999, p. 661). Similar to the experience of TUT, action slips are attributed to situations of boredom and worry (Reason & Lucas, 1984; Robertson, Manly, Andrade, Baddeley, & Yiend, 1997). The research presented in this paper investigates whether states of subjective experience, such as TUT, contribute to the likelihood with which one experiences attentional lapses or action slips under laboratory conditions.

1.1. Task unrelated thinking

Over the last 30 years, TUT has been investigated by sampling the frequency with which an individual's attention departs from the current situation, a state that anecdotally reflects absentmindedness on the part of the individual. A variety of different approaches have been used to measure these drifts in attention and they broadly fall into two categories: (i) the thought probe technique, in which at various points in a task the thinking of the individual is probed and the sample is recorded either by verbal report (Smallwood et al., 2003a, 2003b) or via a button push (Giambra, 1995; Schooler et al., in press) and (ii) the self-report method in which the individual is primed with the purpose of the experiment and responds with a button push whenever they experience a TUT (Antrobus, 1968). The literature identifies a variety of external factors that moderate the frequency of TUT during vigilance. In particular, TUT is higher when: (i) the rate of stimulus presentation is slow (Antrobus, 1968; Giambra, 1995; Grodsky & Giambra, 1991), (ii) the frequency of targets is low (Giambra, 1995), and (iii) task duration is long (Antrobus, 1968; Cunningham, Scerbo, & Freeman, 2000; Giambra, 1995).

Moreover, the frequency with which these thoughts are experienced also depends upon internal factors. Perhaps the most consistent finding is that TUT depends upon the salience of one's current concerns (Klinger, 1978; Klinger, Barta, & Maxeiner, 1980). Current concerns reflect the "hypothetical process active during the time that one has a goal" (Klinger, 1999, p. 439). Antrobus and colleagues demonstrated that an experimental induction which induced a personally salient concern increased the likelihood of task unrelated thinking relative to a neutral control broadcast (Antrobus, Singer, & Greenberg, 1966). In addition, dysphoria can be conceptualised as a state in which current concerns are more salient (Higgins, 1987). In this context it is relevant that higher frequencies of TUT, or off task thinking, has been associated with: (i) positive and negative mood induction procedures (Seibert & Ellis, 1991) and (ii) dysphoria in variety of undergraduate samples (Smallwood et al., 2004; see also Lyubomirsky, Krasri, & Zehm, 2003).

1.2. Attentional lapses

The cognitive failures questionnaire (CFQ, Broadbent & Cooper et al., 1982) measures the frequency with which people fail to maintain attention on ongoing activity and historically was the only reliable measure of attentional lapses. Despite an ability to distinguish individual differences and to predict accidents mishaps and hospital admissions (Larson, Anderton, Neideffer, & Underhill, 1997) initial attempts to develop reliable behavioural correlates of the CFQ in the laboratory were disappointing (e.g., Broadbent & Broadbent et al., 1986). A recently developed task, however, the Sustained Attention to Response Task (SART, Robertson et al., 1997) has been shown to be sensitive to the CFQ.

The SART is similar in many respects to a standard vigilance task, in that a single infrequent target is presented amongst a background of frequent non-targets. Unlike a traditional vigilance task, however, the participant is required to push the space bar to the non-target and *inhibit their response to the target*. To perform this task correctly, the individual must remain "sufficiently attentive to their responses, such that at the appearance of a target they can substitute the directly antagonistic response" (Manly et al., 1998, p. 664). Errors on the SART are associated with a speeding of reaction time to the non-targets preceding an error, but only under conditions of low target probability (Manly et al., 1999; Robertson et al., 1997). Similar to research into TUT, the specificity of errors on the SART to situations of low target probability is attributed to the fact that a "high frequency of targets acts as external (exogenous) support to performance" (Manly et al., 1999, p. 668).

Under circumstances of low target probability, CFQ score has been demonstrated to predict SART performance (Manly et al., 1998; Robertson et al., 1997). These effects are conceptualised by the authors as demonstrating that errors on the SART reflect a lessening of active attention to the task (Robertson et al., 1997). In particular, the speeding of reaction time may be a consequence of the development of an "absentminded and insensitive approach to the task" (Manly et al., 1999, p. 669). Similarly, the 'Oops' phenomenon associated with errors, suggests that the detection of errors tends to re-direct "attentional resources" towards the task slowing responses to a level in which the alternative response can supervene (Manly et al., 1999).

Not only is the SART sensitive to action slips, a potentially important aspect of the SART is that it is possible to distinguish between the effects of different variables on: (i) reaction time in the lead up to an error and (ii) the frequency of errors. For example, Robertson et al. (1997) demonstrated that both cognitive failures score and Traumatic Brain Injury were associated with faster reaction times and differences in the number of errors of commission. By contrast, Manly, Lewis, Robertson, Watson, and Datta (2002) demonstrated that in healthy participants variations in circadian rhythms led to changes in the frequency of errors of commission, without any change in reaction time before such an error. The high discriminative power of the SART raises the possibility that when investigating differences in subjective experience associated with action slips, we may be able to identify the specific relationships between particular components of subjective experience and their respective relationships to RT and error commission described by Robertson and colleagues.

1.3. Experimental aims

Given the overlap between attentional lapses as measured by thought probes, on the one hand, and action slips on the other, the aim of these experiments was to examine subjective experience during the SART. Overall, we sought to explore whether two forms of subjective experience, broadly reflective of task engagement and disengagement, vary with respect to their contributions to SART performance. Subjective experience can be directed towards task relevant material, either in a mindful manner, or in terms of the appraisal of oneself or one's current performance. As a consequence, task focus can be considered to consist of two aspects: (i) attention can either be mindfully directed towards task completion or (ii) can be directed towards task re-appraisal (Task Related Interference, TRI, Sarason, Sarason, Keefe, Hates, & Shearin, 1986; see also Smallwood et al., 2003a, 2003b). Alternatively, attention can be disengaged from the task and is often directed to self-relevant material more or less unrelated to the current task: this form of subjective experience can be operationalised as TUT (Smallwood et al., 2003a, 2003b). Using these dimensions of subjective experience, we will investigate two claims regarding performance on the SART made in the literature: (i) the interpretation that the drift in RT in the lead up to errors on the SART is moderated by task disengagement (Manly et al., 1999; Robertson et al., 1997) and that (ii) errors on the SART are accompanied by the subsequent deployment of attentional resources towards the task (Manly et al., 1999).

The sustained vigilance task employed in these experiments is a modified version of the SART (Robertson et al., 1997). Since previous research demonstrated that during periods of fast stimulus presentation the frequency with which TUT is reported is low (Antrobus, 1968; Giambra, 1995), we modified the presentation pace of the SART task to approximately one stimulus every 2 s. By presenting a stimulus at this slower rate of presentation we anticipated that we would be able to employ a thought sampling methodology to investigate task disengagement in a reliable fashion. Experiment 3 tests this interpretation by contrasting subjective experience and performance on the SART under conditions of fast (one stimulus per second) and slow stimulus presentation (one stimulus every 2 s).

A subsidiary aim was to examine the relationship between subjective experience and on-going behavioural and physiological measures. This serves two purposes. From a methodological perspective, if we can identify a reliable association between task engagement/disengagement and a objective index, this would enhance our confidence in the validity and reliability of the construct being investigated. A reliable, physiological index for subjective experience, for example, will ultimately provide a more accurate indication of the theoretical status of self-report, either by validating current theories of subjective experience, or by revealing limitations which may not be apparent at the level of verbal reports alone. Second, the manner with which subjective experience is associated with these indices may shed light on the phenomenology associated with a given construct. In Experiments 1 and 2, we examine the relationship between subjective experience and two psycho-physiological measures: heart rate and galvanic skin response.

1.4. Statistical analysis

To examine subjective experience during the modified SART we employed the same form of statistical analysis that has been employed elsewhere (Smallwood et al., 2003b; Teasdale et al., 1995). In the first two experiments, we continuously measured physiological and behavioural variables in blocks lasting approximately 2 min. At the end of each of these blocks, we sampled the individual's thinking using thought probes (see Smallwood et al., 2003a, 2003b, 2003c). These verbal reports were subsequently classified as reflecting disengagement from the task (task unrelated thought, TUT) or broadly directed towards task completion (No Task Unrelated Thought, NTUT). This judgement was made using published criteria (Smallwood et al., 2003c) by judges blind to both the behavioural and physiological data. We then summated the relevant behavioural (i.e., mean reaction time) or physiological measure (i.e., galvanic skin response or heart rate) recorded during the last 30 s of each block. Using these classified verbal reports, we summated behavioural and physiological measures recorded over all of the blocks of a particular task into one of two categories: (i) those which preceded a report rated as task disengagement (i.e., TUT) and (ii) those in which the thoughts were rated as reflecting task engagement (i.e., NTUT). The indices recorded during periods of task disengagement can be compared with those recorded during the experience of task engagement using ANOVA.

In Experiment 3, we adapted the analysis employed by Robertson and colleagues to examine errors on the SART. Rather than sample subjective experience using thought sampling, we measured the frequency of TUT and TRI using a retrospective questionnaire (see below). Using these scores, we can examine the relationship between subjective experience and RT in the lead up to an error without explicitly monitoring thinking during the task. If using this approach we see a consistent relationship between the dimensions of subjective experience and SART performance, as is provided by the thought sampling methodology, we can be reasonably sure that the pattern of data is not an artefact of the method of thought sampling. Moreover, because the RT data is categorised on the basis of errors, rather than on the basis of verbal report, such a finding would enhance our confidence on the link between subjective experience and SART performance.

Individuals vary in the frequency with which they experience different aspects of subjective experience (Giambra & Grodsky, 1989; Grodsky & Giambra, 1991; see also Smallwood, Obonsawin, Reid, & Heim, 2002; Smallwood et al., 2004). On this basis we measured subjective experience at the end of the task using the Thinking Content component of the Dundee Stress State Questionnaire (DSSQ). This questionnaire measures the two components of subjective experience: (i) TUT and (ii) TRI. Using these scores individuals were classified into High and Low groups using the median score for each type of subjective experience for each independent experimental condition. These group variables are included in the analysis as between-participants variables, to provide an index of the frequency of subjective experience that complements the categorised behavioural and physiological data. For example, in Experiment 1, the analysis employs a Mixed ANOVA with repeated measures on two factors: Type of Experience (TUT/NTUT) and Type of Vigilance (successive versus simultaneous). This analysis is conducted twice, once with TUT Group (High/Low) included as a between participant factor, the second time with TRI Group (High/Low) included as a between participant factor. Finally to examine the role of individual differences in moderating subjective experience, in the final section of this paper we combine the data from the questionnaires gathered throughout these experiments. Using stepwise regression analysis we explore the dispositional, emotional, and contextual factors that contribute to subjective experience during sustained vigilance.

2. Experiment 1

2.1. Aims of Experiments 1 and 2

The aim of Experiments 1 and 2 was to use thought sampling to examine the claims of Robertson and colleagues (Manly et al., 1999, 2002; Robertson et al., 1997) that: (i) the phenomenology associated with task disengagement (i.e., TUT) corresponds to an acceleration in reaction time and (ii) following these errors the participant deploys strategic resources towards task completion (i.e., NTUT). In certain individuals, this strategic deployment will not reflect mindful task focus and instead will take the form of appraisal of the self/task (e.g., TRI, Matthews, Schwean, Campbell, Saklofske, & Mohammed, 2000).

For the purpose of Experiment 1, we varied the nature of the vigilance condition. Studies examining sustained attention (Parasuraman, 1979; Warm, Dember, & Hancock, 1996) have demonstrated that decrements are often strongest when a stimulus is compared to a standard representation (successive) held in working memory, relative to situations in which targets and non-targets are simultaneously presented (simultaneous). Successive vigilance tasks require the participant to "discriminate between a currently viewed stimulus and a *standard* representation of a specific stimulus held in working memory" (Desmond, Matthews, & Bush, 2001, p. 1386 emphasis added). In contrast, during a simultaneous vigilance task, "all of the necessary information to make the discrimination is presented in the current field" (ibid citation, 2001, p. 1386). The standard SART involves the presentation of a constant single target stimulus (often the digit 3) and so corresponds to successive vigilance. In Experiment 1, the manipulation we used was based on that employed by Parasuraman (1979) and involves the individuals detecting targets under two conditions: (i) under simultaneous vigilance, participants are presented with a target stimulus alongside a simultaneous non-target stimulus, (ii) under conditions of successive vigilance, participants are presented with two examples of a target stimulus. This manipulation makes a moderate working memory demand, although this demand is negligible relative to an N-back design. Such a subtle manipulation was deemed appropriate for the investigation of attentional lapses, which are likely to be highly dependent upon the strategic deployment of attentional resources by the participant.

