Running head: MUSICAL KEY AND CONTEXT-DEPENDENT MEMORY

Music Tonality and Context-Dependent Recall:

The Influence of Key Change and Mood Mediation

Katharine M. L. Mead and Linden J. Ball*

Lancaster University, UK

*Corresponding author:

Department of Psychology, Lancaster University, Lancaster, LA1 4YF, UK.

Email: L.Ball@lancaster.ac.uk

Tel: +44 (0)1524 593470

Fax: +44 (0)1524 593744

Abstract

Music in a minor key is often claimed to sound sad, whereas music in a major key is typically viewed as sounding cheerful. Such claims suggest that maintaining or switching the tonality of a musical selection between information encoding and retrieval should promote robust “mood-mediated” context-dependent memory (CDM) effects. The reported experiment examined this hypothesis using versions of a Chopin waltz where the key was either reinstated or switched at retrieval, so producing minor--minor, major--major, minor--major and major--minor conditions. Better word recall arose in reinstated-key conditions (particularly for the minor--minor group) than in switched-key conditions, supporting the existence of tonality-based CDM effects. The tonalities also induced different mood states. The minor key induced a more negative mood than the major key, and participants in switched-key conditions demonstrated switched moods between learning and recall. Despite the association between music tonality and mood, a path analysis failed to reveal a reliable mood-mediation effect. We discuss why mood-mediated CDM may have failed to emerge in this study, whilst also acknowledging that an alternative “mental-context” account can explain our results (i.e., the mental representation of music tonality may act as a contextual cue that elicits information retrieval).

Music Tonality and Context-Dependent Recall: The Influence of Key Change and Mood Mediation

Context-dependent memory (CDM) refers to the phenomenon whereby a change in context between information learning and retrieval causes some of the originally encoded information to be forgotten. Much laboratory-based experimental research has investigated which aspects of a prevailing context can provide successful retrieval cues when reinstated -- as opposed to being switched -- between information encoding and subsequent testing. Evidence has shown that a wide range of contexts, both external (i.e., environmental) and internal (i.e., mental), can produce CDM effects on recall with a good degree of reliability. Such contexts include olfactory stimuli (e.g., Cann & Ross, 1989; Parker, Ngu, & Cassaday, 2001; Schab, 1990), general physical location (e.g., Godden & Baddeley, 1975; Isarida & Isarida, 2004; Smith, Glenberg, & Bjork, 1978); time of day (e.g., Holloway, 1978); drug or alcohol states (e.g., Eich, 1980); mood states (e.g., Eich, Macauley, & Ryan, 1994; Eich & Metcalfe, 1989) and ambient music (e.g., Balch, Bowman, & Mohler, 1992; Smith, 1985). It is the claim that background music can act as a potent environmental CDM cue that forms the focus of the present study, with a specific issue being the role of music tonality in promoting CDM effects on recall. Before we overview the detailed aims of our study we first review existing research on musical CDM effects.

In his pioneering study of music-based CDM, Smith (1985) presented participants with words in one of three different contexts: a jazz selection, a classical selection, or with no music present. Smith observed a CDM effect when reinstating a music context (but not a quiet context) for word recall 48 hr after encoding. Although switched-music contexts revealed worse performance than music-reinstated contexts, music switching did not, in fact, hamper recall compared to either the quietness-

reinstated condition or the music then quiet conditions. These latter findings indicated that music mismatches did not cause *distraction* at recall.

Balch, Bowman, and Mohler (1992, Experiment 1) found no music-dependent memory effects for delayed recall (after a 48 hr interval), but for immediate recall it was observed that reinstated-music conditions produced reliably better performance than switched-music conditions, indicating a CDM effect. Music reinstatement, however, showed no memory facilitation relative to a no-cue condition, although the linear trend in recall scores across the three ordered means (i.e., switched music < no cue < reinstated music) was reliable, suggesting switched conditions inhibited recall whilst reinstated conditions facilitated recall. Balch et al. (1992, Experiment 2) also observed that when the musical selections were switched in either tempo (fast vs. slow) or form (classical vs. jazz) it was only those pieces having an altered tempo that produced reliably lower immediate recall compared with reinstated pieces.

Balch and Lewis (1996) noted a problem with the methodology that Balch et al. (1992) had used to investigate tempo-dependent memory, which was that *different* musical selections had been employed in the switched-tempo conditions. This meant that switched conditions were potentially changed on multiple dimensions, including instrument timbres, melodic phrases, and harmonic sequences. Altering contexts along multiple dimensions may promote stronger CDM effects (e.g., Dalton, 1993; Eich, Macauley, & Ryan, 1994; Tulving, 1983) and also confuses the interpretation of what manipulations may be *sufficient* causes of CDM. Balch and Lewis (1996, Experiments 1 and 2) therefore tested music-dependent memory by including conditions that reinstated or changed the tempo (slow vs. fast) or reinstated or changed the timbre (piano vs. brass) of the *same* musical piece. Evidence supported the existence of tempo-dependent memory (more words recalled in reinstated-tempo

vs. switched-tempo conditions) but not timbre-dependent memory (no reduction in recall for switched-timbre vs. same-timbre conditions).

In follow-up experiments, Balch and Lewis (1996, Experiments 3 and 4) examined the specific question of why changes in music tempo induce CDM effects. They hypothesised that alterations in tempo might induce changes in mood, and that it is such mood changes that underpin music-dependent memory. To assess this “mood-mediation hypothesis” (cf. Eich, 1995a), Balch and Lewis (Experiment 3) ran an initial study to determine the effects of three music variables – tempo (slow vs. fast), timbre (piano vs. brass), and musical selection (classical vs. jazz) – on people’s mood states, as measured by eliciting mood ratings from participants on the dimensions of *arousal* and *pleasantness*. The only music variable that was significantly related to mood ratings was tempo. Moreover, tempo was observed to influence the arousal dimension of mood (with the fast tempo being more arousing), but not the pleasantness dimension.

