Background Music Stints Creativity: Evidence from Compound Remote Associate Tasks

Journal:	*Applied Cognitive Psychology*
Manuscript ID	ACP-17-0082.R1
Wiley - Manuscript type:	Research Article
Keywords:	Music, Creativity, Insight, Compound Remote Associates Task, Distraction

Background Music Stints Creativity:

Evidence from Compound Remote Associate Tasks

Abstract

Music has a profound effect on the brain, activating emotional and motor areas. It is therefore unsurprising that background music has been claimed to enhance creativity. In three experiments we investigated the impact of background music on performance of Compound Remote Associate Tasks (CRATs). Background music with foreign (unfamiliar) lyrics (Experiment 1), instrumental music without lyrics (Experiment 2), and music with familiar lyrics (Experiment 3) all significantly impaired CRAT performance in comparison to quiet background conditions. Furthermore, Experiment 3 demonstrated that background music impaired CRAT performance regardless of whether the music induced a positive mood, was liked by the participants, or whether participants typically studied in the presence of music. The findings challenge the view that background sound enhances creativity, and are discussed in terms of an auditory distraction account (interference-by-process; Jones & Tremblay, 2000) and the processing disfluency account (Mehta et al., 2012).

Keywords: Music, Creativity, Insight, Compound Remote Associates Tasks, Distraction

Background Music Stints Creativity:
Evidence from Compound Remote Associate Tasks

Creativity is a vital aspect of cognition underpinning activities such as innovative product design, scientific advancement and effective advertising and marketing communications. Background music is an environmental stimulus known to influence cognitive performance, which is commonly believed to enhance people's creativity (Schellenberg, Nakata, Hunter, & Tamoto, 2007). Yet, we argue that there is inadequate empirical support for this belief, with to our knowledge only one study demonstrating that the presence of music facilitates performance on a divergent thinking creativity task (i.e., "list alternative uses for a brick"; Ritter & Ferguson, 2017). Another reason to be cautious regarding the notion that background music can enhance performance on tasks tapping creative cognition is the presence of a substantial research base demonstrating that to-be-ignored background sound impairs task performance (Beaman, 2005; Hughes & Jones, 2003).

In the present paper we critically examine the claim that background music enhances creativity by employing variants of widely used verbal problem solving tasks that are typically used to study creativity (Ansburg, 2000; Fodor, 1999; Mednick & Mednick, 1967; Mehta, Zhu, & Cheema, 2012; Mikulincer & Sheffi, 2000; Storm, Angello, & Bjork, 2011) being indexed by, and solved via, a process of insight: Compound Remote Associate Tasks (CRATs; e.g., see Bowden, Jung-Beeman, Fleck, & Kuonios, 2005). We contrast two competing accounts of the impact of background music on creative problem solving: (1) the processing disfluency account (Mehta et al., 2012), in which background music potentially enhances creativity by engendering processing disfluency and thence increased task engagement; and (2) the auditory distraction (interference-by-process) account (e.g., Jones & Tremblay, 2000; Marsh, Hughes, & Jones, 2009; Perham & Vizard, 2010), which assumes that the presence of any type of auditory distractor sequence will disrupt cognitive task performance providing it demonstrates changing-state characteristics. That is, auditory sequences in which a series of elements differ from one element to the next (such as tones, syllables, words) in terms of frequency/pitch/timbre are more disruptive than a series within which the same element is repeated such as the same tone, syllable or word. It has been shown, for example, that the latter, steady-state stimuli typically fail to disrupt short-term memory performance (e.g., Jones & Macken, 1993).

Creative problem solving is characterised by the ability to perceive a problem space in new ways by discovering hidden patterns or by connecting seemingly unrelated ideas (e.g., Ohlsson, 2011). One key way in which creative problem solving comes about is by means of so-called *insight*, with tasks that tap creativity typically being solved via insight processes. Accounts of insight in problem solving such as the “special-process theory” (e.g., Bowden et al., 2005; Ball & Stevens, 2009) argue that problems that tend to be solved via an insight process call upon very different processing mechanisms to “non-insight” problems. For example, Jung-Beeman et al. (2004) identified neural patterns just prior to the emergence of insight that demonstrate a hemispheric shift in processing occurring at this point. Jung-Beeman et al. (2004) propose that during insight problem solving loose associative processing occurring non-consciously in the right temporal lobe takes precedence over finer-grained processing in the left hemisphere, implying that neural areas linked with diffuse associative processing are critical for the emergence of creative insight (for a recent review of related findings see Shen, Yuan, Liu, & Luo, 2017).

Several researchers suppose that an insight sequence defines creative thinking and that any advance in thought that is not characterised by such a sequence is therefore not creative (e.g., Ohlsson, 2011; Perkins, 2000; Wiley & Jarosz, 2012; but see Weisberg, 2015). This unique sequence of events that defines insight in problem solving comprises: presentation of the problem, repeated failure, impasse, restructuring, and an “Aha!” experience that is associated with solution generation. According to this sequence of events, failed attempts to solve a problem can lead to an *impasse*, whereby the participant, after several unsuccessful attempts at solving the problem, feels they are unable to move forward to reach a solution. After a period of failing to make progress, an abandoning of the original problem structure occurs and a new representation of the problem is formed through restructuring, which may itself be based on processes such as spreading activation in associative networks (see Shen et al., 2017). Such problem restructuring may then lead to the emergence of a solution. Crucially, problems that are typically solved by insight often cannot readily be solved via routine search processes. This is because the starting conditions, goals, and possible sequences of actions are ambiguous (i.e., a heuristic-type search within the *original* problem representation will not yield a solution).

As we have noted, our present research used Compound Remote Associates Tasks (CRATs) as a measure of insight-based creative problem solving (Bowden & Jung-Beeman, 1998). A CRAT involves a participant being shown three words (e.g., dress, dial, flower), with the requirement being to find a single associated word (in this case “sun”) that can be combined with each presented word (either being placed before it or after it) to make a common word or phrase (i.e., sundress, sundial and sunflower in the present example). CRATs are variants of problems referred to as Remote Associate Tasks (RATs; see Mednick, 1962; Mednick & Mednick, 1967), for which the solution can be associated with each of the provided three words in different ways. For example, a RAT (e.g., same, tennis, head), in contrast to a CRAT, can be solved by means of semantic association (tennis-match), synonymy (same = match) and, as with CRATs, the formation of compound words (matchhead).

Problem solving performance on CRATs has been found to correlate with performance on other creative tasks (Bowden & Beeman, 1998; Schooler & Melcher, 1995; Toplyn & Maguire, 1991; but see Webb, Little, Cropper, & Ruze, 2017). Such patterns of association suggest that CRATs represent effective tests of creativity. Moreover, CRATs also appear to involve “the same component processes critical for, and the same phenomenological experience of, insight solutions to more complex problems” (Bowden & Jung-Beeman, 2003, p. 634; see also Bowden & Jung-Beeman, 2007). For example, the problems initially misdirect or fail to direct retrieval processes, thereby leading to an impasse. In addition, solvers often report an “Aha!” experience on completion of a CRAT.

In addition to being characterised by the insight-sequence, CRATs also appear to be underpinned by a range of processes, including unconscious spreading activation in associative networks (Smith, Huber, & Vul, 2013), conscious verbal processes such as subvocal rehearsal (Ball & Stevens, 2009) and executive processes such as those that inhibit incorrect solution ideas and enable the active manipulation of information in working memory (Chein & Weisberg, 2013; Storm & Angello, 2010). Indeed, these problems have also been used to measure creativity in relation to sleep (e.g., Cai, Mednick, Harrison, Kanady, & Mednick, 2009), memory (e.g., Storm, Angello, & Bjork, 2011), attention (e.g., Ansburg & Hill, 2003) and attentional deficit hyperactivity disorder (ADHD; e.g., White & Shah, 2006) and they have additionally been employed in neuroimaging studies of creativity (e.g., Arden, Chavez, Grazioplene, & Jung, 2010).

Nowadays RATs and CRATs are a commonly used test of creativity within psychology and cognitive neuroscience. According to Bowden and Jung-Beeman (2003), their popularity rests in the fact that they have an unambiguous single word answer, and that multiple CRATs can be solved in a single session. Furthermore, CRATs are less complex than classic insight problems such as the candle or two-string problem (see Weisberg, 1995) and therefore less susceptible to confounding of variables. These characteristics made the CRAT appealing for the current investigation.

While there is a paucity of research examining the effects of background music on creativity, there is a small literature on the impact of noise on creative cognition (Hillier, Alexander, & Beversdorf, 2006; Kasof, 1997; Martindale & Greenough, 1973; Mehta et al., 2012). Aperiodic noise such as white noise and pink noise have been shown to affect creativity, as measured using RATs. For example, Martindale and Greenough (1973; 75 dB) and Hillier et al. (2006; 90 dB) showed that a high intensity white noise, compared to a no noise control condition, impaired task performance. Moreover, Kasof (1997) reported that a high level (85 dB[A]) of intermittent, compared to continuous, pink noise reduced creativity as measured with a poetry writing task. In contrast, Toplyn and Maguire (1991) found that highly creative participants (as gauged by their performance on RATs) demonstrated greater creativity on other tasks when exposed to 80 dB white noise, compared to when exposed to 60 dB or 100 dB white noise.

Mehta et al. (2012) used more naturalistic, ambient noises to resemble restaurant noise, wherein distant construction noise, multi-talker babble and roadside traffic were blended, and reported that a moderate level of noise (70 dB), as compared to low level noise (50 dB), improved performance on creative tasks. These tasks included RATs (Experiment 1), a task wherein participants generated novel ideas for improving mattress comfort (Experiment 2), a task requiring the generation of alternative uses for a brick (Experiment 3) and a task concerning how to clean scuffed shoes with no polish (Experiment 4). Of relevance to the present study, participants generated more correct answers to RATs in the presence of moderate noise compared to a low level of noise and a high level of noise (85 dB).