In addition, in both experiments we examined the dynamic interaction between the effects of practice on the SART and changes in subjective experience. There are two reasons for expecting practice on task to moderate attentional lapses. Previous work using thought sampling (Smallwood et al., 2003c; Teasdale et al., 1995) indicates that time on task yields increases in the frequency of TUT. Moreover, the effects of the SART are attributed to the development of an automatic relationship between the stimulus and the non-target, described as "stimulus-press, stimulus-press" (Manly et al., 2002; Robertson et al., 1997). As practice is a key determinant of automaticity (Schneider & Shifrin, 1977), we should expect changes in errors on the SART with practice. Both literatures agree that attentional lapses are mediated by practice, and therefore, we examined the effects of practice in mediating the role of subjective experience on the SART task.

2.2. Method

2.2.1. Participants

Twenty-one participants were recruited from a University Psychology department, eight of whom were male and 13 were female. The mean age of the sample was 27.8 ($SD = 8.1$) years of age. All participants were paid £10 at the end of the experimental session.

2.2.2. Materials

Stimuli for both vigilance tasks were two squares presented either side of the centre of a computer screen and they were non-masked. These squares (approximately $16\,cm^2$) were either white or light grey and were presented against a black background. Regardless of vigilance task, stimuli were presented on the computer screen for 1500 ms. The inter-stimulus interval (ISI) was 2000 ms, during which time the screen was blank. Block duration was approximately 80 (±5) s and was terminated by the appearance of the word "STOP" in the centre of the screen.

In both the successive and simultaneous vigilance conditions, the non-target stimulus was two grey squares. For the successive vigilance task the target stimulus was two white squares, whilst during the simultaneous vigilance, the target stimulus was one white and one grey square. For the simultaneous vigilance condition the left-right position of the target relative to the non-target stimulus was counter balanced. Irrespective of condition, a total of 20 stimuli were presented in each block. Of these, four were targets (20%). The order of stimuli was randomised in each block. Each participant completed nine blocks of successive and simultaneous vigilance in a counter balanced design (ABAB).

2.2.3. Galvanic skin response (GSR) and heart rate (HR)

All electrode attachment sites were cleaned with SkinPure (Nihon Kohden) prior to the attachment of electrodes. Ag–AgCl skin conductance electrodes (6 mm, Biopac Systems) were attached to the palm side of the medial phalanx of the index and middle fingers of the non-dominant hand. The electrodes were secured with Velcro straps. The electrolyte used was a commercially available preparation (KY Lubricating Gel, Johnson and Johnson) with a conductivity similar to that of the 0.051 M NaCl solution recommended for use by Fowles et al. (1981) and Golding (1992). The signal from the electrodes were amplified with a Biopac model GSR100 amplifier.

The ECG was recorded using a three lead setup. Three surface Ag–AgCl electrodes (8 mm, BiopacSystems) were used, with one electrode placed on the palm side of the right and left wrists, and the third electrode placed on the anterior surface of the ankle ipsilateral to the dominant hand. The electrodes were secured with surgical tape (Blenderm, 3 M). The electrolyte used was Sigma Gel (Parker Laboratories). The signals from the electrodes were amplified with a Biopac model ECG100B amplifier, with a high-pass filter set at 1.0 Hz. The sampling rate for the ECG and the SCR channels was set at 500 Hz. Data acquisition and analysis were performed with Acknowledge software (Biopac Systems).

2.2.4. Questionnaires

Upon completion of the task, participants completed a battery of four questionnaires. (1) The Centre for Epidemiological Studies Depression Inventory (CES-D, Radloff, 1977), (2) The short form of the Response to Situations Questionnaire (RSQ, Nolen-Hoeskema, 1991), (3) Thinking Content and (4) the Mood component of the Dundee Stress State Questionnaire (DSSQ, Matthews, Joyner, Gililand, Campbell, & Faulconner, 1999). The CES-D is a measure of dysphoria successfully employed in a non-clinical sample in the past (Radloff, 1977; Smallwood, 2004; Smallwood et al., 2004) and contains 20 items which assess the frequency with which various self-descriptive terms can be applied to oneself over the last week. The CES-D is measured on a four point Likert scale and contains items such as "I was bothered by things which did not normally bother me." The short form of the RSQ is a 10-item measure of rumination that assesses the manner with which an individual responds when they are feeling depressed. It contains items such as "I think about how alone I feel" or "I think about how hard it is to concentrate." Each item is measured on a four point Likert scale.

The Thinking Content component of the DSSQ is a 16-item questionnaire that assesses the content of thinking during a recently completed task and it is divided into two, eight item factors (i) TRI (e.g., "I thought about how I should work more carefully" or "I thought about my level of ability") and (ii) TUT (e.g., "I thought about personal worries" or "I thought about something that happened earlier today"). For the sake of simplicity we will refer to these thoughts as retrospective TUT. Both factors (TUT and TRI) are measured on a five point Likert scale ranging from Never to Very Often. These factors, TUT and TRI were used to categorise the sample into High and Low TUT and TRI groups respectively.

The Mood component of the DSSQ consists of 29 adjectives (such as happy, nervous or tired) divided into three factors: Energetic Arousal, Tense Arousal and Hedonic Tone. The participant rates the extent to which each adjective describes how they felt whilst performing the task on a four point Likert scale. The mean scores for questionnaire responses relative to TUT and TRI Groups are summarised in Tables 1A and 1B. Independent *t* tests were employed to contrast group differences on these indices.

2.2.5. Procedure

Upon arrival, participants were greeted by the experimenter and seated in a comfortable seat in front of a computer. The experimenter outlined the experimental procedure and invited each participant to read and sign an informed consent sheet. Ethical approval had been obtained from the University Psychology Department's Ethics committee. Before beginning the experiment the participants completed a short questionnaire recording basic demographics.

Table 1A
Demographic characteristics of TUT groups for all three experiments

Experiment		1		2		3	
		M	*SD*	*M*	*SD*	*M*	*SD*
Age	Low	30.6	10.7	21.6	0.82	25.6	9.0
	High	25.5	4.45	20.8	0.75	24.1	4.4
CESD	Low	31.0	4.1	35.8	7.77	29.1**	8.4
	High	32.0	10.7	40.3	11.0	37.4**	10.8
RSQ	Low	23.7	13.9	21.8	2.6	18.7	5.0
	High	19	7.56	24.2	6.6	19.5	5.2
Mood							
Energetic	Low	18.1	5.3	19.3	4.0	21.2	4.0
	High	11.2	24.5	17.2	3.54	20.5	4.6
Tense	Low	16.1	4.65	16.8	4.9	16.7	4.7
	High	17.1	6.0	18	1.8	16.3	5.2
Hedonic	Low	24.6	4.53	26.2	4.2	22.2	8.6
	High	23.4	5.0	26.2	3.1	23.7	3.4
Thinking style							
TUT	Low	14.5***	1.8	22.7**	6.2	9.7***	1.4
	High	22.3***	4.8	26.2**	6.5	15.6***	4.44
TRI	Low	21.6	5.4	19	2.8	19.4	6.3
	High	23	5.0	24	1.0	21.7	3.6

** Significant group differences ($p < .01$).
*** Significant group differences ($p < .001$).

Table 1B
Demographic characteristics of TRI groups in all three experiments

Experiment		1		2		3	
		M	*SD*	*M*	*SD*	*M*	*SD*
Age	Low	27.8	5.6	21.7	0.82	27.6**	8.8
	High	27.8	9.9	20.8	0.75	22.0**	2.6
CESD	Low	28.9	4.1	37.5	12.6	32.1	6.8
	High	33.7	10.3	38.7	6.0	34.6	13.4
RSQ	Low	20.1	5.4	21.2	3.7	17.9	3.6
	High	22	14.2	24.7	5.7	20.4	6.1
Mood							
Energetic	Low	9.8	25.7	18.8	4.1	22.9	2.8
	High	18.7	4.0	17.7	3.8	18.5***	4.5
Tense	Low	16	7.2	16.5	4.60	16.8	4.5
	High	17.2	3.30	18.3	2.2	16.2	5.4
Hedonic	Low	24.7	5.46	26	3.6	22.5	7.9
	High	23.3	4.1	26.3	3.7	23.6	4.2
Thinking style							
TUT	Low	16.7	2.6	19.3	2.6	11.9	3.4
	High	20.4	6.53	29.5	4.5	13.4	5.2
TRI	Low	17.7***	2.3	21.2***	2.23	16.6***	3.1
	High	26.3***	2.9	21.8***	4.3	24.5***	3.6

** Significant group differences ($p < .01$).
*** Significant group differences ($p < .001$).

Five minutes before beginning the task electrodes to record GSR and heart rate were fitted to each participant. The electrodes for recording GSR were always attached to the participants' non-dominant hand. The participant was informed that their task was to detect a series of targets (white squares) from a non-target stimulus (light grey squares) both of which were presented sequentially on a computer screen. They were told that in some blocks they were going to be shown sequences of two grey squares interspersed by either: (i) two white squares (successive) or (ii) a white and a grey square (simultaneous). In either case, they were to push the space bar as quickly as possible, when they saw the non-target stimulus on the screen, and to do nothing when they saw a white square (target) on the screen. Before beginning the task, participants completed a short practice block of each type of vigilance task, including thought probes. During the testing procedure the participants was approximately 1 m from the computer screen, although consistent with research using the SART (Manly et al., 1999, 2002; Robertson et al., 1997) no restrictions were made on the participants' movement. Participants were asked to put equal emphasis on performing the task both quickly and accurately.

2.2.6. Thought probes

Before beginning the practice blocks participants were informed that throughout the course of this task they would be asked to report what was passing through their mind.

When you see the word STOP appear on the screen, I would like you to stop what you are doing and tell me exactly what was passing through your mind as you saw the word STOP. I do not want you to tell me what you were thinking about during the trial, just what was passing through your mind when you saw the word STOP.

2.2.6.1. Thought classification. Thoughts were recorded verbatim, onto a sheet, and later classified by the investigator and two judges blind to the hypothesis of the experiment. All classifications were made blind to all the physiological and behavioural data recorded. Thoughts were classified into whether they were directed towards task completion (NTUT) or not (Task Unrelated Thought, TUT). The definition of TUT reflects thoughts that are broadly directed to the self but bear no relationship to the task in hand, or the current situation (see Smallwood et al., 2003c). An example of TUT is "I was thinking about what I was going to do this evening" or "I was thinking of a meeting I have just had." The total number of recorded thought probes in each task was nine, and therefore, eighteen thoughts per individual were recorded for analysis. In the experiments presented in this paper we opted not to classify thought probe data in terms of the extent to which it was concerned with the appraisal of the self/task as in previous work (Smallwood et al., 2003a, 2003b, 2003c). For reasons of simplicity, a two-factor design (TUT/NTUT) was deemed more appropriate, yielding a more reliable distribution of behavioural and physiological data for subsequent analysis. Inter-rater reliability was calculated as described by Smallwood et al. (2003a, 2003b, 2003c). Two judges rated the thoughts independently and when disagreement occurred these issues were discussed until the disagreement was resolved. These agreed ratings were then compared with a third independent rater. The total number of thoughts for which the raters agreed was divided by the total number of thoughts. Inter-rater agreement was high (94%). In this experiment, the frequency of retrospective TUT measured via questionnaire was positively correlated with the classification of verbal reports as TUT during simultaneous [$r = +.55$, $p < .05$] and successive vigilance [$r = +.48$, $p < .05$] when controlling for retrospective TRI frequency.

2.3. Results

2.3.1. Distribution of task unrelated thinking

To compare the distribution of subjective experience, we examined the likelihood of reporting TUT in the first and second halves of each vigilance condition. Initial 2×2 ANOVA indicated that neither the effects of practice nor type of vigilance task had significant effects on the distribution of thinking ($p = \text{ns}$). A subsequent ANCOVA controlling for the retrospective frequency of TRI and the age of the participants indicated a significant effect of practice on the distribution of TUT [$F(1, 17) = 6.0$, $p > .05$]. Irrespective of the nature of the vigilance task, therefore, lower frequencies of TUT were reported in the first half [Mean = 0.42 ($SD = 0.06$)] than in the second half of the session [Mean = 0.46 ($SD = 0.05$)].