Having established a relation between tempo and mood, Balch and Lewis (Experiment 4) went on to examine how this relation influenced memory. Participants learnt and recalled words in the same moods or in switched moods. The mood states in the learning phase were induced by different music tempos (slow vs. fast). In the retrieval phase mood states were induced by verbal instructions and a spoken mood-inducing scenario. Recall was found to be higher in mood contexts at retrieval that were consistent with a previous tempo versus mood contexts inconsistent with a previous tempo. These results support a view of music-dependent memory as being mediated by mood changes, as predicted by Eich’s (e.g., 1995a) mood-mediation hypothesis. An alternative to the mood-mediation account of CDM is the “mental-context hypothesis” (Smith, 1995; Tulving, 1983), which proposes that mental context

can be broadly construed to include many factors (e.g., mood, place, mental set) and that any of these can cause context-dependent retrieval when reinstated (or, alternatively, context-dependent forgetting when switched). Balch and Lewis (1996), however, argue that whilst their tempo manipulation is a sufficient cause for music-dependent memory, they were unable to induce CDM effects with other contextual changes associated with music. As such, their results provide little evidence that music-dependent memory can have multiple causes, as predicted by Smith's (1995) mental-context hypothesis.

The starting point for our own research is the assumption that the mood-mediation hypothesis is generally capable of explaining a wealth of data (including inconsistent findings across studies) relating to CDM (cf. Balch & Lewis, 1996; Eich, 1995b). For example, notable failures in finding CDM effects (e.g., Fenandez & Glenberg, 1985) may well be associated with failures of the different contexts to promote sufficiently distinctive mood or arousal states (Eich, 1995a). We also concur with Balch and Lewis (1996), who note that reliability has not been an issue in studies of music-dependent memory, except in those cases where tempo was purposely kept the same across switched-context conditions. This observation, again, attests to the possible role of mood mediation in music-dependent memory.

Given the apparent importance of mood-induced states to the emergence of CDM effects, it is useful to reflect on the critical factors that promote or hinder the demonstration of mood-dependent memory (e.g., Bower, 2003; Eich, 1995a; Eich, & Forgas, 2003). Mood-dependent memory effects seem to emerge most reliably when: (1) to-be-remembered items are generated by participants or actively processed by them in some way (e.g., Balch, Myers, & Papotto, 1999; Eich & Metcalfe, 1989; Eich et al., 1994); (2) memory is tested via free recall (e.g., Beck & McBee, 1995; Bower,

1981; Eich & Metcalfe, 1989), with recollection arising immediately after learning rather than at a delay (e.g., Balch et al., 1992); and (3) mood induction is done with great care to ensure that the mood manipulation is strong, that it lasts throughout the learning or testing phase, and that where moods are switched they are substantially different (e.g., Bower, 1992; Eich et al., 1994; Isen, 1984).

In more recent research on music-induced mood states and CDM, Balch, Myers, and Papotto (1999) further investigated the links between the arousal and pleasantness dimensions of mood and CDM effects. Balch et al.'s experiments pre-classified selections of "mood music" in terms of pleasantness and arousal, and then systematically manipulated such selections in a CDM paradigm. Switching music on only the pleasantness dimension decreased memory whether or not the *intended* mood of the pieces had been explicitly described to participants by the experimenter. However, just changing the arousal dimension of music decreased memory only when intended moods had been explicitly defined. These results indicate the robustness of pleasantness-dependent memory, but raise some concerns about the reliability of arousal-dependent memory. Still, it should be borne in mind that Balch and Lewis' earlier research (e.g., Balch & Lewis, 1996, Experiment 4) *did* find evidence for tempo-dependent CDM effects being mediated by arousal changes even without participants being directly instructed on what mood to try to attain during music presentations at learning. In sum, mood arousal in music-based CDM seems a real phenomenon, but perhaps a somewhat inconsistent one.

Returning to the present study, we note that music-dependent memory is still a relatively new domain of research, where little is known about what dimensions of music can serve as CDM cues and whether mood-mediation is always a critical element of music-based CDM effects. Whilst Balch and Lewis (1996) found that

tempo-dependent memory was mediated by mood, they did not explicitly consider if the tonality of music (i.e., the use of major or minor keys) is a further dimension of music that may induce robust CDM effects (perhaps also through mood mediation). Indeed, studies of music-dependent memory do not appear to have directly controlled for that tonality of the pieces selected for study. This seems like a curious omission given that tonality is well-known for its use by composers as a central device to modulate mood changes in listeners. The minor key is often described as invoking sadness, whilst the major key is claimed to instil feelings of happiness. Recent studies have also provided empirical evidence that the affective tone of musical chords can influence the speed of evaluation of *affectively congruent* chord--target word pairings relative to *affectively incongruent* pairings (Sollberger, Reber, & Eckstein, 2003), and that musical tone (major vs. minor) may impact on mood valence (i.e., inducing positive vs. negative mood states respectively), but not arousal (Husain, Thompson, & Schellenberg, 2002).

The present experiment used a single selection of music (to control for timbre, form, tempo, melody and the like) and examined whether reinstating or switching its musical key between item learning and recall could produce a CDM effect that was systematically related to mood changes. The main research questions were: can the key of a musical piece act as a retrieval cue for the recall of words when reinstated, relative to when it is switched, and can such a CDM effect, if present, be shown to be mediated by changes in people's mood states? A tonality-based CDM effect on recall would manifest itself as a cross-over interaction between learning environment (minor vs. major) and recall environment (minor vs. major). Furthermore, the mood-mediation hypothesis would gain support if tonality changes were seen to be systematically linked to mood changes, and, more critically, if a regression-based

"mediation analysis" of the data (e.g., Baron & Kenny, 1986) could establish a path between the context manipulation and recall scores that was mediated by mood status. In our study we were mindful to employ a methodology that was most likely to reveal mood-mediated CDM, that is, we included active processing of the to-be-remembered material and an immediate free-recall paradigm.