We note here, however, that in contrast with the RATs, the other tasks used by Mehta et al. (2012) arguably make less demands on verbal working memory. Indeed, these tasks tap divergent thinking in that they require the production of multiple responses in a manner similar to standard verbal fluency tasks. Verbal fluency tasks require the production of numerous responses given a phonemic (produce words beginning with the letter “F”) or semantic (produce as many examples of “Fruit”) cue within a time limit (Jones, Marsh, & Hughes, 2012; Marsh, Crawford, Pilgrim, Sörqvist, & Hughes, 2017). While some aspects of the task, such as the requirement to maintain memory for previously produced responses to avoid repetition tap verbal working memory, these tasks are not characterised by continuous generation and testing of word combinations and maintenance of intermediate solutions that distinguish the convergent thinking underpinning the RAT. Indeed, perhaps it is no surprise that tasks that tap divergent thinking such as category fluency, tend to be immune to disruption produced by changing state background sound, unless it conveys semantic content (Jones et al., 2012). In this respect, our focus was on the variant of the RAT (i.e., the CRAT), since in contrast to divergent thinking tasks, CRATs should be more sensitive to disruption produced by the changing-state acoustic properties of background sound.

An alternative account of the relationship between background sound and creativity holds that benefits to cognitive task performance can be observed through mood and arousal (for a review, see Schellenberg, 2005). For example, Thompson, Schellenberg, and Hussain (2001) showed that performance on tests of spatial abilities was improved when the tasks were executed after listening to music, as opposed to being exposed to quiet. Moreover, the improvement in performance was driven by changes in arousal and mood produced by listening to the music. A recent study by Ritter and Ferguson (2017) required participants to undertake tasks tapping creative cognition while concurrently listening to music or exposure to quiet. In a between-participants design, they showed that a beneficial effect of music on creative task performance was limited to a comparison between a silent condition and a so-called “happy music” condition (Vivaldi’s “Four Seasons”). Exposure to “calm music”, “sad music” and “anxious music” had no impact on creative task performance compared to quiet. Participants assigned more positive mood and higher arousal to the happy music condition in comparison to the other conditions and therefore the benefit to creative task performance could have been driven by mood and arousal rather than the presence of the music per se. However, the authors did not report statistical comparisons between all of the music conditions they used within their between-participants design, which potentially undermines their conclusions. Furthermore, Mehta et al. (2012), propose that arousal-based explanations of the impact of to-be-ignored noise on creativity are insufficient because over a longer period of exposure to the sound, physiological arousal levels should normalise and cease to have a consistent influence. Thus, Mehta et al. argue that arousal is not the key contributing factor to the impact of to-be-ignored noise on creativity. They instead propose that moderate noise levels increase *processing disfluency*, with this processing disfluency increasing construal levels and promoting more abstract thinking (nevertheless we return to the mood and arousal account in our introduction to Experiment 3).

The processing disfluency account has its conceptual basis within research on metacognition, which focuses on processes that monitor and control cognition (Ackerman &

Thompson, 2017a, 2017b). Such metacognitive processes are involved in people's subjective judgements of how well a current task is being, could be, or has been performed. Metacognitive control processes about one's current task can be applied to initiate, terminate or change the allocation of time, effort and cognitive resources to the task (Ackerman & Thompson, 2017a). One of a variety of cues on which metacognitive monitoring is based is the subjective ease of processing (fluent *vs*. disfluent; easy *vs*. difficult) that derives from one's own experience at attempting the task. Subjective experiences of task difficulty can catalyse a shift in processing and engender increased task engagement (e.g., Alter, Oppenheimer, Epley, & Eyre, 2007; Rummer, Schweppe, & Schwede, 2016). Attempts to comprehend metacognitive modulation of thought have typically evoked dual-process theories, which posit the existence of two qualitatively distinct types of thinking: Type 1 and Type 2 processes (Evans & Stanovich, 2013a, 2013b). Type 1 processes are autonomous and undemanding of working-memory (a concept used in ways that links to notions of executive and attentional control) and tend to be fast, non-conscious, intuitive and associative. On the other hand, Type 2 processes rely on working-memory (including executive and attentional control) and are focused on cognitive decoupling and mental simulation, critical for hypothetical thinking. Type 2 processes also tend to be slow, conscious, analytic and deliberative. Type 2 processes can be activated if the monitoring system – as part of the metacognitive architecture – judges that a task is difficult (e.g., Bjork, Dunlosky, & Kornell, 2013; see also Thompson, 2010). Metha et al. (2012) argue that the presence of noise creates processing disfluency and supports a processing shift inducing higher construal levels and more abstract thinking that is presumably linked to more diffuse associative processing of the type that is known to arise in creative insight (e.g., Bowden & Jung-Beeman, 1998; Shen et al., in press).

That background sound can *improve* performance on creative tasks contrasts with a large literature relating to distraction of human cognition through exposure to noise (for reviews, see Beaman, 2005; Hughes & Jones, 2003). The task typically used to illustrate the vulnerability of cognition to disruption by the mere presence of to-be-ignored background sound is short-term visual-verbal serial recall (Colle & Welsh, 1976; Jones & Macken, 1993; Salamé & Baddeley, 1982). This task involves the visual presentation of verbal items (e.g., seven or eight letters or digits) with the requirement to recall these items according to the serial order in which they were presented. Initial work suggested that this disruption arose because the sound was comprised of speech. However, the semantic properties of speech were found to be impotent in their capacity to disrupt serial recall: speech presented in a language understood by the participant produces no more disruption than that produced in a language incomprehensible to the participant (Jones, Miles, & Page, 1990). Thus, semantic properties of the to-be-ignored background sound were irrelevant to the level of disruption caused. Similarly, the notion that the disruption by background sound arose due to a confusion between phonemes derived from the visual items (via their covert articulation) that gain direct (spoken items) and indirect (visual items) access into a phonological store (Salamé & Baddeley, 1982) was undermined by findings that serial recall was shown to be susceptible to disruption by the presence of background music without lyrics, and therefore phonemes (Klatte, Kilcher, & Hellbrück, 1995; Klatte, Lachmann, Schlittmeier, & Hellbrück, 2010; Nittono, 1997; Salamé & Baddeley, 1989; Schlittmeier, Hellbrück, & Klatte, 2008), and by the presence of sequences of tones, provided they change from one successive tone to the next (Divin, Coyle, & James, 2001; Elliott, 2002; Jones & Macken, 1993).

The key empirical referent for this so-called "irrelevant sound effect" is the *changing-state effect.* This concerns the finding that a changing sequence of sounds, regardless of whether the changes occur on a speech carrier (e.g., a sequence of different verbal tokens) or a non-speech carrier (e.g., a sequence of tones of different frequency), disrupts serial recall to a far greater extent than a non-changing or steady-state sound (e.g., a repeated token or tone; Jones & Macken, 1993; Jones, Madden, & Miles, 1992). According to the *interference-by-process* approach (e.g., Jones & Tremblay, 2000), the pre-attentive processing of the order of changes within sound impairs the deliberate serial rehearsal process that supports the ordered recall of to-be-remembered items.

Given that the solving of CRATs appears to be underpinned by verbal processes such as subvocal rehearsal (Ball & Stevens, 2009) in addition to executive processes (Chein & Weisberg, 2013) and spreading activation in associative networks (Smith et al., 2013), we expected CRATs to be susceptible to disruption by the presence of to-be-ignored background music. Our rationale behind suggesting that verbal processing of CRATs will be susceptible to changing-state distraction is supported by the findings that impairment of CRAT performance through concurrent articulatory suppression is observed within this procedure, while facilitation of CRAT performance occurs via encouraging verbalisation through the "think aloud" technique (Ball & Stevens, 2009). Thus, the availability of speech (inner speech or external speech) is necessary for efficient CRAT problem solving performance. In the context of serial recall, the skill of speech (or rather speech-planning) is co-opted because it provides an effective medium for retaining visual-verbal items due to its inherent sequentiality, continuity and prosodic and co-articulatory nature. Inner speech therefore enables the grafting of serial order constraints onto the presented list; the act of covertly co-articulating to-be-remembered items generates sequential information and constraints that do not occur within the list itself. However, this motoric based serial rehearsal process is subject to interference from the automatic, pre-attentive processing of the serial order of changes in a background auditory sequence (e.g., such as music).

In our recent work it is becoming clearer that tasks that may not necessarily involve serial rehearsal, but that do involve the use of inner speech for effective task performance (e.g., face recognition; Marsh, Nakabayashi, Frowd, Skelton, & Ball, 2018) are also vulnerable to the changing-state effect. Of course, speech (inner or outer) involves planning of sequential motor-acts which may render it vulnerable to disruption via changing-state speech in many settings. In the context of CRATs, inner (and outer) speech clearly supports effective performance (Ball & Stevens, 2009). It may even be that participants use serial rehearsal to test out novel solutions. For example, for the problem ("house", "pear", "family" and the solution "tree") it may be the case that "tree-house, pear-tree" might be rehearsed while participants test out the veracity of a final problem word "family-tree". That solution words can either serve as a prefix or suffix to problem words may reinforce this rehearsal strategy since order processing is necessary to obtain an appropriate solution ("tree-pear" being an appropriate combination of problem and solution word but in the reverse order).

We do not claim that CRAT performance is underpinned *entirely* by verbal maintenance processes. Rather, it is clear that spreading semantic activation processes (Smith et al., 2013), and executive processes that are involved in generating response candidates and inhibiting misleading/incorrect solutions (Storm & Angello, 2010) are also central to the production of responses. That said, it is often not clear which component of a multi-component task is associated with CRAT performance. For example, the finding of an association between Working Memory Capacity (WMC) measures and CRAT performance (Chein & Weisberg, 2014) could be due to the role of attentional/cognitive control (which may involve executive control processes such as inhibition) or the requirement to retain serial order information: WMC tasks involve combining the short-term storage of visual/verbal items with a concurrent processing task. Therefore, we hold that sub-vocal maintenance processes involving inner speech can underpin solution of CRATs and that this process is susceptible to disruption via processing of a changing-state auditory sequence. The aim of the series of three experiments that we present here was to investigate the impact of to-be-ignored background sound (i.e., music with foreign, [unfamiliar] lyrics; instrumental music; and music with familiar lyrics) on tasks believed to measure creativity, that is, CRATs.

Experiment 1

The aim of Experiment 1 was to establish if music with unfamiliar lyrics facilitates creativity as measured via performance with CRATs through engendering processing disfluency (Mehta et al., 2012), or impairs creativity, possibly due to the disruption of verbal processes such as subvocal rehearsal (Ball & Stevens, 2009) due to its changing-state characteristics. To investigate these two competing accounts and opposing predictions, we contrasted performance in a quiet condition to a condition with to-be-ignored background music with clearly discernible lyrics in a foreign language that were unfamiliar to the participants (i.e., the musical excerpt contained Spanish lyrics presented to native English speakers, thus the lyrics were both *unfamiliar*, and *meaningless* to the participants, who were unable to process their semantic content). As such, any observed disruption could not be viewed as being attributable to interference between the semantic properties of the to-be-ignored sound and the semantic processes underpinning the solving of CRATs (see Marsh, Hughes, & Jones, 2008, 2009). If unfamiliar music engenders processing disfluency (cf. Mehta et al., 2012) then one should observe better performance when music, as compared with quiet, accompanies problem solving. However, if verbal processes underpinning CRAT performance are susceptible to disruption via changing-state irrelevant sound – as the interference-by-process account would assume (Jones & Tremblay, 2000) – then performance should be poorer in the presence of music as compared with quiet.