2.3.2. Behavioural measures

2.3.2.1. Reaction time. Effects of TUT. Reaction time (RT) to non-target stimuli over the last 30 s of the block was averaged and summated as described in the introduction. We contrasted RT using a Mixed $2 \times 2 \times 2$ ANOVA with repeated measures on Type of Experience (TUT/NTUT) and Task [simultaneous and successive vigilance]. TUT Group was included as a between participants' factor. The ANOVA indicated a Type of Experience by Vigilance task interaction

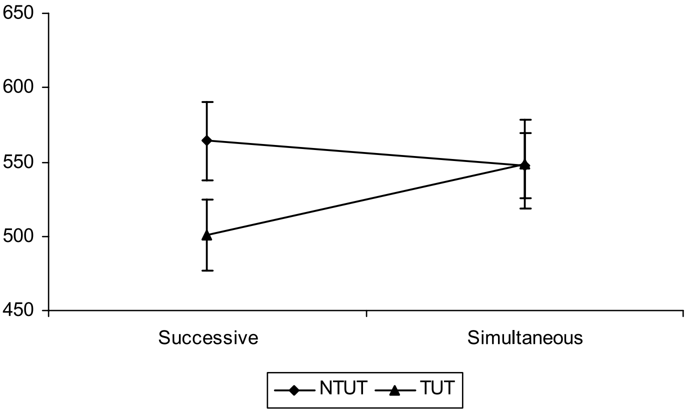

Fig. 1. Experiment 1. The effects of task engagement (NTUT) and disengagement (TUT) on mean RT (milliseconds) in two variations on the successive and simultaneous (SART) employed in Experiment 1.

[$F(1,18) = 4.80$, $p < .05$]. To follow up this interaction, separate repeated measures ANOVA were conducted to examine the effects of each type of Type of Experience in each task. Analysis indicated RT during TUT was significantly faster in the successive than in the simultaneous conditions [$F(1,18) = 7.26$, $p < .05$]. When subjective experience was directed towards the task, no differences in RT were observed between the tasks [$F(1,18) = .93$, $p =$ ns, see Fig. 1]. This is consistent with the notion that attentional lapses, when operationalised as TUT, contribute to performance on the SART by accelerating RT to the non-target stimulus (Robertson et al., 1997).

ANOVA on the effects of practice for the simultaneous vigilance task indicated a Type of Experience × Practice interaction [$F(1,12) = 4.7$, $p < .05$] demonstrating that RT decreased with practice in blocks during TUT [1st Half Mean = 556 (SD = 28), 2nd Half Mean = 539 (SD = 39)] and increased during NTUT [1st half Mean = 524 (SD = 32) and 2nd Half Mean = 539 (SD = 39)]. In the successive vigilance condition, initial 2 × 2 ANOVA indicated no reliable differences. Separate ANOVA indicated an effect of Type of Experience for the 2nd half of the task [$F(1,15) = 6.5$, $p < .05$, TUT Mean = 518 ms (SD = 36) and NTUT mean = 534 (SD = 34)]. No difference was observed in the first half of the task (p = ns).

Effects of TRI. ANOVA with TRI Group, included as a between participants' factor yielded no significant differences for RT scores (p = ns).

2.3.2.2. Errors of commission. Effects of TUT. A Mixed 2 × 2 × 2 ANOVA was used to contrast the distribution of errors of commission, i.e., when an individual fails to withhold a response to a target. Comparison of the frequency of errors of commission yielded a Type of Experience × TUT Group interaction [$F(1,18) = 8.16$, $p < .01$]. Errors of commission were distributed as follows: High TUT Group: Mean Errors of Commission NTUT = .03 (SD = .02), Mean Errors of Commission TUT = .10 (SD = .03). Low TUT Group, Mean Errors of Commission during NTUT = .10 (SD = .03), Mean Errors of Commission during TUT = .03 (SD = .03). This indicates that the High TUT group was more likely to make an error of commission during blocks in which TUT was experienced, whilst the low TUT group was most likely to make an error of commission in blocks in which NTUT was reported. Analysis indicated no effects of practice were reliable.

Effects of TRI. ANOVA indicated a reliable effect of TRI group [$F(1,18) = 5.16$, $p < .05$]. Overall, the Low TRI group made fewer errors [Mean = .03 ($SD = .02$)] than the High TRI Group [Mean = .10 ($SD = 0.02$)]. No other main effects or subsequent interactions reached significance (p = ns).

2.3.3. Physiological measures

2.3.3.1. Heart rate. Effects of TUT. Acquire was used to derive Heart Rate (HR), measured in beats per minute, from the raw ECG data. Mean HR was averaged and analysed in the same manner as RT and errors of commission. The 2 × 2 × 2 Mixed ANOVA indicated two effects: (i) a vigilance task × TUT Group interaction [$F(1,18) = 4.7$, $p < .05$] and (ii) a reliable three-way interaction between Type of Experience, TUT Group and Vigilance task [$F(1,18) = 6.6$, $p < .05$, see Fig. 2]. This interaction was followed up examining the heart rates recorded separately for periods of task engagement and disengagement. A Task × TUT Group interaction was observed for periods of task disengagement [$F(1,18) = 8.1$, $p < .01$, see Fig. 2A]. The difference between HR during task disengagement during each task varied with TUT group [$F(1,18) = 7.5$, $p < .05$]. By contrast no effects of task or a TUT Group interaction reached significance for periods of task engagement (p = ns, see panel B).

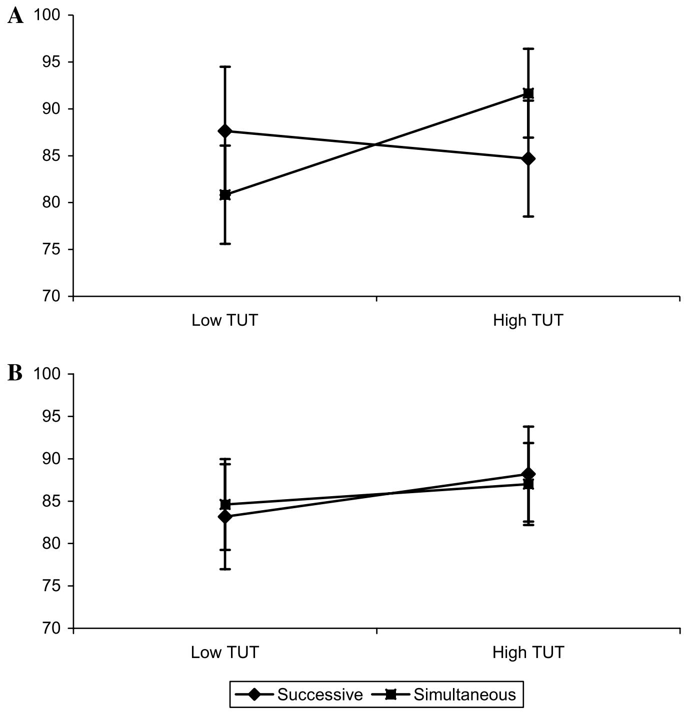

Fig. 2. Experiment 1. The effects of TUT Group (High and Low) in two versions of the successive and simultaneous (SART) on Heart Rate (BPM) recorded during periods of: (A) task disengagement (TUT) and (B) task engagement (NTUT).

ANOVA on the effects of practice in the simultaneous vigilance task confirmed the Type of Experience × TUT Group interaction [$F(1,15) = 5.61$, $p < .05$] described above. In the successive vigilance task, a TUT Group × Practice interaction was observed [$F(1,11) = 9.2$, $p < .01$]. Subsequent analysis yielded a main effect of TUT group [$F(1,11) = 9.2$, $p < .01$] implying that the biggest increase in HR as the task proceeded was in the High TUT group [Increase in HR: High TUT Group = 3.9 BPM ($SD = 1.3$), Low TUT group = −1.9, $SD = 1.4$].

Effects of TRI. ANOVA on the distribution of HR indicated no significant main effects or any subsequent interactions (p = ns).

2.3.3.2. Galvanic skin response (GSR). Effects of TUT. As for HR, Acquire was used to convert Raw SCR data into peak to peak changes over the last 30 s of each block (Galvanic Skin Response, GSR). The Mixed 2 × 2 × 2 ANOVA indicated a main effect of TUT group on GSR [$F(1,17) = 9.4$, $p < .01$] indicating that overall the High TUT group demonstrated lower GSR over the last 30 s of a block (Mean = .23, $SD = .06$)] than the Low TUT group [Mean = 0.51 ($SD = .06$)]. This confirms that the verbal reports of TUT are indicative of task disengagement. No effects of type of experience or any other interactions were significant.

The analysis of practice confirmed that GSR was lower in the successive task in the High TUT group [$F(1,7) = 15.16, p < .01$]. Analysis of the simultaneous task indicated no overall differences (p = ns). Separate ANOVA on each half of the task, however, indicated a reliable effect of TUT Group on the second half of the task [$F(1,13) = 5.2$, $p < .05$: High TUT GSR = .22 ($SD = 0.07$), Low TUT = .46 ($SD = 0.07$)], implying the High TUT group showed smaller increases in GSR with practice.

Effects of TRI. As for Heart Rate, ANOVA on the distribution of GSR indicated no significant main effects or subsequent interactions (p = ns). No reliable effects of practice were observed (p = ns).

2.3.4. Discussion of Experiment 1

The results of Experiment 1 are broadly consistent with the suggestions of Robertson and colleagues regarding subjective experience during the SART: (i) blocks of successive vigilance in which TUT were reported were associated with faster RT (see Fig. 1) and (ii) high frequency of errors on the SART was associated with subsequent verbal reports in which attention was directed towards the appraisal of the task (e.g., TRI). In addition, individuals who reported high frequencies of TUT, and can be considered absent minded, had: (i) higher heart rates in blocks in which TUT was reported in the simultaneous vigilance condition, (ii) demonstrated lower GSR over the task as a whole, combined with smaller increases in GSR as the task proceeds, and (iii) made the majority of their errors of commission during blocks in which the verbal report was classified as TUT.

One possible interpretation of the relationship between TUT and errors in Experiment 1 is that the experience of errors on the SART re-directs the individual's attention towards the task, the so-called 'Oops phenomenon' (Manly et al., 1999). If the individual is aware that they have made a mistake, then over the course of the rest of the block, they are likely to devote strategic resources towards task completion, thereby reducing the likelihood that TUT is reported at the subsequent thought probe. The co-variation of errors with the GSR level in the High TUT group indicates that this interpretation does not rely on self-report alone, because high GSR levels are often ascribed to the deployment of effort towards the task, whilst lower levels of GSR are conceptualised as indexing task dis-engagement (Pechineda & Smith, 1996).

It is plausible that high level of errors of commission during TUT blocks and lower GSR scores over the task as a whole recorded in the High TUT group indicate that these individuals may have failed to detect that they made an error. To test this interpretation of the results of Experiment 1, we conducted a second experiment using the same experimental framework. In Experiment 2, we moderated the stimulus duration of the targets. Errors of commission for targets presented on the screen for a short duration are likely to be less available to awareness than similar errors for targets with a longer duration. This would provide an opportunity to test the notion that *awareness* of errors on the SART might disrupt the experience of TUT, and encourage the individual to attend to the task.

3. Experiment 2

3.1. Aims of Experiment 2

The aims of Experiment 2 was to examine whether the relationship between subjective experience and errors of commission depends upon awareness. To test this notion, we included targets with two durations (short and long). Presumably, the shorter the target is presented on the screen, the less aware the individual will be that their response was correct or incorrect. It is clear from Fig. 1 that the under conditions of successive vigilance RT was faster during periods of task disengagement, and therefore, to facilitate the investigation of awareness, this experiment focused on conditions of successive vigilance.

3.2. Methods

3.2.1. Participants

A further 12 participants were recruited from a University Psychology department, four of whom were male and eight were female. The mean age of the sample was 21.25 ($SD = 0.9$) years of age. As before, all participants were paid £10 at the end of the experimental session.

3.2.2. Methods

All participants completed eighteen blocks of the same successive vigilance condition of Experiment 1. In Experiment 2, half of the stimuli were of shorter duration than in Experiment 1 (short targets, 1000 ms) and half the stimuli were of the same duration as in Experiment 1 (long targets, 1500 ms). All blocks contained 20 stimuli, of which two targets of each type were presented. The total number of targets, therefore, was four, ensuring that the proportion of targets/non-targets was held constant between Experiments 1 and 2.

Thought probes and questionnaires. These were described and administered in the same fashion as in Experiment 1 (see Tables 1A and 1B).