Method

Participants

Participants were 48 Lancaster University undergraduates (10 males, 38 females; age range: 19-25 years) who volunteered to take part in the study for a small financial incentive. All had normal or corrected-to-normal hearing and vision. Participants were randomly assigned to experimental conditions.

Apparatus and Stimuli

Word items and incidental learning task. The to-be-remembered words were 24 common, two- and three-syllable nouns with a concreteness rating of five or above, selected from Spreen and Schultz's (1966) norms (e.g., basket, library, cousin etc.). The learning phase of the study involved an incidental (rather than intentional) word-learning method (i.e., each presented word had to be rated for pleasantness on a six-point scale and the subsequent recall task came as a surprise test of incidental memory). Participants were given a pencil and a booklet for rating the words on the 6-point scale (1 for *very unpleasant*, 2 for *moderately unpleasant*, 3 for *slightly unpleasant*, 4 for *slightly pleasant*, 5 for *moderately pleasant*, and 6 for *very pleasant*). The requirement for participants to rate words for pleasantness enabled semantic thought processes to occur, thereby encouraging deep encoding of the items (e.g., Hyde & Jenkins, 1973). Note that the selection of the 24 words aimed to achieve a roughly equal distribution of items that were believed to have positive,

negative and neutral emotional connotations, so as to encourage a degree of variability in participants pleasantness ratings as well as some genuine semantic processing of the words (see Appendix A for the full list of word items used).

Music selections. Words were presented for rating during the learning phase along with background music either in a major or minor key. The music used in the study was a Frederic Chopin piano score (“Waltz in A Minor, Op. 34 No.2”) originally written in a minor key and lasting approximately 105 s. This score was transposed into a major key by an expert musician and pianist, without changing any other aspects such as tempo or phrasing. Tape recordings were made of the minor and major versions of the musical piece being played by the expert pianist. These recordings were obtained using a high-quality audio-cassette recorder with an external, multi-directional microphone positioned 1 metre in front of the piano. The input volume of the tape recorder was set to its default level to equate the recording volume for the two versions of the piano piece, and recordings were made with Dolby noise-reduction set to Mode C. Five repetitions of each piece (totalling approximately 8 min for each version) were then re-recorded onto a new cassette. A recording was also made of 240 s of birdsong that was to be played between the learning and testing phases. During the experiment the musical pieces and the birdsong were re-played to participants using a high-quality cassette player with external speakers, with the volume setting set at a mid-point such that the sound was neither unpleasantly loud nor overly quiet.

Our study was based on an assumption that the major-key music would be perceived as inducing a more positive mood state than the minor-key music. To examine this issue we obtained mood ratings for each version of the waltz using a 9 x 9 “mood grid” (see Appendix B; cf. Russell, Weiss, & Mendelsohn, 1989) that

depicted the mood dimensions of *arousal* and *pleasantness*. This music pre-test used a sample of 42 undergraduate participants who did not take part in the main experiment. These participants were randomly assigned to listen individually to either the major-key or minor-key version of the waltz for 4 min, after which they were requested to mark a single point on the grid that reflected how the music made them feel in terms of arousal and pleasantness. The music continued to play whilst participants decided on their ratings. The arousal scores revealed a reliable difference between the major- and minor-key pieces, with the major-key music being viewed as more arousing (M = 5.43, SE = 0.38) than the minor-key music (M = 4.19, SE = 0.39), $F(1, 40) = 5.18$, $MSE = 3.11$, $p = .028$, $\eta_p^2 = .12$. Likewise, the pleasantness scores indicated a significant difference between the pieces, with the major-key music being rated as inducing greater feelings of pleasantness (M = 7.24, SE = 0.23) relative to the minor-key music (M = 6.24, SE = 0.30), $F(1, 40) = 7.05$, $MSE = 1.49$, $p = .011$, $\eta_p^2 = .15$. These results support the contention that tonality differences in musical selections can induce predictable differences in mood states, although it is interesting that both of the musical selection were rated toward the “pleasant” end of the pleasantness dimension, whereas the arousal dimension was arguably a better discriminator of predicted mood differences for these pieces (i.e., the minor-key music was rated on the negative side of the arousal midpoint and the major-key piece on the positive side of the midpoint).

This original Chopin waltz in its minor key was considered to be unlikely to sound particularly familiar to our undergraduate participants, few of whom have classical tastes in music. More critical was the need to validate that the transposed major-key version of the waltz did not differ in its perceived familiarity to the minor-key version, or, indeed, seem “odd” in some way. To examine these issues we obtained familiarity and oddness ratings for each version of the waltz using the 9 x 9

"familiarity/oddness grid" depicted in Appendix B. The same participants who rated the two versions of the Waltz for mood-inducing properties were also asked to provide familiarity and oddness ratings for the piece subsequent to generating their mood ratings. Again, the music continued to play whilst the participant decided which square to a mark on the grid. No reliable difference in perceived familiarity was observed for the major-key piece (M = 5.24, SE = 0.40) relative to the minor-key piece (M = 4.90, SE = 0.52), $F(1, 40) = 0.26$, MSE = 4.54, $p = .62$, $\eta_p^2 = .01$. Both versions of the waltz were viewed as being neither particularly familiar nor unfamiliar. In addition, no significant differences emerged in the perceived oddness of the transposed major-key version (M = 3.24, SE = 0.37) relative to the original, minor-key version (M = 3.29, SE = 0.37), $F(1, 40) = 0.01$, MSE = 3.10, $p = .93$, $\eta_p^2 < .001$. Both pieces were viewed as having low oddness. These ratings are encouraging in that they attest to the close similarity of the two versions of the musical piece at the level of their perceived familiarity and oddness.

