Spacing and induction: Application to exemplars presented as auditory and visual text

Norehan Zulkiply,a,*, John McLeana, Jennifer S. Burt,a, Debra Bathb

aUniversity of Queensland, Australia
bGriffith University, Australia

ARTICLE INFO
Article history:
Received 10 February 2011
Received in revised form 8 November 2011
Accepted 10 November 2011

Keywords:
Spacing effect
Inductive learning
Category learning
Category induction
Category discrimination

ABSTRACT
It is an established finding that spacing repetitions generally facilitates memory for the repeated events. However, the effect of spacing of exemplars on inductive learning is not really known. Two experiments using textual material were conducted to investigate the effect of spacing on induction. Experiment 1 and 2 extended the generality of recently observed spacing benefits to texts, regardless of whether they were visually or aurally presented. Interestingly, participants in both experiments judged massing to be more effective than spacing though their performance showed the opposite. Possible explanations for the superiority of the spaced condition over the massed condition in inductive learning, practical implications of the present study and suggestions for future research are discussed.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Inductive learning involves the process of learning by example. In inductive learning, students learn a set of exemplars of a category and try to induce a general concept from this set of observed exemplars. For example, children learn that a dark green exemplar and a light green exemplar belong to the same category of colour, green, after a series of exposures to the different colour tones of green that exist in their world. Inductive learning is one of the many ways that can be used to learn concepts and categories. Very few studies have been conducted to investigate the effect of spacing on inductive learning to date and they have produced mixed findings. The present studies, using textual materials, aimed to further examine the impact of spacing on inductive learning of categories or concepts.

The spacing effect has more generally been the subject of interest in a plethora of research studies since the late 1800's (Ebbinghaus, 1964). The effect demonstrates the situation whereby memory is enhanced when learning events are distributed across time, rather than massing them together (e.g., Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Donovan & Radvansky, 1999; Melton, 1970). Hundreds of articles in the memory literature, including a number of reviews (e.g., Dempster, 1996) and meta-analyses (e.g., Cepeda et al., 2006; Donovan & Radvansky, 1999) have found a spacing effect across a wide variety of tasks and contexts, as well as showing its benefit in learning and memory. Research has found that the spacing effect holds for many different types of stimuli which include nonsense syllables (e.g., Ebbinghaus, 1985/1913), pictures (e.g., Hintzman & Rogers, 1973), words (e.g., Glenberg & Lehmann, 1980), sentences (e.g., Rothkopf & Coke, 1966) and faces (e.g., Cornell, 1980). Additionally, the benefits of spacing have been found in experiments that involved learning complex judgement tasks (e.g., Helsdingen, Van Gog, & Van Merriënboer, 2011).

Although numerous past studies have been conducted to investigate the spacing effect, the majority of those studies have focused on investigating the spacing effect using pure memory experiments—that is, testing the effect on memory retention of repeatedly presented items (e.g., words or pictures) that participants have learnt earlier during the study phase. In a typical study that tests the spacing effect, participants are asked to remember words that are presented multiple times with a variable degree of spacing between them, and at the end of the session, participants are asked to free recall the words (e.g., Childers & Tomasello, 2002; Rea & Modigliani, 1987; Toppino, 1993).

In inductive learning studies, participants are typically asked to study a series of exemplars from a number of categories which are presented either in massed or spaced fashion, and later their
induction is tested by asking the participants to indicate to which category each novel exemplar belongs (e.g., Kornell & Bjork, 2008). Studies that investigated the effect of spacing on inductive learning have produced contradictory results. In two earlier studies, massing was superior to spacing (Gagne, 1950, who used four categories of nonsense-figure/nonsense-syllable pairs; Kurtz & Howland, 1956, who used four categories of drawings). Less direct evidence that massing facilitates induction comes from experiments that compared exact and non-exact repetitions (e.g., Appleton-Knapp, Bjork, & Wickens, 2005; Dellarosa, & Bourne, 1985; Glover & Corkill, 1987; Melton, 1970), and research on motor learning which involved learning complex motor skills (Wulf & Shea, 2002).

Interestingly, recent studies have shown the opposite finding, that induction profited from spacing (Kang & Pashler, in press; Kornell & Bjork, 2008; Kornell, Castel, Eich, & Bjork, 2010; Vlach, Sandhofer, & Kornell, 2008; Wahlheim, Dunlosky, & Jacoby, 2011). Specifically in their experiments, Kornell and Bjork (2008), Kornell et al. (2010), and Kang and Pashler (in press) used paintings from different artists, Vlach et al. (2008) used different categories of novel objects which were constructed using arts and craft supplies and objects from hardware stores, whereas Wahlheim et al. (2011) used different categories of bird families.