3.3. Results

3.3.1. Distribution of TUT

As in the previous experiment agreement amongst the raters was high (93%). An initial ANOVA indicated no significant differences in the distribution of TUT across the first and second halves of

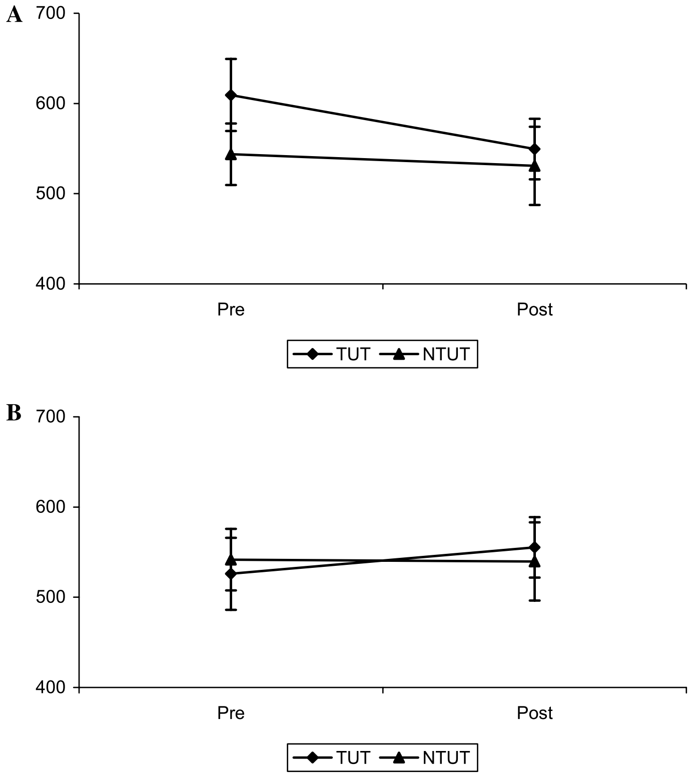

Fig. 3. Experiment 2. The role of practice, type of experience and TUT group (High and Low) on Mean RT during the successive SART condition employed in Experiment 2. (A) The reaction time during in the High TUT Group and (B) the reaction time during the Low TUT group.

the task. Subsequent analysis of the distribution of TUT, controlling for the retrospective TRI approached significance in the predicted direction [$F(1,10) = 4.07$, $p > .05 < .07$][1] implying that TUT was lower in the first half of the task [Mean = .36 ($SD = .04$)] than in the second [Mean = .40 ($SD = .06$)]. This broadly replicates Experiment One, and suggests that similar to other non-demanding tasks, TUT shows a moderate increase with practice during the SART.

3.3.2. Behavioural measures

3.3.2.1. Reaction time. The Effects of TUT. RT was summated in the same fashion as in Experiment 1 and contrasted using a $2 \times 2 \times 2$ Mixed ANOVA with repeated measures on Type of Experience (TUT/NTUT) and Practice (1st and 2nd halves). TUT Group was included as a between participants factor. ANOVA indicated a Type of Experience × TUT group interaction [$F(1,10) = 5.8$, $p < .05$] with higher RT during task disengagement in the high TUT group. It is clear from Fig. 3 that RT was highest during the first half of the task. Follow-up analysis

[1] As the participants in Experiment 2 were all recruited from an undergraduate course with an age range of 20–22, we did not include age in the ANCOVA for the investigation of the effects of practice on TUT.

confirmed this and demonstrated that the effects of TUT groups were reliable for the first half of the task [$F(1, 10) = 5.41$, $p < .05$] not the second half [$F(1, 10) = .129$, $p > .05$]. Moreover, separate analysis of the effect of TUT group on the effects of practice approached significance for blocks in which TUT was reported [$t(10) = 2.10$, $p = .062$]. By contrast, the blocks in which task focus were reported showed no reliable group differences [$t(10) = 0.08$, $p =$ ns]. In the context of Experiment 2, whilst the TUT group initially had slower RT, the effect of practice in those individuals reporting high levels of absentmindedness was to accelerate RT to the non-target stimulus whenever subjective experience was directed away from the task.

The effects of TRI. ANOVA indicated that neither main effects nor any subsequent interactions were reliable for the RT data ($p =$ ns).

3.3.2.2. Errors of commission. Effects of TUT. The likelihood of making an error was contrasted using a $2 \times 2 \times 2 \times 2$ Mixed ANOVA with repeated measures on target duration [short and long], in addition to Type of Experience and Practice. TUT group was included as a between participants' factor. ANOVA indicated a significant four-way interaction: Target Duration × Type of Thought × Practice × TUT Group [$F(1, 10) = 12.12$, $p < .01$]. Separate ANOVA for each target duration indicated that errors of commission for targets of long duration were low and did not vary with Practice or Type of Experience. Moreover, whilst The High TUT group made fewer errors of commission in blocks in which TUT was reported [Mean = .04 ($SD = 0.02$)] than the Low TUT group [Mean = .08 ($SD = 0.02$)], this difference was non-significant ($p =$ ns).

By contrast, ANOVA on targets with a short duration revealed an effect of practice [$F(1, 10) = 6.38$, $p < .05$] indicating that in the first half of the task, errors of commission [Mean = 0.02 ($SD = 0.08$)] were significantly fewer than in the second half of the task [Mean = 0.11, ($SD = .03$)]. In addition, a TUT Group × Type of Experience × Practice interaction was significant [$F(1, 10) = 5.95$, $p < .05$, see Fig. 4A]. Subsequent analysis indicated that in the Low TUT group, errors of commission during NTUT blocks showed a significant increase with time on tasks relative to the TUT blocks [$F(1, 5) = 22.0$ $p < .001$]. In the High TUT group, however, there was no difference between the effects of practice in either TUT and NTUT blocks [$F(1, 5) = .66$, $p =$ ns]. The analysis indicated that across the sample as a whole errors of commission for targets of short duration increased with practice. The difference between High and Low TUT groups was specific to blocks in which TUT was reported and indicated that for targets of short duration, only individuals who reported high levels of absentmindedness (the High TUT group) showed an increase in the proportion of errors of commission as the task proceeded under these circumstances.

Effects of TRI. The initial ANOVA indicated that no main effects or subsequent interactions were statistically reliable. Separate ANOVA for targets of each duration indicated a reliable Type of Experience × TRI group interaction for targets of a long duration [$F(1, 10) = 9.90$, $p < .01$]. Errors of commission for long targets were distributed as follows: Low TRI NTUT Mean = 0.02 ($SD = 0.02$), TUT Mean = 0.09 ($SD = 0.03$), High TRI NTUT = 0.06 ($SD = 0.03$), and TUT Mean = 0.03 ($SD = 0.02$). The difference between errors of commission during task disengagement and task engagement varied with TRI group [$t(10) = -3.1$, $p < .01$] confirming that a higher proportion of errors of commission were made by the High TRI group in blocks in which subsequent verbal reports indicated that attention was directed towards the task. By contrast, ANOVA on errors of commission for the short targets indicated an effect of practice only

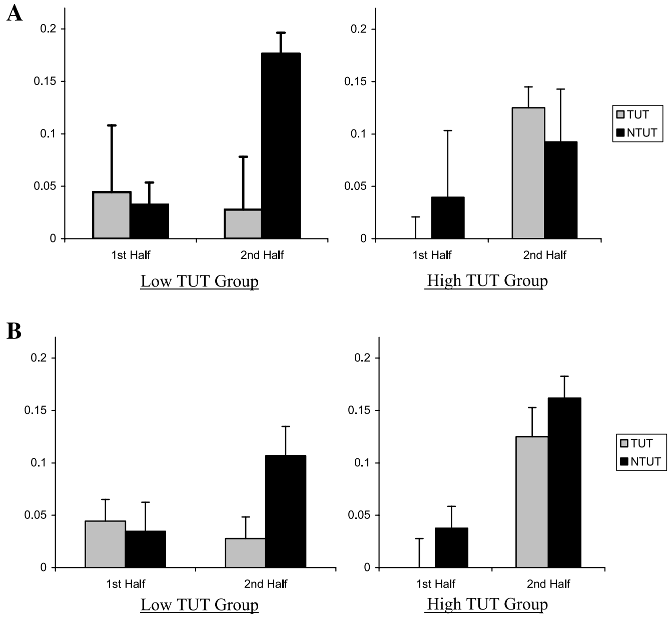

Fig. 4. Experiment 2. Contrasting effects of TUT and TRI on errors of co-mission on targets with a short duration (500 ms) detected in Experiment 2. (A) The effects of TUT group and (B) the effects of TRI group.

[$F(1,10) = 6.0$, $p < .05$, see Fig. 4B]. No other main effects or subsequent interactions were statistically reliable.

3.3.3. Physiological measures

3.3.3.1. Heart rate. Effects of TUT. Heart rate was contrasted using a $2 \times 2 \times 2$ Mixed ANOVA. Overall HR was higher in blocks in which TUT was reported than in blocks in which NTUT was reported [$F(1,9) = 8.0$, $p < .01$, Table 2]. In addition, a Practice × Type of Experience × TUT group interaction was observed [$F(1,9) = 8.2$, $p < .05$]. We followed this interaction up by subtracting HR during TUT from NTUT separately for each half of the task. Subsequent analysis that in the High TUT group [$F(1,5) = 8.2$, $p < .05$] TUT yielded larger increases in HR for the second half of the task [Mean = −2.0 ($SD = .7$)] than in the first half [Mean = .09 ($SD = .39$)]. No other comparisons reached significance [$p =$ ns].

Effects of TRI. ANOVA indicated an effect of type of experience on HR [$F(1,10) = 10.52$, $p < .01$] indicating significantly greater HR during TUT than during NTUT. No other significant main effects or interactions were significant.

3.3.3.2. Galvanic skin response. Effects of TUT. ANOVA indicated a Type of thought × Practice interaction [$F(1,9) = 6.87$, $p < .05$, see Table 2]. Subsequent analysis indicated that GSR increased with practice in the blocks in which NTUT was reported [$F(1,9) = 9.2$, $p < .01$]. By contrast, there was no difference in GSR as the task proceeded in the blocks in which TUT was reported [$F(1,9) = .102$, $p > .05$].

Table 2
Experiment 2. The effects of TUT group, type of experience practice and TUT group on GSR and HR rate

			Low TUT group		High TUT group		Overall sample	
			Mean	*SD*	Mean	SD	Mean	*SD*
Heart rate (HR)	1st Half	TUT	81.1	5.4	80.7	5.0	80.9	3.5
		NTUT	80.0	5.3	80.6	4.8	80.3	3.4
	2nd Half	TUT	80.3	4.5	82.9	4.2	81.7	3.0
		NTUT	80.2	4.8	81.0	4.4	80.6	3.1
	Mean	TUT	80.6	5.3	81.9	4.9	81.3	3.2
		NTUT	80.2	4.7	80.6	4.3	80.5	3.2
Galvanic skin response (GSR)	1st Half	TUT	0.12	0.04	0.12	0.03	0.12	0.02
		NTUT	0.08	0.02	0.09	0.02	0.09	0.01
	2nd Half	TUT	0.10	0.03	0.11	0.03	0.11	0.02
		NTUT	0.10	0.02	0.12	0.02	0.11	0.01
	Mean	TUT	0.10	0.03	0.11	0.03	0.11	0.02
		NTUT	0.10	0.03	0.12	0.02	0.10	0.01

Effects of TRI. ANOVA indicated a Type of Experience and Practice interaction [$F(1,9) = 6.7$, $p < .05$, see Table 2]. No other effects reached significance.

3.3.3.3. Correspondences between measures of task performance. Robertson et al. (1997) suggest that the likelihood of making an error on the SART depends upon RT accelerating towards a non-target as their attention drifts away from the task. We examined this claim by examining the correlations between RT and errors of commission separately for those blocks in which: (i) the verbal reports were indicative of task dis-engagement (TUT) and those indicating task engagement (NTUT) and (ii) for the successive and simultaneous vigilance tasks. To enhance discrimination, for the successive vigilance task we combined the data from Experiments 1 and 2. These correlations are reported in Table 3. The pattern of correlations indicated the predicted negative association between errors of commission and RT. Further analysis indicated that the reaction time for those individuals who only took part in the first experiment was significant and in the same direction [$r = -.48$, $p < .05$], suggesting that the smaller sample size for the simultaneous vigilance task was not responsible for the lack of a reliable association. Moreover, to rule out the possibility that the correlations were skewed by outliers non-parametric correlations confirmed

Table 3
Experiments 1 and 2. Pearson correlations demonstrating the relationship between reaction time and errors of commission during the simultaneous and successive SART

	Mean reaction time			
	Simultaneous SART ($n = 22$)		Successive SART ($n = 32$)	
	TUT	NTUT	TUT	NTUT
Mean errors				
r	.21	.04	−.41	−.24
p	.38	.88	.02*	.19

* Significant association ($p < .05$).

that RT and errors showed a negative relationship during disengagement during blocks in which TUT was reported [$r = -.36$, $p < .05$]. Overall, therefore, the association between RT and errors was only reliable for: (i) conditions of successive vigilance and (ii) in blocks in which the subsequent verbal report indicated that attention was directed away from the task. The pattern of correlations is consistent with the interpretation of the SART data suggested by Robertson et al. (1997) that the relationship between RT and errors of commission is moderated by task disengagement.