Mood measures. During the experiment, participants' mood states were assessed at various phases of the experiment using Watson, Clark, and Tellegen's (1988) "Positive and Negative Affect Schedule" (PANAS), which involves a list of words that describe moods (five negatively and five positively valenced). Participants use a 5-point scale to rate how appropriate each word is for describing how they are feeling at that particular moment (1 denotes that they are *not at all feeling like this*, and 5 denotes that they are *indeed feeling like this*).

Design

Participants were randomly allocated to four conditions in a 2 x 2 between-participants design that manipulated the Learning Context (major vs. minor key) and the Retrieval Context (major vs. minor key). Thus, two conditions involved

reinstatement of the key of the musical selection between learning and retrieval and two conditions involved key switches. Dependent variables were the number of words (out of 24) correctly recalled and participants' mood ratings.

Procedure

Participants were tested in small groups of between two and four in the same room for all conditions. They were informed that the study was about language processing to enable use of the incidental word-learning method, whereby each presented item had to be rated for pleasantness. Once the learning-phase music commenced, participants were asked to get into the mood suggested by the music. After 60 s they completed their first PANAS (which took approximately 60 s) and then commenced the rating of the presented words for pleasantness, whilst the music continued. The 24 words each presented in large typeface on separate index cards at a rate of one item every 5 s. Cards were held up by the experimenter and were clearly visible to all seated participants. The inter-stimulus interval was approximately 1 s. To help give participants sufficient exposure to the material, two different random orders of the 24-word list were run consecutively to generate a complete sequence of 48 items (cf. Balch & Lewis, 1996). During the instructions participants had been informed that each word would be repeated somewhere in the sequence and that they should rate each presented word in terms of their impression of its pleasantness at that particular moment. One of two different 48-word sequences was assigned to half of the participants in each learning condition.

After rating the 48 words for pleasantness participants completed a second PANAS with the background music continuing. They then listened to 240 s of birdsong which acted as a "distraction piece" – following a procedure developed by Balch et al. (1992, Experiment 3). This procedure was intended to help avoid

participants in switched-key conditions simply being differentially *distracted* by the altered key of the second piece compared with participants in reinstated-key conditions. Such added distraction could act to disrupt attention and impair recall performance in a way that is essentially uninteresting from a CDM standpoint. However, an intentionally distracting piece of sound (birdsong in the present case) inserted between word presentation and recall, should have helped eliminate such differences in the level of distraction produced by recall contexts. Birdsong was used as it was distractingly different to the classical music used in the learning and recall phases (i.e., it is essentially a collection of varied sounds with no particular order or key).

After the birdsong, participants progressed to the word-retrieval phase, and were again asked to get into the mood portrayed by the music. Mood states were assessed using the PANAS both before free recall (i.e., 60 s in from the recall-context recording being started) and directly after free recall. The background music continued to play throughout the PANAS-rating and free-recall stages of the second part of the study. Two minutes were allowed for the recall session itself. To initiate recall participants were handed a piece of paper and asked to write down, in any order, as many of the originally presented words that they could recall. After the completion of the study all participants were debriefed.

Results and Discussion

An alpha level of .05 was set for all statistical analyses. Where appropriate, effect-size estimates, measured by partial eta-squared (η_p^2), are also presented.

Word-Pleasantness Ratings

One issue that we sought to address before progressing to an analysis of recall scores and mood data was whether the tonality of the music context at learning had

influenced participants' word-pleasantness ratings (e.g., invoking more pleasant ratings in major-key contexts and less pleasant ratings in minor-key contexts). Such an influence would be indicative of mood congruence (i.e., participants' interpreting the pleasantness of words in a manner that matched their current mood state). Any evidence for mood congruence effects in the present study would complicate the interpretation of any contextual influences on recall.

To examine this issue we first pursued an analysis that involved calculating each person's overall mean word-pleasantness rating across all list items. We grouped these mean ratings in terms of whether the participant had been in the major-key condition at learning or the minor-key condition at learning (see the final row of the table presented in Appendix A). No evidence was found for a mood-congruence effect on participants' overall word-pleasantness ratings. Although words were rated slightly more positively on average in the major-key condition ($M = 3.62$, $SE = .10$) relative to the minor-key condition ($M = 3.51$, $SE = .08$) this effect was far from reliable, $F(1, 46) = 0.72$, $MSE = 0.19$, $p = .40$, $\eta_p^2 = .02$.

To examine the mood-congruence issue further, we conducted item-based analyses that assessed, for each word, whether there was a difference in its rating across minor-key versus major-key conditions (see Appendix A for a full summary of means and standard errors). Only 3 words ("apple", "bush", and "flame") out of 24 showed a reliable difference relating to the presence of major-key versus minor-key music during the pleasantness-rating task. These three differences were all reliable in the expected direction of an increased positive rating in the major-key context relative to the minor-key one. Other differences that came close to significance were for the words "professor" ($p = .080$) and "card" ($p = .072$), but these differences were in the opposite-to-predicted direction. Overall, then, it seems unlikely that changes in

perceived pleasantness for words across different music tonalities would have had any serious or systematic influence on word retention or recall performance in the present study.