Surprisingly, Kornell and colleagues (i.e., Kornell & Bjork, 2008; Kornell et al., 2010) found that the majority of their participants reported massing to be more effective than spacing though their performance showed the opposite, and this finding was extended to the older population in the Kornell et al.’s (2010) study. This finding is consistent with the view that people’s access to their complex mental processes is not very accurate (Nisbett & Wilson, 1977). Nisbett and Wilson argued that typically, we are conscious of the products of our thinking, but only vaguely conscious of the process of our thinking (Nisbett & Wilson, 1977). Thus, in judging whether it was massing or spacing that had helped in learning more, participants in Kornell and Bjork (2008), and Kornell et al. (2010) may have known that they had learnt the categories of artists but may only have had a vague idea of how they learnt them. Prior to conducting the present study, we performed a pilot study to replicate Kornell and Bjork’s (2008) experiment with paintings, and our results validated their findings in both the participants’ performance and their judgements about massing. This successful replication led to the present attempt to generalise the spacing effect found in Kornell and Bjork’s (2008) experiment to textual materials.

Several theories and explanations have been given to explain the spacing effect on inductive learning. Vlach et al. (2008) argued that spacing allows time for forgetting, and forgetting helps to promote abstraction. In induction, when participants were required to identify items that had not been previously presented, abstract memories (of categories) may have helped (Vlach et al., 2008). Additionally, Kornell and Bjork (2008) argued that the interleaving of exemplars that was intrinsic to the spaced condition might have fostered and enhanced discrimination learning, or facilitated comparison and contrast, which gave an advantage to spacing.

1.1. The present study

The specific aim of the present study was to investigate whether the spacing effect in inductive learning generalises to textual materials. Two experiments using visually presented texts (Experiment 1) and aurally presented texts (Experiment 2) were conducted to examine whether or not the effect of spacing can be generalised to texts. The methods followed those developed by Kornell and Bjork (2008) using paintings. Textual presentation is important because of its educational relevance to university teaching and learning. Besides pictures, visually and aurally presented texts are commonly used to present information, as in a lecture or speech, and understanding which text presentation method (spaced condition or massed condition) works best will be beneficial to students. As mentioned previously, the focus was on the inductive learning that occurs during category learning.

1.2. Hypotheses

Based on a view that the spacing effect in induction observed by Kornell and Bjork (2008) is not restricted to their pictorial materials, we devised the following hypotheses:

1) Spacing of categories would produce superior learning of categories from text descriptions presented visually than massed presentation of categories.

2) Spacing of categories would produce superior learning of categories from text descriptions presented aurally than massed presentation of categories.

3) The majority of participants would report massing to be more helpful to learning than spacing.

2. Experiment 1

Experiment 1 examined whether or not the spacing effect found in the past experiments using paintings (e.g., Kornell & Bjork, 2008) can be generalised to texts. Specifically, visually presented texts were used in this first experiment.

2.1. Method

2.1.1. Participants and design

Forty students (22 females, 18 males) from an introductory psychology class participated in the experiment for course credit. Presentation style at study (massed vs spaced) was varied within participants. Following Kornell and Bjork (2008), participants completed a presentation (study) phase, a distractor task, a test phase and a question phase. The test phase required the participants to make categorisations of exemplars not seen in the presentation phase. Items were presented in three blocks, with feedback on accuracy provided after each exemplar. With 36 cases studies from six psychopathological categories (six case studies per category), a slightly different way of arranging the cases in the learning blocks (i.e., in the presentation phase) and in the test blocks was implemented here. For each of the six psychopathological categories, three of the cases were assigned to the study phase and another three were assigned to the test phase. In the presentation phase, the 18 cases were arranged in six learning blocks (three blocks for massed presentation; three blocks for spaced presentation). The order of the blocks was ‘MSMSSM’ (M for massed; S for spaced). The assignment of psychological disorders to condition (massed vs. spaced) was counterbalanced over two lists. Two versions of each list were produced in which there was a different assignment of disorders to blocks. Thus, there were four lists in total. In the test phase, the 18 new cases from the six psychopathological categories were arranged in three test blocks. Each block consisted of one new case from each category, presented in a fixed order across participants.

2.1.2. Materials

The materials were 36 case studies developed from six categories of psychopathological disorders. As noted, 18 cases were used in the presentation/study phase (three cases per category) and another 18 cases were used in the test phase (three cases per category). The psychopathological disorder categories used were identified by nonsense names to minimise the effects of
participants' prior assumptions and expectations. Table 1 below illustrates the six disorder categories chosen as the basis of the case studies as well as the novel names assigned to each of the categories. Each case study was between 100 and 120 words in length and incorporated a description of a few symptoms representative of the four factors of symptoms in general: Cognitive, Behavioural, Emotional, Physical. All the case studies used in Experiment 1 were pilot-tested by ten Clinical PhD students. Samples of case studies used in Experiment 1 are provided in the Appendix.