Method

Participants

Thirty adults (15 female, 15 male) from the University of Central Lancashire participated in the experiment (*M* = 22 years, *SD* = 2.78, age range 19 to 30 years old). The participants were recruited via an opportunity sample. Participants received course credit or the standard department payment rate in exchange for 30 minutes of participation time. All participants spoke English as their first language and reported normal (or corrected-to-normal) vision and hearing. The experiment received Ethical Clearance from the University of Central Lancashire.

Design and Materials

The design was a fully within-participants 2 × 2 design with Sound (Quiet vs. Spanish Music) and CRAT Difficulty (Easy vs. Difficult) as the factors. A set of 38 CRATs were selected from the problems developed by Bowden and Jung-Beeman (1998) using the program "Match" (Van Casteren & Davis, 2007). Match automates the selection of several groups of smaller stimuli sets from a larger pool ensuring the groups are matched on multiple dimensions. Here, the sets of CRATs were matched on solution accuracy and solution time data provided by Bowden and Jung-

Beeman (1998). Each CRAT consisted of the presentation of three single words, with the participant having to find a word that combines with each of the three presented words to make a common word or phrase. For example, if participants are presented with the words *stick / maker / point*, then the word that links with each of these is the word *match* to create the phrases or words *matchstick*, *match maker* and *match point*. Therefore, the answer or target in this instance would be the word "match".

The 38 selected CRATs were divided into a set of 20 easy CRATs (Easy CRATs' solution rates: M = 68.9%, SD = 16.2, Easy CRATs' solution times: M = 8.23s, SD = 2.61) and 18 difficult CRATs (Difficult CRATs' solution rates: M = 26.6%, SD = 14.3, Difficult CRATs' solution times: M = 12.1s, SD = 3.07). Each difficulty set was then further divided into two equal matched sets (Set A solution rates: M = 49.4%, SD = 26.1, solution times: M = 9.98s, SD = 3.6 vs. Set B solution rates: M = 48%, SD = 27, solution times: M = 10.1s, SD = 3.3) as indexed by normative data on solution rate and solution time data for 30 second presentation time provided by Bowden and Jung-Beeman (1998). It is typical to divide CRAT sets into easy and difficult (Ball & Stevens, 2009). Difficult, but not easy, problems can benefit from overt verbalisation, whereas preventing subvocalisation via requiring participants to suppress articulation while problem solving can hinder performance with both easy and difficult problems (Ball & Stevens, 2009). Although not a primary goal of the study, we nevertheless considered it important to investigate the potential differential susceptibility to distraction of easy and difficult problems. The experiment was fully counterbalanced such that each CRAT set appeared within each sound condition.

Procedure

Participants read an information sheet and completed a consent form prior to taking part in the experiment. Participants were given instructions for the CRATs that explained the need to find one target word per problem, that, when combined with the three presented words (either before or after the presented words), created a common word or phrase. Prior to the test problems, participants were asked to tackle five practice problems to ensure familiarity with the task. These practice problems were also selected from Bowden and Jung-Beeman (1998). Participants were allocated 30 seconds per CRAT item. All three problem words were presented simultaneously along the same horizontal plane.

The music was played via Sennheiser HD-202 headphones at approximately 65-70 dB(A). The music was a 30 second segment of a Spanish translation of a 1990's UK chart pop song played via E-Prime Software that contained clearly discernable lyrics and accompanying instruments. The music contains appreciable acoustic variation and satisfied the criterion for being a changing-state stimulus. In the sound condition, this music segment accompanied each CRAT problem, starting with the onset of the problem and ending once the participant indicated they had solved the CRAT by pressing the spacebar. After participants pressed the spacebar, a textbox appeared wherein participants typed their answer. Participants were asked to complete the CRATs while ignoring the background sound. They were also reassured that they would not be asked anything about the background sound. Participants were fully debriefed at the end of the task and thanked for their participation. During debriefing, participants were presented with the auditory stimuli they were exposed to during the experiment and asked if they were familiar with the song or had heard it before in experiment, none replied that they were or had.

Results

The data for each of the three experiments can be found via the following link; https://osf.io/j6hwd/. A 2 (Sound: Quiet vs. Spanish Music) × 2 (CRAT Difficulty: Easy vs. Difficult) within-participants analysis of variance (ANOVAs) was conducted on mean solution rates (i.e., proportion correct). An alpha level of $p < .05$ was adopted for all statistical tests. There was a significant main effect of Sound, $F(1, 29) = 9.91$, $MSE = .01$, $p = .004$, $\eta_p^2 = .25$. Significantly more CRATs were solved in the quiet condition ($M = .43$, $SE = .04$) in comparison to the Spanish music condition ($M = .36$, $SE = .05$). As anticipated, there was a main effect of CRAT Difficulty, $F(1, 29) = 63.36$, $MSE = .02$, $p < .001$, $\eta_p^2 = .67$, with significantly higher solution rates for the easy CRATs ($M = .49$, $SE = .04$) in comparison to the difficult CRATs (*M* = .29, *SE* = .05). There was no significant Sound × CRAT Difficulty interaction, $F(1, 29) = 1.36$, $MSE = .02$, $p = .25$, $\eta_p^2 = .05$.

Discussion

The aim of Experiment 1 was to compare CRAT performance in a quiet condition to performance when ignoring background music with unfamiliar foreign lyrics. Significantly more CRATs were solved in the quiet condition in comparison to the background music condition. As anticipated, there was a significant difference in solution rates for the easy versus the difficult CRATs. Furthermore, there was no significant interaction between Sound and CRAT Difficulty for solution rates. That Sound impaired CRAT performance regardless of problem difficulty coheres with the findings of Ball and Stevens (2009), who demonstrated that articulatory suppression similarly impaired performance with easy and difficult tasks. However, the results are inconsistent with the general view that music enhances creativity, and dispute the prediction that background noise enhances creativity due to the promotion of processing disfluency and subsequent encouragement of abstract thought (Mehta et al., 2012). That CRAT performance was disrupted by the presence of a stimulus that conveyed no meaning to the participants precludes a semantic interference-by-process explanation of the results (cf. Marsh et al., 2008, 2009). However, the results are consistent with previous findings, which demonstrate background sounds that are meaningless to participants, can impair performance of tasks that require verbal working memory components such as serial recall, providing they possess appreciable changing-state properties (e.g., Jones, Miles, & Page, 1990): an acoustic interference-by-process (Jones & Tremblay, 2000).

Experiment 2

Experiment 1 identified that music with foreign (unfamiliar) lyrics had a detrimental effect on the solution rates of CRATs in comparison to a quiet condition. At first glance this finding is at odds with the notion that creative performance can be enhanced in the presence of background sound, through encouraging processing disfluency and promoting the abstract thought believed to be required to solve CRATs. However, Mehta et al. (2012) obtained facilitatory effects on CRAT performance with ambient noise comprising *"multi-talker noise in a cafeteria, roadside traffic, and distant construction noise to create a soundtrack of constantly varying background noise"* (p. 786), whereas the current study used Spanish music with a clearly discernible voice (although in a language foreign to the participant, thereby conveying unfamiliar lyrics). We note that the presence of a clearly discernible voice could be a key difference between our study and that of Mehta et al. (2012) in driving the direction of the effect of background sound on creative performance. One possibility is that the presence of discernible speech in Experiment 1 could somehow prevent participants from achieving the disfluent processing state that could facilitate CRAT performance through abstract thought.

To address this aforementioned issue, Experiment 2 compared a quiet background with music without speech (lyrics) to investigate whether the presence of speech in some way impedes disfluent processing and thus prevents any supposed benefits of such disfluent processing on CRAT performance. In terms of the contrasting view that background sound can impair creative processing through impairing verbal working memory, music without lyrics should impair CRAT performance similarly to Spanish music with a discernible voice. In support of this view, numerous studies in the context of serial-verbal short-term memory (Klatte et al., 1995; Klatte et al., 2010; Nittono, 1997; Salamé & Baddeley, 1989; Schlittmeier et al., 2008) have shown that the presence of speech is not a prerequisite to produce disruption of verbal working memory. Thus, on the interference-by-process account, music without lyrics (speech) would also be expected to disrupt the creative processes necessary for solving CRATs and Experiment 2 sought to determine whether this was indeed the case.

Method

Participants

Eighteen adults (12 female, 6 male) from the University of Central Lancashire aged between 19 and 45 years old participated in the experiment (M = 25 years, SD = 9.31). The participants were recruited via an opportunity sample and received course credit, or the standard department payment rate in exchange for 30 minutes of participation. All participants spoke English as their first language and reported normal (or corrected-to-normal) vision and normal hearing. The experiment received Ethical Clearance from the University of Central Lancashire.

Design and Materials

The design and materials were identical to those outlined in Experiment 1 above, with the exception of a manipulation to one of the levels of the within-participant factors, Sound, which had two levels: Quiet *vs*. Music without Lyrics. The sound used within Experiment 2 was therefore the same as that used in Experiment 1, but without the lyrical content.

Procedure

Each participant read an information sheet and signed a consent form prior to beginning the experiment. The procedure remained identical to that reported in Experiment 1 outlined above. All participants were fully debriefed at the end of the experiment. As with Experiment 1, at debriefing participants were presented with the auditory stimuli they were exposed to during the experiment and asked if they were familiar with the song or had heard it before the experiment. None of the participants reported familiarity with the song, nor hearing it previously.