Recall Scores

Mean recall scores for all conditions are shown in Table 1. It is evident that participants in retrieval conditions where the musical key at learning was reinstated recalled more words than those in switched-key conditions. A 2 x 2 (Learning Context x Retrieval Context) between-participants analysis of variance (ANOVA) revealed neither a main effect of Learning Context, $F(1, 44) = 1.70$, $MSE = 5.95$, $p = .20$, $\eta_p^2 = .04$, nor of Testing Context, $F(1, 44) = 1.70$, $MSE = 5.95$, $p = .20$, $\eta_p^2 = .04$. The *equivalence* of the statistical values obtained for these two main effects was checked in case a calculation error had been made, but these values were, in fact, found to be accurate (i.e., the equivalence is purely coincidental).

Although the ANOVA revealed no main effects, it did indicate a reliable and predicted interaction between the Learning Context and Testing Context factors, $F(1, 44) = 23.56$, $MSE = 5.95$, $p < .001$, $\eta_p^2 = .35$. This interaction, which is indicative of a tonality-based CDM effect, was explored further by a series of Tukey HSD pairwise comparisons. Significantly more words were recalled in the reinstated minor-key condition relative to each of the switched-key conditions (both ps < .001). The difference between the major-key reinstated condition and the two switched-key conditions was marginally reliable (both ps = .072). Finally, and as expected, there was no significant difference between the two reinstated-key conditions ($p = .27$) or between the two switched-key conditions ($p = .99$).

(Table 1 about here)

Mood Scores

Each participant rated their current mood state using the PANAS at four points during the experiment: twice during the learning phase (once directly before incidental word-learning and once immediately after), and twice during the retrieval phase (once before the recall task and once after). A single mood score was computed for each PANAS completed by a participant by subtracting their negative-affect total from their positive-affect total. The resulting mood score signified the participant's mood at that particular point in the study. To simplify analysis and interpretation of mood data we averaged each participant's two mood scores for the learning phase to produce a single learning-phase score; we did the same for the two retrieval-phase mood scores. Statistical analyses were pursued using these single mood scores for the learning and retrieval phases. We consider the analysis of mood effects in the reinstated-key conditions before the effects in the switched-key conditions.

Mood effects in reinstated-key conditions. Mean mood scores at learning and retrieval for the reinstated-key conditions are shown in Table 2. As expected, moods tended to be rated higher (i.e., more positively) when participants heard the music in the major key as opposed to the minor key. A 2 x 2 (i.e., Key x Phase) mixed between-within ANOVA supported this observation and revealed a reliable effect of Key, $F(1, 22) = 11.4$, $MSE = 45.59$, $p = .003$, $\eta_p^2 = .34$. An effect of Phase was also found, $F(1, 22) = 19.73$, $MSE = 11.20$, $p < .001$, $\eta_p^2 = .47$, with overall mood scores being higher in the learning phase than the retrieval phase. In addition, a reliable interaction was found between Key and Phase, $F(1, 22) = 7.15$, $MSE = 11.20$, $p = .014$, $\eta_p^2 = .25$, such that the phase of the study the participant was in (i.e., learning vs. retrieval) had a greater impact when the major key was reinstated than when the minor key was reinstated.

(Table 2 about here)

These last two effects were unexpected, and relate to the way in which people's moods seems to be less positive than anticipated when the major-key music was played at retrieval. The interaction was explored using Tukey HSD pairwise comparisons. The mood scores for the two minor conditions did not differ significantly ($p = .83$), suggesting that the negative mood induced at learning was maintained at retrieval. However, the mood scores for the two major conditions did differ significantly ($p < .001$), indicating that the positive mood induced at learning had dissipated substantially by the retrieval phase. Still, it is noteworthy that mood at retrieval in the major-key condition was still rather higher than either of the mood scores in the minor-key conditions, but not reliably so (i.e., $p = .31$ when compared with the minor key at learning, and $p = .08$ when compared with the minor key at retrieval).

Interestingly, the mood effects depicted in Table 2 may go some way toward explaining aspects of the recall effects shown in Table 1. That is, the reinstated minor-key condition produced the greatest recall score and was also associated with the best overlap of mood state between learning and retrieval phases. The reinstated major-key condition produced the next best recall score, and was associated with a weaker overlap in mood state between learning and retrieval. These patterns of association between retrieval scores and mood scores appear to provide some support for a mood-mediation account of music-dependent memory. We acknowledge, however, that the correlational basis of this evidence does not permit strong conclusions to be drawn concerning the causal link between music tonality, mood scores and recall. We examine the issue of the causal role of mood mediation in promoting CDM effects later in this section using path-analysis techniques.

The basic issue of why the major-key music was less likely to induce a positive mood at retrieval than at learning in the reinstated-key condition is intriguing. Participants' feedback indicated that many had been irritated by the birdsong in the intervening phase between learning and retrieval, whilst others indicated that they had become somewhat fatigued by the experimental procedure by the retrieval phase. As such, participants may have entered the final phase of the study in a more negative mood that was difficult to overturn via the major-key music selection. Alternatively, there have been suggestions in the literature on experimentally induced mood states (e.g., Gerrards-Hesse, Spies, & Hesse, 1994) that it may simply be more difficult to induce a positive mood in participants than a negative mood -- although in the present study it appeared that positive mood induction was very successful at learning and just less effective at retrieval.