3.3.4. Discussion of Experiments 1 and 2

Taken together the results of Experiments 1 and 2 indicate that practice on the SART is associated with a higher frequency of verbal reports reflecting task disengagement. Generally, TUT was associated with higher HR, either in terms of group differences in the effects of practice, or at a block-by-block basis (Fig. 2 and Table 2). In addition, task engagement was associated with high levels of GSR, and this relationship was expressed either in terms of the low level of GSR in the high TUT group (Experiment 1) or the changes in GSR with practice in Experiments 1 and 2. Finally, TUT was associated with accelerations in RT and, moreover, the summated data from Experiments 1 and 2 demonstrated the predicted negative association between RT during TUT blocks and errors of commission (Table 3). It is important to note that many of these effects were specific to blocks in which subsequent verbal reports were indicative of task disengagement, and so do not represent crude, context independent dispositional characteristics of the relevant individuals. Rather, they plausibly reflect time limited changes in the coupling of the cognitive/affective, behavioural and physiological systems of these individuals.

We introduced targets of shorter duration into Experiment 2 to examine whether the relationship between subjective experience and errors of commission was moderated by awareness. Following the results of Experiment 1 we hypothesised that awareness of an error of commission on the modified SART, the 'Oops' phenomenon, might serve to: (i) interfere with the maintenance of material in working memory unrelated to the current situation and (ii) to encourage the individual to focus on task relevant information. In some individuals this might take the form of the appraisal of one's task performance (i.e., TRI).

Consistent with this aim, the High TUT Group represented the only individuals who made a similar frequency of errors on the short targets during task disengagement and engagement. This suggests that during periods of task disengagement high levels of absentmindedness were associated with errors on targets with shorter durations. By contrast, across all participants verbal reports of task focus were associated with increases in GSR and decreases in target detection for short targets. Moreover, the High TRI group contained individuals who were more likely to make an error of commission for long targets in blocks in which the subsequent verbal report indicated attention was directed towards the task. Taken together, these findings help disambiguate the effects of effort during task engagement and absentmindedness during periods of disengagement in the context of the SART. In particular they support the notion that the 'Oops' phenomenon associated with an error on the SART is associated with the subsequent redeployment of effort towards the task. Whilst our evidence that errors on the short targets were less available to awareness is broadly circumstantial, it is certainly consistent with the suggestions of Manly et al. (1999) that we can distinguish between the two forms of subjective experience (TUT and TRI) in terms of their respective relationships to performance on the SART.

4. Experiment 3

4.1. Aims

The final experiment seeks to address several possible experimental confounds in the previous studies presented in this paper. The most pressing concern is that it is possible that by asking participants to monitor their thoughts, we may have been providing tacit endorsement for them to dissociate their attention from the task. This may not necessarily occur because the individuals strategically decide to do so, it may occur because the methodology raises the participant's accessibility of their own thinking, irrevocably changing the subjective experience (see Wenzlaff & Wegner, 2000, for a discussion of this issue; see also Nisbet & Wilson, 1977). In this context, instruction to monitor one's own thinking may produce a situation that differs from a standard vigilance situation. Any reliability in the experimental findings, therefore, may result from the consistent yet artificial situation which thought monitoring encourages.

Experiment 3, therefore, examines the role of drifts in attention using a SART task in the absence of the explicit instructions to monitor thinking during task performance. In this experiment, subjective experience is measured by retrospective questionnaire at the end of the task and no mention is made to the participant before or during the task regarding the specific aims and objectives of the investigation. The rationale behind this experiment is that we can categorise the behavioural data, e.g., RT, on the basis of errors of commission rather than via thought probes. The results of previous experiments would suggest that participants who score highly on the retrospective questionnaires measuring TUT should show an acceleration of RT in the lead up to an error with no differences in RT across the task as a whole (e.g., Fig. 1 and Table 3). Similarly, differences in TRI frequency should be expressed as higher levels of error.

In Experiment 3, participants complete the SART under one of two conditions: (i) one stimulus each second (fast, see Robertson et al. (1997)) and (ii) one stimuli every 2 s (slow). As noted earlier, studies that employ thought sampling indicate that TUT occurs more frequently in situations in which stimuli are presented slowly (Antrobus, 1968; Giambra, 1995). On this basis we can test our assumption that TUT will affect RT preceding errors on the SART only under slow stimulus presentation conditions. We made an additional minor methodological change. To ensure that we have recorded as many examples of TUT as would allow us to examine the consequences of TUT on task behaviour and physiology, we have employed tasks which lasted approximately 40 min. In all of these comparisons, therefore, it is possible that the effects attributed to TUT are confounded by the effects which could be otherwise attributable to experimental fatigue. This task is shorter (10 min) and should allow us to rule out this interpretation.

4.2. Methods

4.2.1. Participants

Forty-one Participants were recruited from a University Psychology department, 12 of whom were male and 29 who were female. The mean age of the sample was 24.8 ($SD = 7.3$) years of age. Due to the shorter nature of this task (10 min), in this experiment all participants were paid £5 at the end of the experimental session. As in previous experiments, the mean scores for the each questionnaire measure relative to the TUT/TRI groups are presented in Tables 1A and 1B. In addition to the questionnaires employed in previous studies (Experiments 1 and 2), we included the CFQ (Broadbent et al., 1982). CFQ score varied with TUT Group status [$t(36)^2 = -3.2$, $p < .001$] with higher CFQ scores in the High TUT Group [Mean = 48.0, $SD = 10.6$] than in the Low TUT Group [Mean = 35.0, $SD = 8.8$], confirming our assumption that TUT relates to absentmindedness. Neither TRI Group nor Pacing condition was associated with reliable differences in CFQ score ($p =$ ns).

4.2.2. Procedure

As in the previous experiments, testing took place on an individual basis. Unlike previous experiments, participants performed the tasks alone. Before beginning the task participants read the following set of instructions, based on the instructions described in Robertson et al. (1997).

"You will see a series of digits (0–9) appear on the computer screen. We want you to respond as quickly as possible by pushing the space bar when you see a digit appear on the screen. Respond to all digits except the digit '3' by pushing the space bar with your preferred hand. When the digit '3' appears on the screen please DO NOT PUSH THE SPACE BAR.

We want you to give equal weight to responding quickly as possible to the stimulus, and also to minimising errors.

The task will be divided into seven sections. At the end of each section you can take time to have a break. The first block is a practice session. Press the space bar when you have read and understood the instructions and are ready to proceed with the task."

4.2.3. Stimuli

Stimuli for the fast condition were presented on screen for 250 ms with an Inter Stimulus Interval (ISI) of 950 ms (see Robertson et al., 1997). Stimuli in the slow condition were presented for 250 ms and in this case the ISI was 2050 ms. In both tasks stimuli were unmasked and during the interval between stimuli the screen was blank. Stimuli were presented in seven blocks. The first block was a practice block containing 20 stimuli, two of which were targets. In the slow condition, blocks contained 40 stimuli. In the fast condition, blocks contained 80 Stimuli. In both fast and slow conditions blocks lasted approximately 80 s and alternated (ABABAB) between a Low probability condition (LP, 10% targets) and a High Probability condition (HP, 50% Targets). Block order was counterbalanced.

4.3. Results

4.3.1. Distribution of subjective experience

The distribution of subjective experience was as follows: TUT fast Mean = 11.50 ($SD = 3.6$), slow Mean = 13.75 ($SD = 4.8$). TRI fast Mean = 20.3 ($SD = 6.5$) and TRI slow Mean = 20.8 ($SD = 3.6$)]. Separate uni-variate ANOVA indicated that neither TUT nor TRI varied between the fast and slow versions of the modified SART task ($p =$ ns).

[2] Three individuals failed to complete the CFQ.

4.3.2. Overall RT to non-targets[3]

Effects of TUT. A $3 \times 2 \times 2 \times 2$ ANOVA was used to contrast the RT to all correctly detected non-targets with repeated measures on two factors: Practice (Blocks 1, 2, and 3) and Target Probability (HP and LP). TUT Group (High and Low) and Stimulus Presentation Rate (fast and slow) were included as between participant factors. This analysis indicated a main effect of Probability $[F(1,26) = 70.6,\ p < .001]$ indicating that RT was faster in the LP condition [Mean = 373 ($SD = 9.9$)] than in the HP condition [Mean = 416 ($SD = 8.1$)]. This main effect was clarified by a subsequent Probability × Stimulus Presentation interaction $[F(1,26) = 11.51,\ p < .001]$ indicating that the difference between the LP and HP conditions was greater for the fast condition [LP Mean = 351 ($SD = 14.1$) and HP Mean = 411 ($SD = 11.6$)] than for the slow condition [LP Mean = 395 (14.1) and HP Mean = 421 ($SD = 11.6$)]. Similar to Experiments 1 and 2, TUT Group was, therefore, not associated with gross changes in overall RT to non-targets.

Effects of TRI. In addition to the effects of Probability $[F(1,62) = 86.9,\ p < .001]$ and the Pace × Probability interaction $[F(1,62) = 13.13,\ p < .001]$ ANOVA indicated a main effect of TRI group $[F(1,31) = 6.12,\ p < .05]$. In addition, a Probability × TRI Group interaction was marginally significant $[F(1,62) = 3.13,\ p > .05 < .09]$. We followed this interaction up separately in each pacing condition to examine whether the effects of TRI on RT were moderated by stimulus presentation pace. This analysis allows us to examine whether the effects of TRI in the slow condition parallel those observed in Experiments 1 and 2. No main effect of TRI $[F(1,32) = .810, p = \text{ns}]$ nor the Probability × TRI interaction $[F(1,32 = .07)\ p = \text{ns}]$ was reliable for the RT in the slow presentation condition ($p = \text{ns}$). By contrast, in the fast condition the TRI Group × Probability interaction was reliable $[F(1,32) = 4.56,\ p < .05]$ indicating that under conditions of fast stimulus presentation the High TRI group showed faster RT during LP [High TRI LP Mean = 314 ms ($SD = 16$), HP Mean = 385 ($SD = 17$), Low TRI LP Mean = 387 ($SD = 14$) and HP Mean = 428 ($SD = 16$)].

4.3.3. Reaction time preceding an error

The analysis described by Robertson et al. (1997) demonstrated that the RT over the last four stimuli preceding an error of commission distinguished CFQ score. As we varied the presentation rate of stimuli between participants, this analysis would confound time preceding the error with the number of stimuli. To account for this difficulty we conducted two separate pieces of analysis: (i) to examine the distribution of RT over a constant period of time (4 s) and (ii) to examine the pattern of RT over a constant number of stimuli (eight).

4.3.3.1. Constant time. For the purpose of analysis over a constant stimulus period, RT was summated in the following manner. The first four RT and the subsequent four RT were summated from the fast condition, reflecting the respective mean RT in the first four and second four stimuli preceding an error. In the slow condition, the first two stimuli and the subsequent two stimuli occurred in the same temporal window, and were therefore summated. We compared RT over the eight second window preceding an error of commission using a $2 \times 2 \times 2 \times 2$ Mixed ANOVA with

[3] Due to a hard drive failure after the completion of data collection the raw data was lost from five individuals. This data had already been analysed for the examination of RT proceeding errors.

repeated measures on the period of time before the target (2 and 4 s) and Probability condition. Stimulus Presentation [fast and slow] and TUT group were included as between participant factors. This ANOVA indicated two reliable effects. First, a Pace × Probability Interaction was observed [$F(1,30) = 8.50$, $p < .01$]. Follow-up analysis indicated that there was no difference between RT preceding an error in the HP condition according to stimulus presentation pace ($p =$ ns). By contrast, RT in the LP conditions was significantly quicker in the fast condition [Mean = 315 ms ($SD = 8$)] than in the slow condition [Mean 350 ms ($SD = 8$), $F(1,32) = 9.32$, $p < .01$].