Results

A 2 (Sound: Quiet vs. Music without Lyrics) x 2 (CRAT Difficulty: Easy vs. Difficult) within participants analysis of variance (ANOVA) was conducted on the dependent variable of mean solution rate. An alpha level of $p < .05$ was adopted for all statistical tests. There was a significant main effect of Sound, $F(1, 17) = 8.60$, $MSE = .02$, $p = .009$, $\eta_p^2 = .34$. Significantly more CRATs were solved in the Quiet condition ($M = .39$, $SE = .04$) in comparison to the Music without Lyrics condition ($M = .29$, $SE = .03$). As anticipated, there was a main effect of CRAT Difficulty, $F(1, 17) = 61.05$, $MSE = .03$, $p < .001$, $\eta_p^2 = .78$, with significantly higher solution rates for the easy CRATs ($M = .48$, $SE = .05$) in comparison to the difficult CRATs ($M = .19$, $SE = .03$). There was no significant Sound × CRAT Difficulty interaction, $F(1, 17) = 2.51$, $MSE = .02$, $p = .13$, $\eta_p^2 = .13$.

Discussion

The aim of Experiment 2 was to compare CRAT performance in a quiet condition versus performance with to-be-ignored background music without lyrics. This was to investigate whether the presence of speech in Experiment 1 produced disruption of CRAT performance, and if this effect would hold for music without any speech content (in other words, music without lyrics). Experiment 2 supported the findings of Experiment 1 in that significantly more CRATs were solved in the quiet condition in comparison to the music without lyrics condition. Consistent with Experiment 1, there was a significant difference in solution rates for the easy versus the difficult CRATs. Furthermore, there was no significant interaction between Sound and CRAT Difficulty for solution rates.

The results are again inconsistent with the general view that music enhances creativity, and instead we demonstrate a deficit to CRAT performance in the presence of to-be-ignored background music with unfamiliar lyrics (Experiment 1) as well as in the absence of lyrics (Experiment 2). Moreover, the results oppose the view that background noise leads to processing disfluency, which in turn promotes creativity by engendering increased abstract thought (Mehta et al., 2012). The results run counter to the idea that the failure to find facilitation of CRAT performance via background music was due to the presence of speech in Experiment 1. Here we demonstrate in Experiment 2 that music without lyrics (i.e., in the absence of any speech content) still failed to produce a facilitation in creativity, and in fact resulted in a decrement in creativity in comparison to a quiet condition.

Taken together, the findings of Experiments 1 and 2 support the notion that verbal working memory is necessary for CRAT performance (Ball & Stevens, 2009), and that this is susceptible to disruption by the presence of to-be-ignored background sound, regardless of whether this background sound is speech or non-speech based (e.g., Jones & Macken, 1993).

Experiment 3

Both the interference-by-process view (Jones & Tremblay, 2000) and the processing disfluency view (Mehta et al., 2012) eschew the role of mood and arousal in mediating the effect of background sound on creative task performance. However, there is a literature showing that increased mood and arousal that derives from listening to music may affect cognitive task performance. For example, Thompson et al. (2001) demonstrated a benefit to subsequent visual-spatial task performance from *prior* listening to music as compared to exposure to quiet that was entirely dependent on the change in mood and arousal that the music produced. Furthermore, Ritter and Ferguson (2017) reported that music presented 15s prior to, and concurrently with, the performance of a task tapping creative cognition facilitated performance on that task. However, this facilitatory effect occurred only for “happy” music that engendered positive affect and increased arousal. Therefore, it remains possible that given the pleasure that individuals usually derive from music, the music one *choses* to listen to might typically induce a positive mood and increased arousal yielding a positive impact on task performance (Thompson et al., 2001), particularly for tasks that involve creativity (Ritter & Ferguson, 2017). Indeed, previous research has established that positive mood can improve performance on RATs (Rowe, Hirsch, & Anderson, 2007). In both Experiments 1 and 2 our use of arbitrary music with foreign or “unfamiliar” lyrics, and music with the absence of lyrics, could have induced a neutral or even *negative* mood state in participants, which might have hindered the emergence of creative insight.

To investigate any potential mediating impact of mood on CRAT performance in Experiment 3, participants tackled CRAT problems in the presence of music with positive lyrics and upbeat melody that we considered should increase positive affect and arousal. To identify support for this assertion, we measured mood states at two different time points (i.e., before and after each background sound condition) using the Profile of Mood States (PoMS) questionnaire (McNair, 1971). In Experiment 3 we also acquired data relating to participants’ musical preferences (i.e., whether they liked or disliked the presented background sound) and their study habits (i.e., whether they tend to study with music or without background music). These data were intended to be peripheral to the main findings, but they nevertheless had the capacity to provide an indication of whether the impact of mood on CRAT performance is influenced by either musical preference or study habits.

We also note that in explaining the findings arising in Experiments 1 and 2 yet another possibility is that the promotion of creativity through processing disfluency in the presence of background noise (Mehta et al., 2012) is a specific effect that is limited to the presence of relatively "steady-state" sound – unlike the background music conditions used in our conditions which clearly satisfy the criteria for changing-state sound. Therefore, in Experiment 3 we included a "library noise" condition that resembled that used by Mehta and colleagues (Mehta et al., 2012). We contrasted this library noise condition with a music condition (i.e., popular music with familiar lyrics) and with a quiet condition.

In terms of the outcomes of Experiment 3, if the mood and arousal account (Ritter & Ferguson, 2017; Thompson et al., 2001) is correct, then we expected to observe an increase in CRAT performance in the presence of the background music condition compared to the quiet and library noise condition, assuming that the music condition reliably increases mood and arousal compared to the library noise condition. We note here that studies exploring the mood and arousal effect usually present music prior to, rather than concurrently with, the cognitive task of interest (Thompson et al. 2001). However, effects of music on creative task performance that are reportedly mediated through mood and arousal have also been shown when music is presented concurrently with the target task (Ritter & Ferguson, 2017). Moreover, participants within mood and arousal studies are free to attend the music, rather than instructed to ignore the background sound, as is the case with studies of the irrelevant sound effect (Jones & Macken, 1993). We make the assumption, however, that changes to mood and arousal induced by the presentation of music occurs regardless of whether participants are free to attend the music or requested to ignore it and explore this proposition.

The processing disfluency account (Mehta et al., 2012) would predict that both library noise and music conditions should increase CRAT performance, while the modified processing disfluency account only predicts a positive effect of background library noise on CRAT performance. Finally, the interference-by-process view (e.g., Jones & Tremblay, 2000) predicts that CRAT performance should be reduced in the music condition relative to the library noise and quiet condition because this condition comprises a changing-state auditory stimulus, the library noise condition constituting a steady-state stimulus.

Method

Participants

Thirty-six adults (23 female, 13 male) from the University of Central Lancashire aged between 19 and 56 years old participated in the experiment (*M* = 24 years, *SD* = 8.36). The participants were recruited via an opportunity sample. Participants received course credit, or the standard department payment rate in exchange for 30 minutes of participation. All participants spoke English as their first language and reported normal (or corrected-to-normal) vision and hearing.

Design

In relation to the assessment of CRAT performance the design was a 3 (Sound: Quiet *vs*. Music *vs*. Library Noise) × 2 (CRAT Difficulty: Easy *vs*. Difficult) × 2 (Musical Preference: Like *vs*. Dislike) × 2 (Study Habit: Music *vs*. No Music) mixed design. For the purpose of mood evaluation, the following within-participants design was used to determine mood changes using the PoMs questionnaire: 3 (Sound: Quiet *vs*. Music vs. Library Noise) × 2 (Time: Before *vs*. After) × 6 (Mood State: Tension *vs*. Depression *vs*. Anger *vs*. Confusion *vs*. Fatigue *vs*. Vigour). Music chosen for background sound was a popular 2013 release, midtempo soul and neo-soul song which contained positive lyrics and upbeat melody. The library noise consisted of distant (non-intelligible) speech, photocopier noise, typing, and rustling of papers.

Materials

Before undertaking the CRATs, participants were asked: “Do you ordinarily study in the presence of background music?” and responded yes or no. The PoMS questionnaire is designed to measure fluctuating feelings and affective states (for further details see McNair, 1971). The questionnaire measures six different aspects of mood state: tension, depression, anger, confusion, fatigue and vigour. According to instructions of administration, the six mood states can be combined in the following way to produce a Total Mood Disturbance (TMD) score: tension + depression + anger + confusion + fatigue – vigour. However, for the purposes of this design we were interested in the specific mood profile and therefore the six specific profile scores were used rather than a general TMD measure (McNair, 1971).

Using the norming data on solution rate and solution time for 30 second presentation time, an additional set of 19 CRAT problems (10 easy, 9 difficult) matching accuracy and solution times to Set A and Set B, were selected using the program “Match” (Van Casteren & Davis, 2007) to create Set C (solution accuracy *M* = 47.9%, *SD* = 25.7, solution times: *M* = 9.6s, *SD* = 3.3). The experiment was fully counterbalanced such that each CRAT set appeared within each sound condition. After undertaking the CRATs, participants were asked: “Did you like the music?” and responded yes or no.

Results

Like Experiments 1 and 2, the dependent variable was the mean solution rate for the CRAT problems. As mentioned in the foregoing, Experiment 3 included a number of further dependent variables. These were responses to the PoMS questionnaires administered before and after the completion of each set of CRATs. The PoMS contains measures of six mood states: tension, depression, anger, confusion, fatigue and vigour. There was also a brief questionnaire related to musical preference (whether participants liked the music played during the music condition) and study habits (whether they regularly listened to music when studying). Twenty-nine participants responded that they liked the music and 7 responded that they disliked the music. Furthermore, 18 participants responded that they ordinarily studied in the presence of music, while 18 preferred to study without the presence of music. Participants were assigned to Like *vs*. Dislike for Musical

Preference and Music vs. No Music for Study Habit, accordingly. An alpha level of $p < .05$ was adopted for all statistical tests used.

Solution Rates

A 3 (Sound: Quiet vs. Music vs. Library Noise) × 2 (CRAT Difficulty: Easy vs. Difficult) × 2 (Musical Preference: Like vs. Dislike) × 2 (Study Habit: Music vs. No Music) mixed ANOVA was conducted on the solution rate data. There was a significant main effect of Sound on solution rates, $F(2, 64) = 3.79$, $MSE = .07$, $p = .03$, $\eta_p^2 = .11$. Pairwise comparisons revealed that significantly more CRATs were solved in the Quiet condition ($M = .34$, $SE = .05$) in comparison to the Music condition ($M = .28$, $SE = .05$, $p < .05$). There were also significantly more CRATs solved in the Library Noise condition ($M = .36$, $SE = .05$, $p < .05$) in comparison to the Music condition. However, there was no significant difference between the mean number of CRATs solved in the Quiet and Library Noise conditions ($p = .60$).