Mood effects in switched-key conditions. Mood ratings in the switched-key conditions would be expected to reveal a cross-over interaction pattern, with participants' mood states switching from happier to sadder when the major key is present at learning and the minor key is present at retrieval, whilst the opposite switch in mood should arise when the minor key is present at learning and the major key is present at retrieval. The mood data in Table 3 support this expected pattern of results. A 2 x 2 (i.e., Key x Phase) mixed between-within ANOVA revealed a marginally significant effect of Key, $F(1, 22) = 3.93$, $MSE = 45.40$, $p = .06$, $\eta_p^2 = .15$, with the major--minor switch promoting higher overall mood ratings than the minor--major switch. This unexpected effect again seems to be a result of the major key at retrieval being less potent at inducing a positive mood state than the major key at learning. The ANOVA revealed no main effect of Phase, $F(1, 22) = 2.53$, $MSE = 51.98$, $p = .13$, $\eta_p^2 = .10$. However, the predicted cross-over interaction between Key and Phase was

highly reliable, $F(1, 22) = 13.79$, $MSE = 51.98$, $p = .001$, $\eta_p^2 = .39$, indicating that that the moods of participants changed in opposite directions in the two switched-key conditions.

Tukey post-hoc tests showed that the switch in moods between learning and retrieval had successfully occurred for the major--minor switched condition (p = .006), but not for the minor--major switched condition (p = .45), although the mood change was in the right direction. The difficulty in inducing a strongly positive mood using the major key in the second phase of the study is supported by the fact that the mood score for the major key at retrieval was no different from either of the minor-key conditions (ps = .45 and .56). In contrast, the mood score for the major key at learning was reliably different from both minor-key conditions (ps = .006 and .004).

(Table 3 about here)

Path analysis of mood-mediation effect. Although our recall and mood analyses provide some intriguing hints that mood changes may mediate between tonality manipulations and the emergence of CDM effects, the data are far from clear-cut in supporting the mood-mediation hypothesis. One way to examine the role of mood states in mediating between music-tonality manipulations and recall scores is to employ path-analysis based around linear regression techniques. To this end we undertook a standard "mediation analysis" (e.g., Baron & Kenny, 1986) to determine if a mediated relationship existed that was based on mood states.

To pursue this path analysis we first computed a "change score" for each participant that captured the *stability* or *instability* of their mood across the learning and retrieval phases of the study. This mood change score is arguably the most appropriate mediating variable to enter into the analysis as it is predicted that those participants with a stable mood between learning and testing (i.e., in the reinstated-

tonality conditions) should evidence better recall than those with a non-stable mood between learning and retrieval (i.e., in the switched-tonality conditions). The predictor variable in the analysis was designated as contextual status (i.e., reinstated vs. switched learning and retrieval contexts), and the outcome variable was a participant's recall score. The results of this mediation analysis are summarised in Table 4, which shows the relationship between contextual status (reinstated vs. switched), mood change, and recall. The amount of mood mediation was very small (i.e., 0.576 - 0.561 = .015), and a Sobel test revealed the clear lack of reliability of this mediating effect, Goodman (I) = -0.27, $p = .785$. Thus, the upshot of this path analysis is that the mood-mediation hypothesis of tonality-based CDM finds little support in the present study. Indeed, it is music-tonality changes per se that seem most closely linked to the emergence of CDM effects in our dataset, with any associated mood induction seemingly playing no convincing causal role in recall performance.

(Table 4 about here)

General Discussion

The aim of the present study was twofold. First, we wanted to extend research on music-dependent memory (e.g., Balch et al., 1992) by investigating music-tonality manipulations (i.e., the presence of major vs. minor musical keys as environmental contexts), as opposed to other music dimensions that have been examined previously such as tempo (which produces a CDM effect) and instrument timbre or musical genre (which do not appear to produce a CDM effect). Second, we wanted to examine the underlying mechanism associated with music-based CDM. In particular, Eich (1995a) has provided evidence that all CDM effects -- whether related to manipulations of location or other environmental contexts such as ambient music -- may be mediated by induced mood states (see also Lewis & Critchley, 2003, for

evidence for a possible neural basis of mood-dependent memory effects). Balch and Lewis (1996) supported this “mood-mediation hypothesis” in their study of tempo-dependent memory effects. They showed, for example, that the tempo of the music influences the arousal dimension of mood (i.e., fast tempos are arousing), and that this mood dimension has to be reinstated at retrieval to aid recall.

Our particular interest in music tonality ties in with the potential role of mood-mediation in CDM, as tonality is well known for its use by composers as way to communicate atmosphere and to modulate people’s mood states. Intriguingly, however, there appears to be no published research that has directly examined musical key and CDM. Moreover, in previous studies that have successfully demonstrated CDM effects with music (e.g., Balch et al., 1992; Balch & Lewis, 1996; Smith 1995) the actual tonality of music selections was left uncontrolled, potentially confusing the interpretation of resulting effects. In our research we were particularly mindful to exercise tight control over our tonality manipulation, such that the same piece of music (a Chopin waltz) was used in all contexts, but with it being transposed from a minor key to a major key to afford the necessary contextual permutations.

In relation to our first aim, the results of our experiment provided clear evidence for tonality-dependent memory, thus extending evidence beyond the one, well-established musical dimension (i.e., tempo) that is known to produce CDM effects on immediate recall. Our study showed that when participants learnt a list of words when listening to a piece of music of a particular tonality their recall at testing was enhanced if the music tonality was reinstated rather than switched. Reinstating the minor-key music selection produced a particularly marked CDM effect.

In relation to our second aim, the assessment of participants’ mood states during the experiment provided some initial, suggestive evidence that the observed

CDM effects may have been mediated through the mood states induced by the different music tonalities. Thus, participants in the switched-key conditions demonstrated switched moods between learning and retrieval phases, and recalled fewer words than participants in the reinstated-key conditions, whose moods were more consistent throughout the learning and retrieval phases. It was also established that the contrasting musical keys induced contrasting moods in participants: The music in a minor key induced a more negative (i.e., sad) mood than the identical piece of music in a major key, as predicted. Despite this apparent support for the mood-mediation account of CDM effects, a subsequent path analysis of the mediating role of mood in tonality-based CDM provided no support for this hypothesis. Indeed, the amount of mood mediation was extremely small, with the variability in the recall data being determined almost solely by the music-tonality manipulation itself.