Second, the ANOVA demonstrated a Pace × TUT Group interaction [$F(1,32) = 9.31$, $p < .001$] indicating that, irrespective of probability of target, High TUT group was associated with faster RT in the slow condition than the Low TUT group [$F(1,14) = 4.9$, $p < .05$, Low Probability, High TUT 323 ms, Low TUT 375 ms; High Probability, High TUT 377 ms, Low TUT 400 ms]. By contrast, no difference was observed as a result of TUT group in the fast condition [$F(1,16) = 1.9$, $p > .05$, Low Probability: High TUT 323 ms, Low TUT 307 ms; High Probability High TUT 419 ms, Low TUT 377 ms].

Effects of TRI. ANOVA indicated that TRI group made no difference to RT preceding an error in either fast or slow conditions, nor in high or low probability conditions ($p =$ ns).

4.3.4. Constant stimuli

Effects of TUT. RT for both fast and slow conditions were calculated by summating the RT into the following windows: (i) 1st and 2nd (ii) the 3rd and 4th (iii) 5th and 6th, and (iv) 7th and 8th. A 4 × 2 × 2 × 2 Mixed ANOVA compared the RT to the last eight stimuli preceding an error. Number of stimuli preceding an error (8, 6, 4, and 2) and target probability (Hp and LP) were included as within participant factors. Stimulus presentation rate (fast and slow) and TUT Group (High and Low) were included as between participant factors. ANOVA indicated the following reliable effects. First, an effect of the number of stimuli preceding the error was observed [$F(3,96) = 20.0$, $p < .01$]. This was, however, subsequently clarified by an interaction between the number of stimuli proceeding an error and the pace condition [$F(3,96) = 5.50$, $p < .01$]. Finally, a TUT Group × Pace interaction was observed [$F(1,32) = 6.2$, $p < .05$]. These effects were followed up by ANOVA on each probability condition.

In the LP condition, ANOVA indicated two effects: (i) an effect of period [$F(3,105) = 4.8$. $p < .05$] and (ii) a Period × Pace × TUT group interaction [$F(3,48) = 3.42$, $p < .05$, see Fig. 5]. Paired samples t tests were employed to follow up the effects of Period on RT. These indicated in the LP condition RT eight stimuli before an error was higher than six stimuli [$t(38) = 3.2$, $p < .01$], four stimuli [$t(38) = 3.2$, $p < .01$] and two stimuli [$t(38) = 2.2$, $p < .05$] proceeding an error. Separate post hoc uni-variate ANOVA were employed to follow up the TUT group and Pace interaction in the LP condition. Uni-variate ANOVA in the slow SART indicated reliable effects of TUT Group at 4 stimuli [$F(1,17) = 12.1$, $p < .01$] and 2 stimuli [$F(1,17) = 4.4$, $p < .05$], indicating that the High TUT group showed faster RT in these periods before an error. No effects of TUT Group were reliable in the fast condition ($p =$ ns). In the HP condition a Pace × TUT group interaction was observed [$F(1,32) = 6.4$, $p < .05$]. Separate uni-variate ANOVA in the Slow SART indicated reliable TUT group differences in RT for four stimuli [$F(1,17) = 7.0$, $p < .05$] only. As for the LP condition, no reliable effects of TUT Group were observed for the fast condition ($p =$ ns).

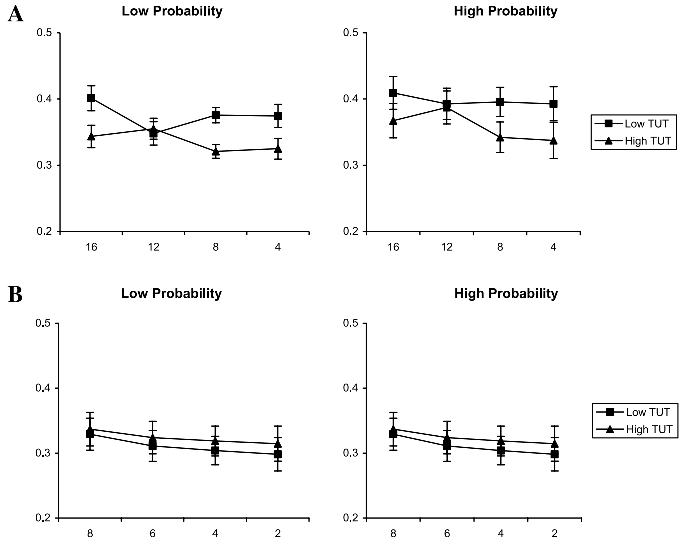

Fig. 5. Experiment 3. Effects of TUT group and time preceding an error on the reaction time (s) in the LP and HP conditions. (A) The data from the slow condition and (B) the fast condition.

Effects of TRI. The analysis of RT over the eight stimuli indicated an effect of Period that was marginally significant [$F(1,96) = 3.89$, $p > .05 < .07$, see above].

4.3.5. Errors of commission

Effects of TUT. A Mixed $3 \times 2 \times 2 \times 2$ ANOVA was used to contrast the distribution of errors of commission. This ANOVA included repeated measures on practice (1st, 2nd or 3rd Blocks) and Probability (HP and LP). Stimulus presentation pace [fast and slow] and TUT Group (High and Low) were included as between participants' factors. ANOVA indicated two reliable effects: (i) a Probability $\times$ Practice $\times$ TUT Group interaction [$F(2,64) = 3.20$, $p < .05$] and (ii) a Probability $\times$ Practice $\times$ TUT Group $\times$ Pace [$F(2,64) = 3.10$, $p < .05$]. These interactions were followed up by separate ANOVA on each probability condition. In the LP condition a Practice $\times$ TUT Group $\times$ Pace interaction was observed [$F(2,64) = 3.12$, $p < .05$]. This interaction was followed up by conducting separate ANOVA on each stimulus presentation condition (see Table 4). In the slow pacing condition, no reliable effects of either block or TUT group, nor the subsequent interaction were observed ($p =$ ns). The analysis of the fast stimulus presentation condition indicated a Block $\times$ TUT group interaction [$F(2,30) = 3.7$, $p < .05$]. Separate post hoc LSD comparisons indicated group difference in block 2 only ($p < .05$, see Table 4) indicating that errors decreased in the High TUT group more readily than the Low TUT group.

By contrast, ANOVA on the HP condition revealed an effect of Pace [$F(1,32) = 9.76$, $p < .01$] although this was clarified by a subsequent TUT Group $\times$ Pace interaction [$F(1,32) = 4.08$, $p < .05$, see Fig. 6A]. Post hoc LSD tests indicated that the Low TUT Group in the fast condition made more errors than either the High [$p < .01$] or the Low TUT Group [$p < .001$] in the slow condition. No other effects or interactions were observed. As in Experiments 1 and 2, group differences in the experience of TUT, especially during fast stimulus presentation generally reflects a low frequency of errors of commission.

Table 4
Experiment 3. The effects of stimulus presentation pace (fast and slow), block and TUT group on the mean likelihood of an error of commission in the low probability condition

Stimulus presentation		Mean errors of commission			
		Slow		Fast	
Block		Low TUT	High TUT	Low TUT	High TUT
1	Mean	0.19	0.20	0.40	0.45
	SD	0.18	0.15	0.27	0.28
2	Mean	0.22	0.25	0.45*	0.20*
	SD	0.16	0.19	0.40	0.21
3	Mean	0.13	0.09	0.35	0.23
	SD	0.35	0.17	0.32	0.13

* Significant group differences ($p < .05$, post hoc LSD test).

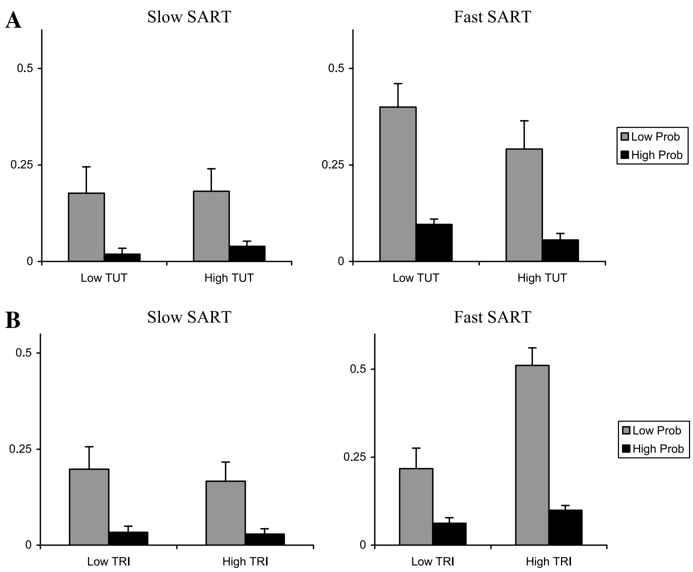

Fig. 6. Experiment 3. The contrasting effects of two types of subjective experience: (A) TUT and (B) TRI on errors of commission. Separate panels display errors of commission for the Low Probability conditions during the two versions of the SART task (Slow and Fast). The Y-axis describes the likelihood of an error of commission.

Effects of TRI. ANOVA yielded a TRI Group × Pace × Probability condition interaction [$F(1,32) = 8.70$, $p < .01$, see Fig. 6B]. Separate ANOVA on the HP and LP conditions respectively indicated a TRI Group × Pace interaction for the LP condition [$F(1,32) = 8.6$, $p < .01$]. Separate ANOVA on each stimulus presentation condition indicated an effect of TRI Group in the fast condition [$F(1,15) = 8.1$, $p < .01$] suggesting that in the fast SART the high TRI Group made significantly greater number of errors of commission in the fast SART task [Mean = .51, $SD = .07$] than the Low TRI Group [Mean = .21, $SD = .07$]. No difference was observed in the LP condition in the slow SART. By contrast, in the HP condition there was no effect of TRI group on errors of commission, nor any reliable interactions ($p =$ ns).

4.3.6. Discussion of Experiment 3

In Experiment 3, we sought to explore the relationship between subjective experience and attentional lapses in the absence of thought monitoring. The analysis of both a constant number of stimuli and over a constant period of time before an error indicated that TUT was reliably associated with time-limited accelerations in RT preceding an error on the slow SART. This result is consistent with those from Experiments 1 and 2 and the effects of TUT are observable without the individual being asked to monitor their ongoing conscious experience during a shorter task. Second, Experiment 3 demonstrates that we can employ a manipulation that has been shown elsewhere to moderate TUT frequency (presentation pace, Giambra, 1995) and thereby attenuate the effects of TUT on RT in the SART task, providing theoretical evidence to suggest that the two phenomenon show reasonable overlap. Third, the positive association between TRI and errors of commission was consistent with the results of Experiments 1 and 2. Finally, by contrast with previous experiments, TRI was associated with reliable effects on RT, notably during the fast stimulus condition, although these changes were reflected by chronic changes rather than time limited shifts, as was the case for TUT.

4.3.7. Combined analysis

The final analysis we report in this paper is concerned with investigating the relationship between the various self-report measures that we have collected over the course of the three experiments. Using stepwise multiple regression we explored the dispositional, emotional and contextual predictors of the two components of subjective experience measured in these studies (see Tabachnick & Fidel, 2000). To achieve this aim we combined the data from all three experiments into one super-ordinate data set ($n = 75$). To control for the differences in length of testing session and stimulus presentation rate in this series of experiments we created two dummy variables which reflect the stimulus presentation speed (fast and slow) and testing duration (short and long). Using this framework, data from Experiment 1, for example, was classified as slow and long. By contrast, data from the fast condition of Experiment 3 was classified as fast and short.

The participants' age, the various dispositional measures (Gender, CESD, and RSQ), current mood (DSSQ: energetic arousal, tense arousal, and hedonic tone) in conjunction with the two contextual variables (speed and testing duration) were entered into the regression analysis as predictors of the two components of subjective experience (TUT and TRI). Finally, the reciprocal dimension of subjective experience was included (i.e., TRI was included in the regression for TUT). Stepwise regression with TUT as the dependent variable indicated that four variables accounted for 58% of the variance [Age, rumination, CESD and duration, $R = 0.76$, $R^2 = .58$, $F(4, 64) = 20.8$, $p < .0001$, see Table 5]. By contrast, the step-wise regression conducted with TRI as a dependent variable indicated that only one variable, TUT, predicted TRI. TUT accounted for 14% of the variance of TRI [TUT, $R = .38$, $R^2 = .14$, $F(1, 64) = 10.6$, $p < .01$, see Table 5].