As expected, there was a significant main effect of CRAT difficulty on solution rates, $F(1, 32) = 95.75$, $MSE = .02$, $p < .001$, $\eta_p^2 = .75$, with significantly more easy CRATs solved ($M = .46$, $SE = .05$) than difficult CRATs ($M = .19$, $SE = .05$). There was no significant main effect of Musical Preference on CRAT solution rates, $F(1, 32) = .09$, $MSE = .26$, $p = .76$, $\eta_p^2 = .003$. Participants who specified liking the music ($M = .34$, $SE = .04$) did not solve significantly more CRATs than those who specified a dislike for the music ($M = .31$, $SE = .09$). There was also no significant main effect of Study Habit on CRAT solution times, $F(1, 32) = .09$, $MSE = .26$, $p = .77$, $\eta_p^2 = .003$. Participants who specified that they preferred to study without music ($M = .31$, $SE = .08$) did not solve significantly more CRATs than those who specified that they preferred to study with music ($M = .34$, $SE = .05$).

The remaining interactions and three-way interactions all failed to reach significance, with the exception of the Sound × Difficulty interaction, $F(2, 64) = 3.93$, $MSE = .01$, $p = .03$, $\eta_p^2 = .11$. Pairwise comparisons revealed the source of the Sound × Difficulty interaction. For the easy CRATs there were significantly more CRATs solved in the Quiet condition ($M = .51$, $SE = .06$) in comparison to the Music condition (*M* = .38, *SE* = .06, $p < .001$) and significantly more solved in the Library Noise (*M* = .49, *SE* = .06) condition in comparison to the Music condition ($p = .01$). However, there was no significant difference between the number of easy CRATs solved in the Quiet and Library Noise conditions ($p = .70$). This suggests that CRAT difficulty only influenced both the Quiet and Library Noise conditions. All further pairwise comparisons failed to reach significance (all *p*s > .05).

Figure 1. The Sound × Difficulty interaction with standard error bars.

Profile of Mood States (PoMs) Questionnaire

To ascertain any changes in mood before and after completing the CRATs in each sound condition, a PoMS questionnaire was administered to participants at two different points (before and after completing the CRATs in each of the three sound conditions). Therefore, each participant completed the PoMS questionnaire a total of six times. A 3 (Sound: Quiet vs. Music vs. Library Noise) × 2 (Time: Before vs. After) × 6 (Mood State: Tension vs. Depression vs. Anger vs. Confusion vs. Fatigue vs. Vigour) within-participants ANOVA was conducted on the mood state scores. The ANOVA revealed that there was no significant main effect of Sound, $F(2, 70) = 1.66$, $MSE = .20$, $p = .20$, $\eta_p^2 = .05$, with no significant difference in the mean mood state score in the Quiet (*M* = 5.89, *SE* = .54), Music (*M* = 5.65, *SE* = .58) and Library Noise (*M* = 6.10, *SE* = .54) conditions. There was a main effect of Time, $F(1, 35) = 6.10$, $MSE = 9.98$, $p = .02$, $\eta_p^2 = .15$, with mean mood state scores significantly higher Before (*M* = 6.10, *SE* = .58) in comparison to After (*M* = 5.66, *SE* = .49) completing the CRATs. There was a significant main effect of Mood, $F(5,175) = 28.17$, $MSE = 80.60$, $p < .001$, $\eta_p^2 = .45$. Pairwise comparisons indicated that there were significant differences between all mood states with the exception of tension (*M* = 5.05, *SE* = .67) versus fatigue (*M* = 5.13, *SE* = .78), depression (*M* = 2.86, *SE* = .87) versus anger (*M* = 5.13, *SE* = .78), anger versus fatigue, and fatigue versus vigour (*M* = 11.96, *SE* = .91) (all *ps* > .05). There was no significant interaction between Sound × Time, $F(2,70) = 2.43$, $MSE = 11.88$, $p = .10$, $\eta_p^2 = .07$, Sound × Mood, $F(10,350) = 1.42$, $MSE = 8.25$, $p = .10$, $\eta_p^2 = .07$, or Time × Mood, $F(5,175) = .48$, $MSE = 5.18$, $p = .79$, $\eta_p^2 = .01$.

As expected, there was a significant three way interaction between Sound × Time × Mood, $F(10, 350) = 3.50$, $MSE = 5.846$, $p < .001$, $\eta_p^2 = .09$. Pairwise comparisons indicated the source of this significant interaction. For the Quiet condition, there was no significant change in any of the mood states after completing the CRATs (all *p*s > .05). However, for the Music condition, there was a significant change in mood for four out of the six mood states. There was a significant decrease in tension ($M = 5.36$, $SE = 0.82$ vs. $M = 4.00$, $SE = 0.69$, $p = .01$), anger ($M = 4.08$, $SE = 1.04$ vs. $M = 2.72$, $SE = 0.69$, $p = .049$), confusion ($M = 6.92$, $SE = 0.75$ vs. $M = 5.69$, $SE = 0.59$, $p = .02$) and fatigue ($M = 4.81$, $SE = 0.87$ vs. $M = 3.72$, $SE = 0.80$, $p = .03$). There was no significant change in depression ($M = 3.25$, $SE = 1.23$ vs. $M = 2.61$, $SE = 0.98$, $p = .33$), or vigour ($M = 11.86$, $SE = 0.97$ vs. $M = 12.72$, $SE = 1.13$, $p = .22$). For the Library Noise condition, there was no significant change in five of the six mood states (all *p*s > .05). However, there was a significant decrease in vigour ($M = 12.67$, $SE = 1.08$ vs. $M =10.67$, $SE = 0.94$, $p < .001$).

These findings indicate that the significant changes in mood states before and after completing CRATs occurred within the Music condition and not in the Quiet or Library Noise condition, thus indicating that music altered a number of mood states, and as measured by the PoMS, provided a general increase in *positive* mood. Since previous research (Rowe et al., 2007) has shown that positive mood can improve performance on RATs, the present observation that music increased mood but *decreased* CRAT performance suggests that a mood-based explanation of the detrimental effect of music on creative insight seems implausible.

Profile of Mood States and Solution Rates

In the previous section the PoMS scores and CRAT solution rates were examined independently. However, it is useful to consider the possible impact of mood as a mediating influence on the relationship between to-be-ignored background sound and CRAT performance. Unfortunately, the current dataset is unsuitable for mediation analysis (i.e., to ascertain mood as a possible direct or indirect mediator in the relationship between background sound and CRAT performance) given the implementation of Sound as a within-participants factor rather than a between-participants factor. However, an Analysis of Covariance (ANCOVA) was performed to examine CRAT solution rates when mood score was included as a covariate. Here we focused on the relationship between the quiet and music conditions, given our particular interest in the disruption caused by to-be-ignored background music.

In order to establish the mood score for entry as a covariate, the "before" score for each of the six mood score dimensions (tension, depression, anger, confusion, fatigue and vigour) was subtracted from the "after" score to provide a "mood change" score for each of the six profile of mood state dimensions listed above. This resulted in a mood state dimension change score for both the music and quiet conditions, for each mood state (tension, depression, anger, confusion, anger and vigour). The change score for the music condition was subtracted from the change score for the quiet condition, to provide a single change score for each of the six mood states. A 2 × 2 ANCOVA was conducted (Sound: Quiet vs. Music) × 2 (CRAT Difficulty: Easy vs. Difficult) on solution rates, with tension, depression, anger, confusion, fatigue and vigour each entered as a covariate. The findings revealed that all covariates failed to reach significance (all *p*s >.05), indicating that each of the six mood state measures failed to have a significant impact on CRAT solution rates, either directly or in interaction with the Sound and CRAT Difficulty factors.

Discussion

Experiment 3 demonstrated that popular music with familiar lyrics disrupted CRAT performance (in terms of solution rates) in comparison to a quiet condition or library noise condition. However, there was no significant difference in CRAT performance between the quiet and library noise conditions. Mehta et al. (2012) previously demonstrated a beneficial effect to creativity with what could be termed "steady-state sound". Whilst these findings demonstrate that a steady-state sound such as library noise does not result in a relative enhancement to creativity, we also find that the decrement was not significant, particularly in comparison to the background music with familiar lyrics.

Furthermore, the findings imply that music with familiar lyrics resulted in a decrement in CRAT performance, despite an apparent overall positive increase in mood as identified by six mood states recognised in the PoMS. Given that previous research has identified that positive mood can lead to an improvement in RAT scores (Rowe et al., 2007), the findings here demonstrate that the decrement in performance in the music condition does not appear to be driven by mood. Indeed, these findings further support the notion that CRAT performance relies on verbal working memory, and that this is susceptible to disruption by non-steady-state sound, with or without the presence of speech.

General Discussion

In a series of three experiments we investigated the impact of background music (with varying semantic properties) on creativity, as measured by performance on CRATs. In Experiment 1, background music with foreign (and therefore "unfamiliar") lyrics resulted in a significant decrement to CRAT performance; significantly fewer CRATs were solved in the music with foreign lyrics condition in comparison to the quiet condition. This finding was replicated in Experiment 2 with the implementation of background instrumental music in comparison to a quiet condition. In Experiment 3 familiar music was found to impair CRAT performance regardless of whether the music induced a positive mood or whether those participants typically studied in the presence of music. Moreover, disruption occurred despite the fact that the music was liked by the participants, which coheres with the findings of Perham and Vizard (2010) and Perham and Currie (2014), who showed equivalent disruption by liked and disliked background music in the context of serial recall and reading comprehension, respectively.

Prior to discussing the implications of Experiments 1 to 3 for the theoretical accounts entertained, we undertook a Bayesian meta-analysis of the collective findings. Bayes factors were calculated to quantify the evidence for two hypotheses: the hypothesis from the processing disfluency account (cf. Mehta et al., 2012) that CRAT performance would be better in the music conditions relative to the quiet conditions (H1), and the hypothesis from the auditory distraction account (Jones & Tremblay, 2000) that CRAT performance would be better in the quiet conditions relative to the music condition (H2).

To calculate Bayes factors, one must specify the plausibility of effect sizes given one's theory. For both H1 and H2, we model the sort of effect size considered plausible on the results of Mehta et al. (2012), who reported that hearing a moderate level of noise resulted in participants solving a significantly greater proportion of RATs (M = .73) than participants in the quiet control condition (M = .56). These results provide an approximate effect size that could be expected for a noise manipulation (such as music) on measures of creativity, from .73 to .56 = .17. Following Dienes' (2011, 2014) guidelines, the experimental hypotheses in the current analyses were modelled using a half-normal distribution with a mode of 0 and a standard deviation of .17. Bayes factors < 0.33 are interpreted as moderate evidence for the null hypothesis and Bayes factors > 3 as evidence for the experimental hypotheses. Bayes factors around 1 are conventionally considered inconclusive (Dienes, 2011, 2014). Bayes factors were calculated using Dienes and McLatchie's (2018) calculator.