In summary, our study has demonstrated the existence of clear-cut tonality-based CDM effects, but has failed to established support for a mood-mediation theory of tonality effects, despite evidence for an apparent association between tonality changes and mood changes. As such, our findings, can be interpreted as falling in line with the "mental-context hypothesis" of CDM (e.g., as espoused by Smith, 1995). This hypothesis claims that mood states, if they arise, have no special status or necessary role in mediating between context manipulations and retrieval, but are instead just another form of mentally-represented cue that can work alongside other represented cues (e.g., tonality) to aid retrieval, if reinstated.

This latter interpretation, however, is somewhat weakened by its inability to accommodate previous evidence for mood mediation determining the impact of tempo changes in music-based CDM (e.g., Balch & Lewis, 1996). Of course, it might be proposed that mood mediation arises with *some* contextual changes (e.g., music

tempo) but not with other musical properties (i.e., music tonality). However, this proposal begs the question of what it is about music tempo that promotes mood-mediated CDM memory and what it is about music tonality that does not. Previous evidence has indicated that tempo can influence mood arousal but not pleasantness (e.g., Balch & Lewis, 1996), whilst tonality may shows the reverse pattern, impacting on mood pleasantness but not arousal (Husain et al., 2002). Perhaps, then, it is only arousal changes that lead to mood-mediated CDM effects whilst pleasantness changes do not? Although this is an interesting possibility, our music pre-test data actually run counter to this interpretation, as the minor-key version of the waltz promoted reliably lower feelings of both pleasantness *and* arousal than the major-key version.

The pre-test data do, however, give rise to a third account of why mood-mediation was not found to be reliable in the present study. This account relates to our use of the PANAS mood checklist as a way to measure mood changes during the experiment. The PANAS seems to be an effective method for detecting the pleasantness aspects of mood, but it may have lacked sufficient sensitivity to detect the arousal dimension of mood. Moreover, if it is mood arousal that mediates between contextual manipulations and recall (as in the case of tempo-based CDM) then the use of the PANAS may have underestimated the role of arousal as a mediating variable between tonality and recall in the present study. Indeed, it is interesting that in the music pre-test it was the arousal dimension that actually polarised the two pieces most effectively (i.e., the minor-key music was rated on the negative side of the arousal mid-point and the major-key piece on the positive side of the arousal mid-point).

Examining the possible role of mood arousal as a mediating variable between music-tonality manipulations and recall scores would seem to be a particularly useful avenue for future research. As things stand, however, we have to acknowledge that

the evidence from our study does not allow us to go beyond the position that tonality-based influences on memory may simply relate to mental context changes (cf. Smith, 1995; Tulving, 1983). Notwithstanding the uncertainty over the role of mood-mediation in music-based CDM, we believe that one of the most important contributions of our study remains its unique demonstration that the tonality dimension of music (i.e., whether it is in a minor or a major key) can lead to striking CDM effects similar to those observed with manipulations of music tempo.

References

Balch, W. R., Bowman, K., & Mohler, L. A. (1992). Music-dependent memory in immediate and delayed word recall. *Memory & Cognition, 20,* 21-28.

Balch, W. R., & Lewis, B. S. (1996). Music-dependent memory: The roles of tempo change and mood mediation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1354-1363.

Balch, W. R., Myers, D. M., & Papotto, C. (1999). Dimensions of mood in mood-dependent memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 70-83.

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

Beck, R. C., & McBee, W. (1995). Mood dependent memory for generated and repeated words: Replication and extension. *Cognition & Emotion, 9,* 289-307.

Bower, G. H. (1981). Mood and memory. *American Psychologist, 36,* 129-148.

Bower, G. H. (1992). How might emotions affect learning? In S.-A. Christianson (Ed.), *Handbook of emotion and memory* (pp. 3-31). Hillsdale, NJ: Erlbaum.

Bower, G. H. (2003). Mood and memory. *American Psychologist, 36,* 129-148.

Cann, A., & Ross, D. A. (1989). Olfactory stimuli as context cues in human memory. *American Journal of Psychology, 102,* 91-102.

Dalton, P. (1993). The role of stimulus familiarity in context-dependent memory. *Memory & Cognition, 21,* 223-234.

Eich, E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory & Cognition, 8,* 157-173.

Eich, E. (1995a). Mood as a mediator of place dependent memory. *Journal of Experimental Psychology: General, 124,* 293-308.

Eich, E. (1995b). Searching for mood dependent memory. *Psychological Science, 6,* 66-75.

Eich, E., & Forgas, J. P. (2003). Mood, cognition, and memory. In I. B. Weiner, A. F. Healy & R. W. Proctor (Eds.), *Handbook of psychology: Experimental psychology, Vol. 4* (pp. 61-83). New York: Wiley.

Eich, E., Maculey, D., & Ryan, L. (1994). Mood dependent memory for events of the personal past. *Journal of Experimental Psychology: General, 123,* 201-215.

Eich, E., & Metcalfe, J. (1989). Mood dependent memory for internal versus external events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 443-455.

Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition, 13,* 333-345.

Gerrards-Hesse, A., Spies, K., & Hesse, F. W. (1994). Experimental inductions of emotional states and their effectiveness: A review. *British Journal of Psychology, 85,* 55-78.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in the two natural environments: On land and underwater. *British Journal of Psychology, 66,* 325-331.

Holloway, F. A. (1978). State-dependent retrieval based on time of day. In B. T. Ho, D. W. Richards III, & D. L. Chute (Eds.), *Drug discrimination and state-dependent learning* (pp. 319-343). New York: Academic Press.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception, 20,* 151-171.