4.4. Discussion

The analysis of the distribution of retrospective reports of subjective experience indicates two issues worthy of comment. First, the effects of contextual factors on the distribution of task-unrelated thinking are comparable with previous literature: (i) TUT is higher in longer testing sessions (Cunningham et al., 2000; Smallwood et al., 2003c; Teasdale et al., 1995), (ii) TUT is

Table 5
Combined analysis. Results of the two separate step-wise multiple regression on the measures of retrospective subjective experience recorded throughout these four experiments (TUT and TRI)

Dependent	Variable	Unstandardised co-efficient		Standard co-efficient		
		B	*SE*	*β*	*t*	*p*
TUT	(Constant)	6.00	1.70		3.54	.00
	Duration	7.49	0.93	0.66	8.08	.001
	CESD	0.26	0.06	0.42	4.65	.001
	Age	−0.18	0.07	−0.22	−2.72	.008
	RSQ	−0.17	0.07	−0.22	−2.40	.020
TRI	(Constant)	15.39	1.69		9.10	.00
	TUT	0.37	0.10	0.41	3.78	.00

higher in younger individuals (Giambra, 1989, 1993), (iii) TUT is higher in periods of dysphoria (Smallwood et al., 2004; see also Lyubomirsky et al., 2003), and (iv) is negatively associated with an individual's tendency to ruminate (Smallwood et al., 2004). This consistency demonstrates that we can generalise from the results of these studies to previous research using TUT and also that we can generalise the data generated by retrospective questionnaires employed in these experiments to data generated using thought probes/online self-report (i.e., Antrobus, 1968; Cunningham et al., 2000; Giambra, 1995; Smallwood et al., 2003c). Second, the stepwise regression indicates that the only predictor of task appraisal (i.e., TRI) was the frequency of TUT, confirming that in the context of the SART TRI may reflect strategic deployment of attention to the task in response to an attentional lapse (Manly et al., 1999).

5. General discussion

In this series of experiments we set out to investigate the relationship between subjective experience and attentional lapses. As a broad framework, we investigated the claims of Robertson and colleagues (Manly et al., 1999, 2002; Robertson et al., 1997) regarding the association between components of action slips during the SART (RT and errors of commission) and the direction of subjective experience (either task disengagement or task engagement). In particular, Robertson et al. (1997) suggested that under conditions of low target probability attention tends to drift away from task relevant material, and that in the context of the SART this is expressed by an acceleration of RT to the non-target stimulus. Second, they suggest that detection of an error of commission tends to re-direct "attentional resources" towards the task slowing responses to a level in which the alternative response can supervene (Manly et al., 1999).

The investigation of subjective experience associated with performance on the SART provides support for both of these positions. First, the results of all three experiments demonstrate that under conditions of slow-moderate stimulus presentation pace, task disengagement, expressed as TUT, was associated with time-limited accelerations in reaction time. The results of Experiment 3 imply that this phenomenon was strongest under conditions of low target probability. Moreover, the correlations presented in Table 3 suggest that only under situations in which verbal reports indicate task disengagement can we see the reliable negative association between RT and errors, implied by Robertson and colleagues (Manly et al., 1999, 2002; Robertson et al., 1997). Under the same stimulus conditions, verbal reports indicative of task focus (NTUT) were associated with higher frequencies of errors of commission (Experiments 1 and 2) in individuals who reported low levels of TUT. Moreover, those individuals who reported high frequencies of subjective experience associated with appraisal of the self/task, i.e., TRI, consistently made higher frequencies of errors of commission. When TRI was associated with RT (Experiment 3), it was by and large associated with chronic changes in RT, rather than time limited changes, as was the case for TUT. The degree of discrimination between these two components of subjective experience is surprising given the statistical overlap between these two constructs (Table 5).

5.1. Potential limitations

Before dealing with the implications of the results of this series of experiments, it is worth considering the limitations of the data presented. The most compelling criticism of the studies presented is that they all, to a greater or lesser extent, rely on the self-report of the participants in the study. Whilst the results of Experiment 3, for example, suggest that the relationships between subjective experience and task performance described in Experiments 1 and 2 in particular, are observable without the employment of thought sampling, we cannot rule out the possibility that some tertiary variable, such as response bias is responsible for the relationship in question. Such a criticism is likely to be most applicable to the physiological data generated. For example, it is possible that the association between high HR and TUT is a consequence of the act of monitoring one's thinking. Alternatively it may indicate arousal as the participant reflects on the fact that they will be subsequently required to report their thinking. Whilst, at present we cannot rule out these possibilities there are several reasons why this may not be the case.

First, an explanation at the level of demand characteristics alone would be unlikely to account for the dissociation between physiology, behaviour and verbal report. Whilst the relationship between subjective experience and physiological measures was not directly replicated over the two experiments, there is certainly enough consistency to draw some general conclusions. Consistently, throughout these studies task disengagement was positively associated with HR (Experiments 1 and 2) and negatively associated with RT. By contrast, task engagement was broadly associated with GSR (Experiments 1 and 2) and errors of commission.

Second, it is important to note that the specific co-variation between physiological measures recorded during blocks of task engagement/dis-engagement in our methodology requires agreement, not just between the verbal reports of the individuals themselves, it also rests upon agreement with the independent, experimental raters employed in each study, who were unaware of the physiological/behavioural data when classifying the thoughts. Nonetheless, until the physiological variables are investigated in the absence of thought monitoring the issue of the relationship between physiological arousal and subjective experience should be treated with a degree of caution. A recent study, however, provides some confidence in this interpretation of our data. O'Keefe, Dockree, and Robertson (2004) demonstrated that during the SART errors of commission were associated with subsequent increases in GSR in controls but not participants with TBI, who were less aware that they had made an error. In the future, it is important that this issue of demand characteristics be resolved by investigating changes in psychophysiology on the SART using a methodology similar to Experiment 3 that does not rely on on-line thought monitoring.

A second, potential criticism of these experiments concerns relationship between TUT and target detection. In particular, the experience of TUT appears to be related to superior performance in some of the experimental situations presented in this paper (particularly the fast condition, Experiment 3). On this basis it is unclear whether the verbal reports reflect task disengagement, or alternatively, the high frequency of TUT is a consequence of a higher quantity of available cognitive resources in those individuals who performed the task in a relatively error free manner. Whilst we cannot rule out this interpretation, there are several reasons why this particular account does not explain all the experimental results presented in this paper in satisfactory manner. For example, it is unclear why a higher availability of cognitive resources, in isolation, would relate in a reliable fashion to time limited changes in RT, particularly when accelerations in RT have been shown to impair performance on this task (Robertson et al., 1997) or why an excess of resources would encourage the context dependent variations of these errors (Experiments 1 and 2). Nor does it follow that we should expect a consistent pattern between the results of Experiments 1 and 2, in which data is categorised using thought sampling and Experiment 3 in which the analysis is based on the distribution of errors. As the analysis in the final experiment demonstrates, TUT is not associated with chronic changes in RT, as one would expect on the basis of the resource account, rather the High TUT group show accelerations in RT only before an error is made. Finally, it is important to note that previous research has validated TUT by comparison with task performance (for a discussion see Smallwood et al., 2003b). For example, previous literature demonstrates that in the context of random number generation (Teasdale et al., 1995, Experiment 4), the encoding and generation of verbal material (Smallwood et al., 2002, 2003a, 2003b) and text comprehension (Schooler et al., in press) TUT is associated with inferior task performance. Taken together these points suggest that the experience of TUT is associated with measurable, albeit subtle, impairments in the ability of the individual to sustain attention over a prolonged period of time.

Perhaps the most powerful refutation of this account of the TUT data, however, follows from the theoretical interpretation of the nature of TUT, proposed by Singer (1975) and exemplified in the following quotation: "By storing and manipulating internal information we organise what could not be organised during stimulus presentation, solve problems that require computation over long periods of time, and create effective plans governing behaviour in the future. These capabilities have surely made no small contribution to human survival and the invention of technology" (Binder et al., 1999, p. 85). It is clear that internal processes such as TUT, provide a potential vehicle for the anticipation of complex situations and are, therefore, likely to play an important role in facilitating problem solving in the less hazardous environment of the neural workspace (Cleeremans & Jiménez, 2002; Dehaene & Naccache, 2001). As this explanation has clear evolutionary significance, it is important to emphasise that the adaptive value of this system would be limited if the organism engages in intractable cognitive activity to the exclusion of all alternative sources of external information. Clearly, far greater adaptive flexibility would be provided by a system in which during periods of low environmental stimulation subjective awareness can co-ordinate internal information in a stimulus independent fashion, and yet is readily interrupted by salient signals from the external environment. Such an account is clearly consistent with data presented in this paper which describes a cognitive process which, whilst capable of manipulating internalised information, is readily disrupted by important environmental events as in the case of an error. The evidence in this paper suggests that it is only those individuals who report high levels of absentmindedness (High TUT Group, Experiments 1 and 2) or individuals with TBI (Robertson et al., 1997) whose internal cognitive experience may not be disrupted by an error.

Finally, it is important to note that we are not suggesting that the experience of TUT is a causal phenomena in determining errors on the SART task. Whilst our data, suggest that the relationship between subjective experience and task performance co-varies with attentional lapses, it is possible that TUT is merely an epiphenomenona associated with failures in sustained attention. The regression analysis indicates that constructs such as dysphoria play an important role in mediating the likelihood that TUT is experienced and would be expected to contribute to failures of sustained attention via the mechanisms of increasing task disengagement. The relationship between TBI and SART performance supports this interpretation as these individuals show organic deficits (Robertson et al., 1997). A second important aim for research in the future, therefore, is to document the factors that moderate the relationships described in this paper. One advantage of the approach described in this paper, however, is that it encourages us to conceptualise attention lapses as a complex dynamic interaction between internal (such as dysphoria) and external influences (such as target probability and/or presentation pace).

5.2. Phenomenology of subjective experience

The evidence presented in this paper confirms that the experience of TUT is best conceptualised as absentmindedness: a time limited phenomenon in which the individual's attention becomes decoupled from the current task. This state of mind is readily disrupted by external events because: (i) the effects on RT are limited to situations of moderate stimulus presentation (Experiment 3) and (ii) whilst associated with errors it does not predict crude error rates in any of the three studies. In addition to information processing factors, however, investigations of day-to-day thinking, suggest that the experience of task unrelated thinking is often directed towards one's current concerns (e.g., Klinger, 1999). Several aspects of the results presented in this paper support this perspective. First, the higher HR associated with TUT (Experiments 1 and 2) suggests that some degree of physiological activation is associated with the experience of TUT. This interpretation is consistent with previous research which demonstrated that higher body temperatures were associated with periods of high TUT frequency (Giambra, Rosenberg, Kasper, Yee, & Sack, 1988). Second, the fact that high levels of dysphoria accompany verbal reports of TUT during these experiments and elsewhere (Lyubomirsky et al., 2003; Seibert & Ellis, 1991; Smallwood et al., 2004) implies that these thoughts may well involve the processing of information of personal salience. Taken together, these data provide an important, additional line of evidence that the processing of TUT involves the processing of internalised information with personal salience (Klinger, 1999).

Unlike TUT, the results of this series of experiments suggest that the experience of task focus and the sub-component, TRI, may reflect a strategic response to environmental appraisal. Generally, task focus was reported in blocks in which errors and the physiological indices associated with effort (GSR) were higher. In these experiments, High TRI was consistently associated with: (i) a high frequency of errors on the SART task and (ii) chronic accelerations of reaction time under circumstances of higher workload (fast condition, Experiment 3). A parsimonious interpretation of these disparate findings might be that TRI reflects a form of strategic self-regulation which is expressed by certain individuals when environmental circumstances either exceed, as in the case of failure to withhold a go response when faced with a target, or are perceived to exceed the individual's capabilities, as in the fast condition of Experiment 3. Previous work investigating the appraisal of the self/task suggests that both self-consciousness (Matthews, Mohammed, & Lochrie, 1998) and rumination (Smallwood et al., 2004) may be important characteristics of the individuals who respond to workload in this manner. Consequently it would be useful for future research to identify whether these personality dimensions play a similar role in determining the relationship between subjective experience and failures on the SART.