The results of the Bayes factors for each study are presented in Table 1. There was strong evidence across all three studies for H0 relative to H1, indicating that playing music did not enhance creativity as measured by CRAT performance. In contrast, there was moderate to strong evidence across all three studies for H2 relative to H0, suggesting that music decreased creativity to approximately the same extent one might have expected it to have increased by. The right-most column in Table 1 provides the Bayes factors comparing H2 to H1, and they indicate strong evidence for H2, that listening to music diminishes rather than increases CRAT performance.

Table 1. Bayes factors testing the effect of Music versus Quiet conditions from Experiments 1 to 3

The experiment-level analyses were followed up with a fixed-effects meta-analysis using Dienes' (2008) calculator (see Goh, Hall & Rosenthal, 2016, for an overview of the benefits of such internal meta-analyses within studies). The meta-analytic posterior mean and standard deviation, along with the 95% credible intervals, are shown in Table 2. These provide the best estimate of the population parameter and its uncertainty following the completion of all three studies (meta-analytic results following each study are reported on each row of Table 2). The meta Bayes factors, calculated on the results across all three studies to test H1 and H2, revealed that the overall body of evidence provides substantial support for H2, that listening to music diminishes rather than enhances creativity, $B_{H1/H2} = 3.36 \times 10^4$, thereby supporting the interference-by-process account (Jones & Tremblay, 2000) over the processing disfluency account (Mehta et al., 2012).

Taken together, the findings from Experiment 1 to 3, supported by our Bayesian meta-analysis, contradict the popular opinion that background music enhances creativity. Instead, they demonstrate that background music, with or without familiar semantic content (i.e., lyrics), or in the absence of speech, disrupts performance on CRATs, which represent a class of highly researched verbal problem solving tasks that are often solved creatively through insight-based processes. Furthermore, the findings of Experiments 1 to 3 undermine the processing disfluency account (Mehta et al., 2012), which predicts superior performance in the presence of moderate intensity background noise in comparison to quiet. We note that a reprieve for the account might be offered if one were to assume that music (at least pleasant music) impairs creativity via inducing processing *fluency* (Mehta et al., 2012, p. 796). However, it is not immediately obvious why, on the processing disfluency account, noise and music should differ in relation to the processing fluency that they hypothetically engender. The experiment-level analyses were followed up with a fixed-effects meta-analysis using Dienes' (2008) calculator (see Goh, Hall & Rosenthal, 2016, for an overview of the benefits of including internal meta-analyses within studies). The meta-analytic posterior distribution (*M* = -0.06, *SD* = 0.01, *95%CI* [-0.08, -0.03] provide the best estimate of the population parameter and its uncertainty in light of all three studies. The meta-analysis suggests that music reduces creativity as measured using the CRAT. Bayes factors were calculated on the meta-analytic to test H1 and H2, and revealed that the overall body of evidence provided substantial support for H2, that listening to music diminishes rather than enhances creativity, $B_{H(0, .17)} = 3.36 \times 10^4$, thereby supporting the interference-by-process account (Jones & Tremblay, 2000) over the processing disfluency account (Mehta et al., 2012).

We contend that the deficit in CRAT performance in the presence of background music appears altogether more consistent with the interference-by-process framework (e.g., Marsh et al., 2009) than with the processing disfluency account (Mehta et al., 2012). According to the interference-by-process approach, the disruption of CRAT performance is attributable to the *changing-state effect*, which refers to the finding that a changing sequence of sound (regardless of whether the changes occur on a speech or non-speech carrier), disrupts serial recall to a far greater extent than a non-changing or steady-state sound (e.g., a repeated token or tone; Jones & Macken, 1993; Jones et al., 1992). The pre-attentive perception of changes between elements in the sound, as a by-product of acoustic-based perceptual organisation processes (Bregman, 1990), gives rise to irrelevant order cues. These cues compete with the process responsible for subvocally maintaining the to-be-remembered items in sequence. In support of this suggestion, we have recently found that changing-state letters (c, t, g, u) produce more disruption to CRAT performance than steady-state letters (c, c, c, c; Marsh, Threadgold, Barker, & Ball, 2017). Music, of course, is a changing-state, rather than a steady-state, sound. Therefore, the findings presented here, which attest to the disruption to verbal insight problem solving as measured through CRAT performance, are entirely consistent with findings that have revealed a disruption to serial recall by changing-state sounds that include music (e.g., Perham & Vizard, 2010). Furthermore, the findings imply that the presence or absence of semantic content (i.e., lyrics) – and indeed the familiarity of the lyrics (e.g., lyrics in foreign and unfamiliar language) – does not alter the disruptive influence of background music. Moreover, that background music successfully increased mood and arousal in Experiment 3, but led to poorer, rather than better, CRAT performance compared with the quiet control and library noise condition undermines the mood and arousal account (Ritter & Ferguson, 2017; Schellenberg et al., 2007).

One explanation for why background music impairs CRAT performance in the same way as it impairs serial recall (e.g., Salamé & Baddeley, 1989) relates to the processes of verbal working memory and their importance for insight problem solving. Indeed, Ball and Stevens (2009) identified a strong verbal component to the solving of insight-based CRATs, in that implementing a "think aloud" process during problem solving reliably enhanced solution rates. If verbal working memory is important for CRAT performance, and any non-steady-state background sound (such as music) is disruptive to the creative and analytic processes necessary to solve CRATs, a decrement to CRAT performance would be expected. This, therefore, suggests that it is not the type of semantic content within the to-be-ignored background per se that is disruptive to insight problem solving, but rather the presence of changing-state sound and its impact on verbal working memory processes (such as rehearsal) underpinning the solving of CRATs. It might be, for example, that participants rehearse various target solutions, before obtaining an appropriate solution word. However, since semantic associative processes are likely to be involved in CRAT solving (Smith et al., 2013), it would be reasonable to predict that meaningful background speech in a participant's mother tongue, as compared with meaningless background speech (speech in a language foreign to the participant), could produce additional disruption to CRAT solving, superimposed on the changing-state effect, as a consequence of a semantic interference-by-process. The general notion is that disruption over and above the changing-state effect can occur when there is a conflict between semantic processing of the sound and semantic processing in the focal task (Jones, Marsh, & Hughes, 2012; Marsh et al., 2009). For example, in the context of mental arithmetic, Perham, Marsh, Clarkson, Lawrence, and Sörqvist (2014) argued that the additional disruption due to an ascending sequence of number distracters as compared with a random sequence was produced due to an additional priming process that was superimposed upon the interference-by-process that underpinned the changing-state effect.

One challenge that stems from the finding that background music impairs CRAT solving relates to determining the generalisability of the observed effect beyond CRATs alone. CRATs are but one example of a verbal insight problem solving task, albeit a popular and widely used example that is believed to provide an effective test of creativity (e.g., Bowden & Jung-Beeman, 1998; Mednick, 1962). However, many other types of verbal and non-verbal insight problem solving tasks exist (e.g., see Gilhooly, Fioratou, & Henretty, 2010). An important consideration is to what extent do the explanations presented here in terms of interference-by-process and working memory generalise to further insight problem solving tasks in both the verbal and visual domain?

In relation to this latter question, preliminary evidence from Ball, Marsh, Litchfield, Cook, and Booth (2015) suggests that non-verbal (i.e., visuo-spatial) insight problem solving may in fact be *facilitated*, rather than hampered by background sound. Participants solved “classic” non-verbal insight problems (such as the pigs-in-pens problem; see Schooler, Ohlsson, & Brooks, 1993) more accurately, and faster, when background sound was presented (in a form that involved repeated, canonical counting of the digit sequence 1 to 7) compared to a quiet background. According to Ball et al. (2015), these results can be interpreted in terms of background sound impairing inner speech, thereby permitting the operation of non-reportable, “special processes” (e.g., problem restructuring) that are critical for enabling successful solutions to emerge in classic insight-based problem solving tasks involving predominantly visuo-spatial components. It is our conjecture that background music would have a similar positive effect to this, yet it would be the *changing-state properties* of the to-be-ignored background music, rather than the music per se, that would underpin an apparent enhancement in creativity. In sum, boundary effects are likely to be evident in terms of: (1) the relationship between background music and creativity; (2) the negative consequences of background music on verbal insight problem solving tasks as demonstrated here with CRATs; and (3) the positive consequences on visuo-spatial insight problem solving tasks (e.g., Ball et al., 2015).

One further point of consideration is that Mehta et al. (2012) found facilitatory effects of background sound for a number of different tasks that are thought to tap creativity. Although we found no evidence for the processing disfluency account (Methta et al., 2012) in the context of our current study, it is possible that other tasks, or their component processes, are more sensitive to the potential engendering of processing disfluency – and thus enhancement of cognition – via the presence of background noise. Arriving at a CRAT solution may involve multiple processes that include a delicate balance between top-down processes (e.g., rehearsal and executive control, such as inhibition in the case of excluding incorrect response candidates) and bottom-up processes, such as spreading activation in associative, semantic networks that provide the candidate responses (cf. Benedek, Kenett, Umdasch, Anaki, Faust, & Neubauer, 2016). Since it is likely that the balance between these top-down and bottom-up processes may differ substantially between different insight problems, different susceptibility to distraction by background sound is likely across different types of problem solving tasks solved via a process of insight. In this way it is possible that any advantage to creative problem solving promoted by processing disfluency due to the presence of background noise may be dependent upon the particular strategy and processes used to solve a CRAT. It is reasonable to suggest that such strategies and processes could differ on a problem-by-problem basis.