Hyde, T. S., & Jenkins, J. J. (1973) Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior,* 12, 471-480.

Isarida, T., & Isarida, T. K. (2004). Effects of environmental context manipulated by the combination of place and task on free recall. *Memory, 12,* 376-384.

Isen, A. M. (1984). Towards understanding the role of affect in cognition. In R. S. Wyer, & T. K Srull (Eds.), *Handbook of social cognition, Vol. 3* (pp 179-236). Hillsdale, NJ: Erlbaum.

Lewis, P. A., & Critchley, H. D. (2003). Mood-dependent memory. *Trends in Cognitive Sciences, 7,* 431-433.

Parker, A., Ngu, H., & Cassaday, H. J. (2001). Odour and Proustian memory: Reduction of context-dependent forgetting and multiple forms of memory. *Applied Cognitive Psychology, 15*, 159-171.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57,* 493-502.

Schab, F. R. (1990). Odors and the remembrance of things past. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 648-655.

Smith, S. M. (1985). Background music and context-dependent memory. *American Journal of Psychology, 98,* 591-603.

Smith, S. M. (1995). Mood is a component of mental context: Comment on Eich (1995). *Journal of Experimental Psychology: General, 124,* 309-310.

Smith, S. M., Glenburg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6,* 342-353.

Sollberger, B., Reber, R., & Eckstein, D. (2003). Musical chords as affective priming context in a word-evaluation task. *Music Perception, 20,* 263-282.

Spreen, O. & Schultz, R. W. (1966). Parameters of abstraction, meaningfulness, and pronunciability for 329 nouns. *Journal of Verbal Learning & Verbal Behavior, 5,* 459-468.

Tulving, E. (1983). *Elements of episodic memory.* Oxford, England: Clarendon Press.

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

Author Note

We would like to thank David Ridout for his expert assistance in both transposing and playing the piano music used in this study.

Table 1

Mean Number of Words Recalled in Reinstated-Key and Switched-Key Conditions

	Key of learning context				
	Major		Minor		
Key of retrieval context	*M*	*SE*	*M*	*SE*	Overall mean
Major	10.5	.54	8.0	.76	9.3
Minor	8.0	.60	12.3	.86	10.2
Overall mean	9.3		10.2		9.7

Table 2

Mean Mood Scores at Learning and Retrieval for Reinstated-Key Conditions

	Phase				
	Learning		Retrieval		
Reinstated key	*M*	*SE*	*M*	*SE*	Overall mean
Major	14.2	1.60	7.3	1.38	10.8
Minor	5.0	1.05	3.3	1.97	4.2
Overall mean	9.6		5.3		7.5

Table 3

Mean Mood Scores at Learning and Retrieval for Switched-Key Conditions

	Phase				
	Learning		Retrieval		
Switched key	*M*	*SE*	*M*	*SE*	Overall mean
Major--Minor	15.3	1.52	4.3	1.68	9.8
Minor--Major	3.8	1.68	8.2	2.88	6.0
Overall mean	9.5		6.2		7.9

Table 4

Summary of Mediation Analysis of the Relationship Between Contextual Status (Reinstated vs. Switched), Mood Change Score, and Recall Score

		B	Beta	SE(B)	p value
Contextual Status onto Recall Score		3.417	.576	0.715	< .001
Contextual Status onto Mood Change Score		-5.437	-.448	1.602	= .001
Contextual Status and Mood Change Score onto Recall Score	Contextual Status	3.329	.561	0.807	< .001
	Mood Change Score	-0.016	-.033	0.066	= .809

Appendix A

The 24 Words Used During Incidental Learning, with Mean Pleasantness Ratings Given by Participants in Major-Key and Minor-Key Learning Contexts.

	Major Key		Minor Key	
	M	*SE*	*M*	*SE*
fence	3.58	.20	3.15	.22
professor	2.96	.17	3.42	.19
shadow	3.21	.26	3.27	.24
cousin	4.23	.19	4.15	.21
library	2.90	.21	3.10	.19
maid	3.29	.24	3.35	.20
council	2.48	.18	2.83	.20
clothing	4.23	.23	4.33	.24
wool	4.27	.20	4.42	.14
apple**	4.56	.19	3.69	.25
cheek	4.10	.23	3.69	.19
prison	1.50	.15	1.79	.25
sheep	4.13	.22	3.79	.24
knife	2.23	.23	1.81	.16
nurse	4.19	.22	4.17	.24
basket	3.88	.15	3.50	.18
lamp	4.29	.16	4.12	.15
card	3.65	.18	4.19	.23
witness	2.50	.20	2.88	.23
flour	3.63	.18	3.38	.19
bush*	3.79	.23	3.19	.17
gift	5.19	.19	5.04	.19
flame*	4.50	.22	3.77	.25
tribe	3.56	.23	3.27	.18
Overall	3.62	.10	3.51	.08

Note. Only statistically reliable ANOVA comparisons for major-key and minor-key word-pleasantness ratings are indicated. $^{*}p < .05$. $^{**}p < .01$.

Appendix B

Mood and Familiarity/Oddness Grids Used in the Music Pre-Rating Study

Participants used the following grids to characterise: (1) how the music made them feel on the dimensions of *pleasantness* and *arousal,* and (2) whether or not the music seemed *familiar* or *odd*. Participants placed a cross in each grid to register a response.

EXTREMELY HIGH AROUSAL

EXTREMELY UNPLEASANT

EXTREMELY PLEASANT

EXTREMELY LOW AROUSAL

EXTREMELY HIGH FAMILIARITY

EXTREMELY LOW ODDNESS

EXTREMELY HIGH ODDNESS

EXTREMELY LOW FAMILIARITY