Finally, it is worth speculating on the time course of subjective experience as it occurs throughout this task. Recent conceptions of TUT in terms of 'zoning out' imply that in certain circumstances, the individual lacks momentary awareness that their attention has become dissociated from the current situation (Schooler, 2002; see also Schooler et al., in press). Subject to individual differences, such drifts of attention occur across a wide sample of individuals with reasonable frequency, particularly under situations of low environmental support. In the context of the SART task these drifts of attention will be periodically interrupted by an error, providing a source of feedback which may indicate the individual's current state of meta awareness (Schooler, 2003). Unlike traditional sustained attention tasks, the 'Oops' phenomenon associated with errors on the SART reflects a situation in which the individual's failure to detect a target is available to their awareness. By contrast in "many putative failures in signal detection will tend to be, by definition, unnoticed" (Manly et al., 2002, p. 669). This form of feedback is only available in a task such as the SART in which both correct and incorrect responses are placed in opposition with one another. The positive association between TRI and error rates may, therefore, be indicative of attempts by certain individuals to prevent the expression of these drifts of attention when the task indicates their attention has lapsed. This interpretation is certainly consistent with the experimental data presented in this paper, particularly: (i) the time limited nature of variations in RT which accompany TUT, (ii) the positive association between TRI and errors and (iii) the results of the step wise regression which indicate that the only reliable predictor of TRI was TUT. If this time course of subjective experience is supported by subsequent experimental data, it may shed important light on the nature of self-focus in clinical conditions. In particular, it offers a possible explanation of the paradoxical finding that rumination is associated with lower frequencies of TUT (Smallwood et al., 2004), high degrees of cognitive interference (e.g., TRI, Lyubomirsky et al., 2003; Smallwood et al., 2004) and longer depressive episodes (Nolen-Hoeskema, 1991). It is plausible that rumination may extend the duration of a depressive episode because the individual employs a self-regulative strategy which attempts to control the haphazard wandering of attention which seems to accompany emotional states in general (Seibert & Ellis, 1991) and is particularly frequent in dysphoria (Smallwood et al., 2004). This perspective is supported by the efficacy of mindfulness based cognitive therapy in reducing the length of depressive episodes by training individuals to identify when their attention has drifted from the current situation and accept this with out attempting to control it (Williams, Teasdale, Segal, & Soulsby, 2000).

6. Conclusion

Overall, these experiments demonstrate that it is possible to combine the techniques of thought monitoring with sensitive cognitive and psycho physiological measures to investigate the phenomenological status of lapses of attention. In particular, these experiments highlight the advantages of developing reliable indices for moment to moment changes in the focus of attention. It is worth noting, however, that whilst it may be tempting to suggest that this paper is primarily concerned with validating the technique of thought monitoring, Jack and Roepstorff (2003) argue that introspective evidence can actually enrich validity of the objective claims of cognitive science. They remind us "So long as cognitive science continues to doubt the face validity of introspective reports it will never conduct the investigations necessary to provide full validation of those measures" (Jack & Roepstorff, 2003, p. xiii). In this sense the results of these experiments not only validate the thought sampling methodology employed in this paper, they also provide a source of validation for the claims of a variety of theorists who have investigated the attentional lapse over the last three decades (Broadbent et al., 1982; Manly et al., 1999, 2002; Reason & Lucas, 1984; Robertson et al., 1997).

Acknowledgments

The authors thank Professor John Antrobus and two anonymous reviewers for their helpful comments on an earlier version of the manuscript. Thanks also to Gordon McAlpine for writing the computer programme employed in Experiment 3. Finally, we thank the participants who took part in these experiments.

References

Antrobus, J. (1968). Information theory and stimulus independent thought. *British Journal of Psychology, 59*, 423–430.

Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness: Experimental suppression of spontaneous cognitive processes. *Perceptual and Motor Skills, 23*, 399–417.

Broadbent, D. E., Broadbent, M. H., et al. (1986). Performance correlates of self-reported cognitive failure and of obsessionality. *The British Journal of Clinical Psychology, 25*(Pt. 4), 285–299.

Broadbent, D. E., Cooper, P. F., et al. (1982). The Cognitive Failures Questionnaire (CFQ) and its correlates. *The British Journal of Clinical Psychology, 21*(Pt. 1), 1–16.

Binder, J. R., Frost, J. A., Hammeke, P. S., Bellgowan, S. F., Rao, S. M., & Cox, R. W. (1999). Conceptual processing during the conscious resting state: A Functional MRI Study. *Journal of Cognitive Neurosciences, 11*(1), 80–93.

Cleeremans, A., & Jiménez, L. (2002). Implicit learning and consciousness: A graded, dynamic perspective. In R. M. French & A. Cleeremans (Eds.), *Implicit learning and consciousness* (pp. 1–40). Hove, UK: Psychology Press.

Cunningham, S., Scerbo, W., & Freeman, F. (2000). The electrocortical correlates of daydreaming during vigilance tasks. *The Journal of Mental Imagery, 24*, 1–2.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition, 79*, 1–2, see also pp. 1–31.

Desmond, P. A., Matthews, G., & Bush, J. (2001). Sustained vigilance during simultaneous and successive vigilance tasks. In *Proceedings of the human factors and ergonomics society, 45th annual meeting*, pp. 1386–1389.

Fowles, D. C., Christie, M. J., Edelberg, R., Grings, W. W., Lykken, D. T., & Venables, P. H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology, 18*, 232–239.

Giambra, L. M. (1989). Task-unrelated-thought frequency as a function of age: a laboratory study. *Psychol. Aging, 4*(2), 136–143.

Giambra, L. M. (1993). The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Exp. Gerontol., 28*(4-5), 485–492.

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition, 4*, 1–21.

Giambra, L., & Grodsky, A. (1989). Task unrelated images and thoughts whilst reading. In J. Shorr, P. Robin, J. A. Connek, & M. Wolpin (Eds.), *Imagery: Current perspectives*. NY: Plenum Press.

Giambra, L. M., Rosenberg, E. H., Kasper, S., Yee, W., & Sack, D. A. (1988). A circadian rhythm in the frequency of spontaneous task-unrelated images and thoughts. *Imagination, Cognition, and Personality, 8*, 307–312.

Golding, J. F. (1992). Phasic skin conductance activity and motion sickness. *Aviation, Space, and Environmental Medicine, 63*, 165–171.

Grodsky, A., & Giambra, L. (1991). Task unrelated images and thoughts whilst reading. In J. Shorr, P. Robin, J. A. Connek, & M. Wolpin (Eds.), *Imagery: Current perspectives*. New York: Plenum Press.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319–340.

Jack, A., & Roepstorff, A. (2003). Why trust the subject? *Journal of Consciousness Studies, 10*, 9–10, v-xx (1).

Klinger, E. C. (1978). Modes of normal conscious flow. In K. S. Pope & J. L. Singer (Eds.), *The stream of consciousness: Scientific investigations into the flow of human experience* (pp. 225–258). New York: Plenum.

Klinger, E. C. (1999). Thought flow: Properties and mechanisms underlying shifts in content. In J. A. Singer & P. Salovey (Eds.), *At play in the fields of consciousness: Essays in the honour of Jerome L Singer*. Mahwah, NJ: Erlbaum.

Klinger, E., Barta, S. G., & Maxeiner, M. E. (1980). Motivational correlates of thought, content, frequency and commitment. *Journal of Personality and Social Psychology, 39*, 1222–1237.

Larson, G. E., Anderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and the Cognitive Failures Questionnaire. *British Journal of Psychology, 8*, 29–38.

Lyubomirsky, S., Krasri, F., & Zehm, K. (2003). Dysphoric rumination impairs concentration on academic tasks. *Cognitive Therapy and Research, 27*(3), 309–330.

Manly, T., Lewis, G. H., Robertson, I. H., Watson, P. C., & Datta, A. K. (2002). Coffee in the cornflakes: Time-of-day, routine response control and subjective sleepiness. *Neuropsychologia, 40*(1), 1–6.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37*, 661–670.

Matthews, G., Joyner, L., Gililand, K., Campbell, S. E., & Faulconner, S. (1999). Validation of a comprehensive stress state questionnaire: Towards a state Big three?. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Handbook of coping: Theory, research and applications* (pp. 333–350). Tilburg Netherlands: Tilburg University Press.

Matthews, G., Mohammed, A., & Lochrie, B. (1998). Dispositional Self-focus of attention and individual differences in appraisal and coping. In I. J. Bermudez, A. M. Perez, & G. L. van Hecks (Eds.), *Personality psychology in Europe* (Vol. 6, pp. 278–285). Tilburg Netherlands: Tilburg University Press.

Matthews, G., Schwean, V. L., Campbell, S. E., Saklofske, D. H., & Mohammed, A. A. R. (2000). Personality, self-regulation and application: A cognitive-social framework. In *Handbook of Self-Regulation* (pp. 171–207). Academic Press.

Nisbet, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on Mental Processes. *Psychological Review, 84*, 3.

Nolen-Hoeskema, S. (1991). Responses to depression and their and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology, 100*, 569–582.

O'Keefe, F., Dockree, P., & Robertson, I. H. (2004). Awareness deficits in traumatic brain injury mediated by impaired error processing? Evidence from electrodermal activity. *Journal of Cognitive Neuroscience*, supplement, 31.

Parasuraman, R. (1979). Memory load and event rate control sensitivity decrements in sustained attention. *Science, 205*, 924–927.

Pechineda, A., & Smith, C. A. (1996). The Affective Significance of Skin Conductance Activity during a difficult problem solving task. *Cognition and Emotion, 10*(5), 481–503.

Radloff, L. S. (1977). The CES-D Depression Inventory. *Applied Psychological Measurement, 1*(3), 385–401.

Reason, J., & Lucas, D. (1984). Absent-mindedness in shops: its incidence, correlates and consequences. *The British Journal of Clinical Psychology, 23*(Pt. 2), 121–131.

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neurospsychologia, 35*(6), 747–758.

Sarason, I. G., Sarason, B. R., Keefe, D. E., Hates, B. E., & Shearin, E. N. (1986). Cognitive interference: situational determinants and trait-like characteristics. *Journal of Personality and Social Psychology, 51*(1), 215–226.

Schooler, J. W. (2002). Representing consciousness: Dissociations between consciousness and meta-consciousness. *Trends in Cognitive Science, 6*, 339–344.

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (in press). Zoning-out during reading: Evidence for dissociations between experience and meta-consciousness. In D. Levin (Eds.), *Visual meta-cognition: Thinking about seeing*. Praeger.

Schneider, W., & Shifrin, R. M. (1977). Controlled and automatic processing 1: Detection, search and attention. *Psychological Review, 84*, 1–66.

Seibert, P. S., & Ellis, H. C. (1991). Irrelevant thoughts, emotional mood states and cognitive performance. *Memory & Cognition, 5*, 507–513.

Singer, J. L. (1975). *The inner world of daydreaming*. New York: Harper & Row.

Smallwood, J. (2004). Ambiguity, self-reference and dysphoria. *Cognition and Emotion*. Forthcoming.

Smallwood, J., Baraciaia, S. F., Lowe, M., & Obonsawin, M. C. (2003b). Task unrelated thought whilst encoding information. *Consciousness and Cognition, 12*, 452–484.

Smallwood, J., Obonsawin, M. C., Baracaia, S. F., Reid, H., O'Connor, R. C., & Heim, S. D. (2004). The relationship between rumination, dysphoria and self-referent thinking: Some preliminary findings. *Imagination, Cognition and Personality*. Forthcoming.

Smallwood, J., Obonsawin, M. C., & Heim, S. D. (2003a). Task Unrelated Thought: the role of distributed processing. *Consciousness and Cognition*, 169–189.

Smallwood, J., Obonsawin, M. C., & Reid, H. (2003c). The effects of block duration and task demands on the experience of task unrelated thought. *Imagination, Cognition and Personality, 22*(1), 13–31.

Smallwood, J., Obonsawin, M. C., Reid, H., & Heim, S. D. (2002). An investigation into the role of personality and situation in the maintenance of subjective experience in a laboratory. *Imagination, Cognition and Personality, 21*(4), 319–332.

Tabachnick, B., & Fidel, L. (2000). *Using Multivariate Statistics* (4th ed.). New York: Allyn & Bacon.

Teasdale, J. D., Dritschell, B. H., Taylor, M. J., Proctor, L., Lloyd, C. A., Nimmo-Smith, I., & Baddeley, A. D. (1995). Stimulus Independent Thought depends upon central executive resources. *Memory & Cognition, 28*(5), 551–559.

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 183–200). Hillsdale, NJ: Erlbaum.

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology, 51*, 59–91.

Williams, J. M. G., Teasdale, J. D., Segal, Z. V., & Soulsby, J. (2000). Mindfulness-based cognitive therapy reduces overgeneral autobiographical memory in formerly depressed patients. *Journal of Abnormal Psychology, 109*(1), 150–155.