While we argue that our results can be explained in terms of the interference-by-process account (Jones & Tremblay, 2000), it is important to note that there are two competing explanations for a deficit in CRAT performance with background noise, which both stem from well-established theories of creativity. One such explanation derives from the "broad attentional scope" (BAS) view of creativity (Zabelina, O'Leary, Pornpattananangkul, Nusslock, & Beaman, 2015), whilst the other derives from the "focused attentional scope" (FAS) perspective (Gilhooly, Fioratou, Anthony, & Wynn, 2007). The BAS view supposes that background sound may reduce the overall amount of attention that one can apply to the problem solving task, resulting in a *diffuse attentional state* that can facilitate insight problem solving (Jarosz, Colflesh, & Wiley, 2012). In comparison, the FAS view proposes that if background sound reduces overall attention, then insight problem solving will be impaired. This FAS view is consistent with theories of distraction that assume the presence of background sound captures attention away from the focal task (Cowan, 1995), or reduces the overall amount of attention applied to the focal task (Neath, 2000), thus impairing observed performance. Across three experiments we demonstrate a deficit to CRAT performance with distraction via to-be-ignored background music. Such a deficit to creativity is consistent with the importance of focused attentional scope (FAS) in verbal insight problem solving such as with CRATs.

One might assume that the findings reported here suggest that a disruption to focused attentional scope is generated by to-be-ignored background music, and that this is disruptive to the analytic and associative processes necessary to solve insight problems. However, there are two key problems with this assumption. First, there was nothing inherent within the presented background sound that was likely to capture attention, and thus disrupt the attention directed toward the task at hand. Moreover, the notion that the overall amount of attention that could be applied to the focal task is reduced in the presence of background sound is inconsistent with the literature showing that only tasks that require verbal rehearsal are susceptible to distraction (Beaman & Jones, 1997; Jones & Macken, 1993). Second, if background sound impaired a focused attentional state necessary to solve insight problem solving tasks, then one would expect disruption, not facilitation, on tasks that require visuo-spatial insight problem solving, similar to that found with verbal insight problem solving. However, evidence from Ball et al. (2015) and the findings presented here, suggest a dissociation in the facilitatory and disruptive effects of to-be-ignored background sound on verbal and visuo-spatial insight problems, respectively.

To conclude, the findings here challenge the popular view that music enhances creativity, and instead demonstrate that music, regardless of the presence of semantic content (no lyrics, familiar lyrics or unfamiliar lyrics), consistently disrupts creative performance in insight problem solving as measured by CRATs.

Author Note

The first and second authors contributed equally to the work undertaken here. The research reported in this article was supported by funding from the British Academy and Leverhulme Trust (Grant No: SG162930).

References

Ackerman, R., & Thompson, V. A. (2017a). Meta-reasoning: Monitoring and control of thinking and reasoning. *Trends in Cognitive Sciences, 21*, 607–617. doi: 10.1016/j.tics.2017.05.004

Ackerman, R., & Thompson, V. A (2017b). Meta-Reasoning: Shedding metacognitive light on reasoning research. In L. J. Ball, & V. A. Thompson (Eds.), *The Routledge international handbook of thinking and reasoning* (pp. 1-15). Abingdon, Oxford: Routledge.

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology*: *General*, *136*, 569-576. doi: 10.1037/0096-3445.136.4.569

Ansburg, P. I. (2000). Individual differences in problem solving via insight. *Current Psychology, 19*, 143-146. doi: 10.1007/s12144-000-1011-y

Ansburg, P. I., & Hill, K. (2003). Creative and analytic thinkers differ in their use of attentional resources. *Personality & Individual Differences, 34,* 1141-1152. doi: 10.1016/S0191-8869(02)00104-6.

Arden, R., Chavez, R. S., Grazioplene, R., & Jung, R. E. (2010). Neuroimaging creativity: A psychometric review. *Behavioural Brain Research, 214,* 143-156. doi: 10.1016/j.bbr.2010.05.015

Ball, L. J., Marsh, J. E. M., Litchfield, D., Cook, R., & Booth, N. (2015). When distraction helps: Evidence that concurrent articulation and irrelevant speech can facilitate insight problem solving. *Thinking & Reasoning, 21,* 76-96. doi: 10.1080/13546783.2014.934399

Ball, L. J., & Stevens, A. (2009). Evidence for a verbally-based analytic component to insight problem solving. In N. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1060-1065). Austin, TX: Cognitive Science Society.

Beaman, C. P. (2005). Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments. *Cognitive Psychology, 19,* 1041-1064. doi*:* 1041-1064. 10.1002/acp.1134

Beaman, C. P., & Jones, D. M. (1997). The role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 23,* 459-471. doi: 10.1037/0278-7393.23.2.459

Benedek, M., Kenett, Y. N., Umdasch, K., Anaki, D., Faust, M., & Neubauer, A. C. (2016). How semantic memory structure and intelligence contribute to creative thought: A network science approach. *Thinking & Reasoning*, *23*, 158-183. doi: 10.1080/13546783.2016.1278034

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417-444. doi: 10.1146/annurev-psych-113011-143823

Bowden, E. M., & Beeman, M. J. (1998). Getting the right idea: Semantic activation in the right hemisphere may help solve insight problems. *Psychological Science, 6,* 435-440. doi: 10.1111/1467-9280.00082

Bowden, E. M., & Jung-Beeman, M. (1998). Normative data for 144 compound remote associate problems. *Behavior Research Methods, Instruments, & Computers, 35*, 634-639. doi: 10.3758/BF03195543

Bowden, E. M., & Jung-Beeman, M. (2003). Aha! Insight experience correlates with solution activation in the right hemisphere. *Psychonomic Bulletin & Review, 10,* 730-737. doi: 10.3758/BF03196539.

Bowden, E. M., & Jung-Beeman, M. (2007). Methods for investigating the neural components of insight. *Methods, 42,* 87-99. doi:10.1016/j.ymeth.2006.11.007

Bowden, E. M., Jung-Beeman, M., Fleck, J., & Kounios, J. (2005). New approaches to demystifying insight. *Trends in Cognitive Sciences, 9,* 322-328. doi: 10.1016/j.tics.2005.05.012

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT Press.

Cai, D. J., Mednick, S. A., Harrison, E. M., Kanady, J. C., & Mednick, S. C. (2009). REM, not incubation, improves creativity by priming associative networks. *Proceedings of the National Academy of Sciences of the Unites States of America, 106,* 10130-10134. doi: 10.1073/pnas.0900271106

Chein, J. M., & Weisberg, R. W. (2013). Working memory and insight in verbal problems: Analysis of compound remote associates. *Memory & Cognition, 38,* 473-481. doi: 10.3758/s13421-013-0343-4

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning & Verbal Behavior, 15,* 17–32. doi: 10.1016/S0022-5371(76)90003-7

Cowan, N. (1995). *Working memory capacity*. Hove, East Sussex, UK: Psychology Press.

Dienes, Z. (2008). *Understanding psychology as a science: An introduction to scientific and statistical inference.* London, UK: Palgrave Macmillan International Higher Education.

Dienes, Z. (2011). Bayesian versus orthodox statistics: which side are you on? *Perspectives on Psychological Science, 6*, 274-290. doi: 10.1177/1745691611406920

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, Article 781, 1-17. doi:10.3389/fpsyg.2014.00781

Dienes, Z., & McLatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin & Review, 25*, 207-218. doi: 10.3758/s13423-017-1266-z

Divin, W., Coyle, K., & James, D. T. (2001). The effects of irrelevant speech and articulatory suppression on the serial recall of silently presented lipread digits. *British Journal of Psychology, 92*, 593-616. doi: 10.1348/000712601162365

Elliott, E. M. (2002). The irrelevant speech effect in children: Theoretical implications of developmental change. *Memory & Cognition, 30,* 478-487. doi: 10.3758/BF03194948

Evans, J. St. B. T., & Stanovich, K. E. (2013a). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223-241. doi: 10.1177/1745691612460685

Evans, J. St. B. T., & Stanovich, K. E. (2013b). Theory and metatheory in the study of dual processing: Reply to comments. *Perspectives on Psychological Science, 8*, 263-271. doi: 10.1177/1745691613483774

Fodor, E. M. (1999). Subclinical inclination toward manic-depression and creative performance on the Remote Associates Test. *Personality & Individual Differences, 27*, 1273-1283. doi: 10.1016/S0191-8869(99)00076-8

Gilhooly, K. J., Fioratou, E., Anthony, S. H., Wynn, V. (2007). Divergent thinking: strategies and executive involvement in generating novel uses for familiar objects. *British Journal of Psychology, 98,* 611-625. doi: 10.1111/j.2044-8295.2007.tb00467.x

Gilhooly, K. J, Fioratou E., & Henretty, N. (2010). Verbalization and problem solving: insight and spatial factors. *British Journal of Psychology, 101,* 81-93. doi: 10.1348/000712609X422656

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10, 535–549.

Hillier, A., Alexander, J. K., & Beversdorf, D. Q. (2006). The effect of auditory stressors on cognitive flexibility. *Neurocase, 12,* 228-231. doi: 10.1080/13554790600878887

Hughes, R. W., & Jones, D. M. (2003). A negative order-repetition priming effect: Inhibition of order in unattended sequences? *Journal of Experimental Psychology: Human Perception & Performance, 29,* 199-218. doi: 10.1037/0096-1523.29.1.199

Jarosz, A. F., Colflesh, G. J. H., & Wiley, J. (2012). Uncorking the muse: Alcohol intoxication facilitates creative problem solving. *Consciousness & Cognition, 21,* 487-493. doi: 10.1016/j.concog.2012.01.002

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 19,* 369-381. doi: 10.1037/0278-7393.19.2.369

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *Quarterly Journal of Experimental Psychology, 44A,* 645-669. doi: 10.1080/14640749208401304

Jones, D. M., Marsh, J. E., & Hughes, R. W. (2012). Retrieval from memory: Vulnerable or inviolable? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 38,* 905-922. doi: 10.1037/a0026781

Jones, D., Miles, C. & Page, J. (1990). Disruption of proofreading by irrelevant speech: effects of attention, arousal or memory? *Applied Cognitive Psychology, 4,* 89-108. doi: 10.1002/acp.2350040203

Jones, D. M., & Tremblay, S. (2000). Interference by process or context? A reply to Neath (2000). *Psychonomic Bulletin & Review, 12,* 550-558. doi: 10.3758/BF03214370

Jung-Beeman, M., Bowden, E. M., Haberman, J., Frymiare, J. L., Arambel-Liu, S., Greenblatt, R., Reber, P. J., & Kounios, J. (2004). Neural activity observed in people solving verbal problems with insight. *PLoS Biology*, *E97,* 500-510. doi: 10.1371/journal.pbio.0020097

Kasof, J. (1997). Creativity and breadth of attention. *Creativity Research Journal, 10,* 303-315. doi: 10.1207/s15326934crj1004_2

Klatte, M., Kilcher, H., & Hellbrück, J. (1995). Wirkungen der zeitlichen Struktur von hintergrundschall auf das arbeitsgedächtnis und ihre theoretischen und praktischen implikationen. *Zeitschrift für Experimentelle Psychologie, 42,* 517-544.

Klatte, M., Lachmann, T., Schlittmeier, S., & Hellbrück, J. (2010). The irrelevant sound effect in short-term memory: Is there developmental change? *European Journal of Cognitive Psychology, 22,* 1168-1191. doi: 10.1080/09541440903378250

Marsh, J. E., Crawford, J. C., Pilgrim, L. K., Sörqvist, P., & Hughes, R. W. (2017). Trouble articulating the right words: Evidence for a response-exclusion account of distraction during semantic fluency. *Scandinavian Journal of psychology, 58*, 367-372. doi: 10.1111/sjop.12386

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory & Language, 58,* 682-700. doi: 10.1016/j.jml.2007.05.002

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition, 110,* 23-38. doi: 10.1016/j.cognition.2008.08.003

Marsh, J. E., Nakabayashi, K., Frowd, C. D., Skelton, F. C., Vachon, F., & Ball, L. J. (2018). In the face of distraction: Changing-state background sound impairs person recognition. Manuscript submitted for publication.

Marsh, J. E., Threadgold, E., Barker, M., & Ball, L. J. (2017). *Can susceptibility to distraction inform the component processes underpinning verbally-based insight problem solving?* Paper presented at 20th Conference of the European Society for Cognitive Psychology, Potsdam, Germany, 3-6th September.

Martindale, C., & Greenough, J. (1973). The differential effect of increased arousal on creative and intellectual performance. *Journal of Genetic Psychology, 123,* 329-335. doi: 10.1080/00221325.1973.10532692

McNair, D. M. (1971). *Manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service.

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review, 69,* 220-232. doi: 10.1037/h0048850

Mednick, S. A., & Mednick, M. T. (1967). *Examiner's manual, Remote Associates Test: College and adult forms 1 and 2.* Boston: Houghton Mifflin.

Mehta, R., Zhu, R. J., & Cheema, A. (2012). Is noise always bad? Exploring the effects of ambient noise on creative cognition. *Journal of Consumer Research, 39,* 784-799. doi: 10.1086/665048

Mikulincer, M., & Sheffi, E. (2000). Adult attachment style and cognitive reactions to positive affect: A test of mental categorization and creative problem solving. *Motivation and Emotion, 24*, 149-174. doi: 10.1023/A:1005606611412

Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review*, 7, 403–423. doi:10.3758/BF03214356

Nittono, H. (1997). Background instrumental music and serial recall. *Perceptual & Motor Skills, 84,* 1307-1313. doi: 10.2466/pms.1997.84.3c.1307

Ohlsson, S. (2011). *Deep learning: How the mind overrides experience*. New York, NY: Cambridge University Press.

Perham, N., & Currie, H. (2014). Does listening to music improve reading comprehension performance? *Applied Cognitive Psychology, 28,* 279-284. doi: 10.1002/acp.2994

Perham, N., Marsh, J. E., Clarkson, M., Lawrence, R., & Sörqvist, P. (2016). Distraction of mental arithmetic by background speech: Further evidence for the habitual-response priming view of auditory distraction. *Experimental Psychology, 63*, 141-149. doi: 10.1027/1618-3169/a000314

Perham, N., & Vizard, J. (2010). Can preference for background music mediate the irrelevant sound effect? *Applied Cognitive Psychology*, *25*, 625-631. doi: 10.1002/acp.1731

Perkins, D. (2000). *The eureka effect: The art and logic of breakthrough thinking*. New York, NY: W. W. Norton.

Ritter, S. M., & Ferguson, S. (2017). Happy creativity: Listening to happy music facilitates divergent thinking. *PLoS ONE, 12*, e0182210, 1-14. doi: 10.1371/journal.pone.0182210

Rowe, G., Hirsh, J. B., & Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 383-388. doi: 10.1073/pnas.0605198104

Rummer, R., Schweppe, J., & Schwede, A. (2016). Fortune is fickle: Null-effects of disfluency on learning outcomes. *Metacognition & Learning*, *11*, 57-70. doi: 10.1007/s11409-015-9151-5

Salamé, P., & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning & Verbal Behavior, 21,* 150-164. doi: 10.1016/S0022-5371(82)90521-7

Salamé, P., & Baddeley, A. D. (1989). Effect of background music on phonological short term memory. *Quarterly Journal of Experimental Psychology, 41,* 107-122. doi: 10.1080/14640748908402355

Schellenberg, E. G. (2005). Music and cognitive abilities. *Current Directions in Psychological Science, 14,* 317-320. doi: 10.1111/j.0963-7214.2005.00389.x

Schellenberg, E. G., Nakata, T., Hunter, P. G., & Tamoto, S. (2007). Exposure to music and cognitive performance: tests of children and adults. *Psychology of Music, 35, 5*-19. doi: 10.1177/0305735607068885

Schlittmeier, S., Hellbrück, J., & Klatte, M. (2008). Does irrelevant music cause an irrelevant sound effect for auditory items? *European Journal of Cognitive Psychology, 20,* 252-271. doi: 10.1080/09541440701427838

Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97-133). Cambridge, MA: MIT Press.

Schooler, J. W., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. *Journal of Experimental Psychology: General, 122*, 166-183. doi: 10.1037/0096-3445.122.2.166

Shen, W., Yuan, Y., Liu, C., & Luo, J. (2017). The roles of the temporal lobe in creative insight: An integrated review. *Thinking & Reasoning*, *23*, 321-375. doi: 10.1080/13546783.2017.1308885

Smith, K. A., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the remote associates test. *Cognition*, *128*, 64-75. doi: 10.1016/j.cognition.2013.03.001

Storm, B. C., & Angello, G. (2010). Overcoming fixation: Creative problem solving and retrieval-induced forgetting. *Psychological Science, 21*, 1263-1265. doi: 10.1177/0956797610379864

Storm, B. C., Angello, G., & Bjork, E. L. (2011). Thinking can cause forgetting: Memory dynamics in creative problem solving. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 37,* 1287-1293. doi: 10.1037/a0023921

Thompson, V. A. (2010). Towards a metacognitive dual process theory of conditional reasoning. In M. Oaksford, & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 335-354). Oxford: Oxford University Press.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science*, *12*, 248-251. doi: 10.1111/1467-9280.00345

Toplyn, G., & Maguire, W. (1991). The differential effect of noise on creative task performance. *Creativity Research Journal, 4,* 337-347. doi: 10.1080/10400419109534410

Van Casteren, M., & Davis, M. (2007). Match: A program to assist in matching the conditions of factorial experiments. *Behavior Research Methods, 39*, 973-978. doi: 10.3758/BF03192992

Webb, M. E., Little, D. R., Cropper, S. J., & Ruze, K. (2017). The contributions of convergent thinking, divergent thinking, and schizotypy to solving insight and non-insight problems. *Thinking & Reasoning*, *23*, 235-258. doi: 10.1080/13546783.2017.1295105

Weisberg, R. W. (1995). Prolegomena to theories of insight in problem solving: A taxonomy of problems. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 157-196). Cambridge, MA: MIT Press.

Weisberg, R. W. (2015). Toward an integrated theory of insight problem solving. *Thinking & Reasoning, 21,* 5-39. doi: 10.1080/13546783.2014.886625

White, H. A., & Shah, P. (2011). Creative style and achievement in adults with attention-deficit/hyperactivity disorder. *Personality & Individual Differences, 40,* 1121-1131. doi: 10.1016/j.paid.2010.12.015

Wiley, J., & Jarosz, A. F. (2012). Working memory capacity, attentional focus and problem solving. *Current Directions in Psychological Science, 21,* 258-262. doi: 10.1177/0963721412447622

Zabelina, D. L., O'Leary, D., Pornpattananangkul, N., Nusslock, R., & Beaman, M. (2015). Creativity and sensory gating indexed by the P50: Selective versus leaky sensory gating in divergent thinkers and creative achievers. *Neuropsychologia, 69,* 77-84. doi: 10.1016/j.neuropsychologia.2015.01.034

Table 1. Bayes factors testing the effect of Music versus Quiet conditions from Experiments 1 to 3

Effect	**Study**	**$B_{H1/H0}$: Music increases Creativity**	**$B_{H2/H0}$: Music decreases Creativity**	**$B_{H2/H1}$**
CRAT Performance	Experiment 1	0.04	21.89	547.25
	Experiment 2	0.08	11.58	144.75
	Experiment 3	0.04	3.31	82.75

Note: *H1* = Music > Quiet, *H2* = Quiet > Music, *H0* = Quiet = Music. Alternative hypotheses specified using a half-normal distribution with a mode of zero and a standard deviation of .17 (Mehta et al., 2012, Experiment 1).

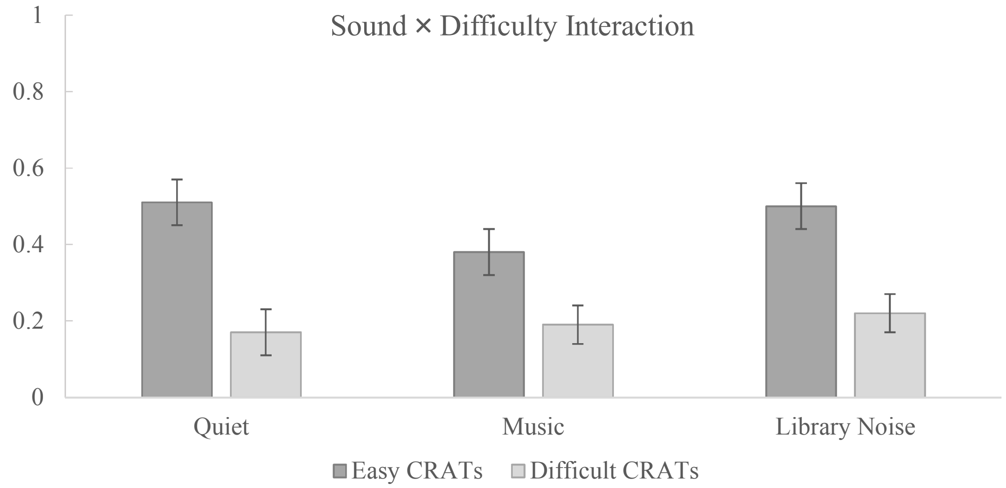

Figure 1. The Sound × Difficulty interaction with standard error bars.