2.1.3. Procedure

Participants were tested in a small group with a maximum of four persons per session. Participants first were instructed about the nature of the experiment that they were going to do, before they entered the presentation phase that was subject to experimental manipulation. When the participants began the presentation phase, they were first asked to read 18 case studies and study the nature of the cases. Each case was shown on a computer screen for 30 s, with the novel label of the corresponding category displayed underneath. Next, participants were asked to complete a distractor task, during which they were asked to count backward by 3 s starting from 547, for 15 s, while typing the numbers in the designated box on the computer screen. Later, participants were shown 18 new case studies in the test phase, which they had not read before, and they had to identify to which category each case belonged. During the test phase, participants were shown one case study at a time on the computer screen, with seven buttons displayed below the case study. Six of the buttons were labelled with the names of the case categories and one button was labelled ‘I don’t know’. Participants selected the category for each case study by clicking their computer’s mouse on one of the seven buttons. Feedback was given after each response. If participants clicked on a correct category name for each displayed case study, the word ‘correct’ would appear in the middle of the computer screen. If they were wrong, the correct category name would be presented on the computer screen. Participants completed the test phase at their own pace. After the test phase, participants read a description about the meanings of the terms ‘massed’ and ‘spaced’ on the computer screen. They were asked, ‘Which option do you think helped you learn more?’ and were provided with three possible answers: ‘massed’, ‘about the same’, or ‘spaced’. This question phase ended the experimental manipulation. Participation in the experiment took approximately 40 min and participants were debriefed about the experiment.

2.2. Results

A two-way repeated measures ANOVA was conducted on the data for Experiment 1. Interestingly, Experiment 1 revealed findings that were consistent with Kornell and Bjork (2008). Pertaining to Hypothesis 1, participants’ performance in spaced study was significantly better than their performance in massed study, $F(1, 39) = 41.40, p < .001$, $\eta^2 = .52$, and participants’ performance also increased significantly across test blocks, $F(1.94, 75.90) = 7.84, p = .001$, $\eta^2 = .17$ (as illustrated in Fig. 1). There was also a significant interaction between presentation condition and test block which indicates that the type of presentation style, massed and spaced, had different effects over the test blocks, $F(1.9, 74.09) = 3.68, p = .032$, $\eta^2 = .09$. The interaction reflects the fact that there was less opportunity for feedback to improve performance over blocks in the spaced than in the massed condition. Table 2 shows the means and standard deviations for all test blocks in each presentation condition (i.e., massed vs. spaced) for Experiment 1.

As depicted in Fig. 2, the pattern of participants’ judgements regarding which study presentation helped them learn more was similar to that obtained in Kornell and Bjork (2008), and was in line with Hypothesis 3, with a majority of 32 (80%) claiming massed was most effective, five (12.5%) claiming spaced was most effective, five (12.5%) claiming spaced, and three (7.5%) of the participants judging the two conditions equally effective, regardless of their performances in the two conditions (i.e., massed and spaced). A one-way Chi-square analysis was conducted to compare the proportion of participants who judged massed to be most useful with the proportion preferring the spaced condition, and the proportion affirming that the two conditions contributed equally in helping them to learn more during the study phase. The result confirmed Hypothesis 3, $\chi^2 (2, N = 40) = 39.35, p < .001$. In terms of categorisation performance, 28 (70%) of the participants performed better in the spaced condition and three (7.5%) participants performed better in the massed condition, with the remainder performing equally in the two conditions.

2.3. Discussion

Students were only moderately successful in the categorisation task during the first test block, suggesting that they were not able to solve the category discrimination on the basis of their knowledge of

![Fig. 1. Proportion of psychopathology disorder categories selected correctly on the test in Experiment 1, as a function of presentation condition (massed or spaced) and test block. Error bars represent standard errors.](image)

![Table 1](image)

<table>
<thead>
<tr>
<th>Categories of psychopathological disorders</th>
<th>Novel names assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obsessive Compulsive Disorder</td>
<td>Duv</td>
</tr>
<tr>
<td>Phobia Disorder</td>
<td>Baj</td>
</tr>
<tr>
<td>Schizophrenia Disorder</td>
<td>Tem</td>
</tr>
<tr>
<td>Attention Deficit Disorder (Inattentive Type)</td>
<td>Pliq</td>
</tr>
<tr>
<td>Attention Deficit Disorder Hyperactive and Impulsive Type</td>
<td>Hix</td>
</tr>
<tr>
<td>Depression Disorder</td>
<td>Wos</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test block</th>
<th>Presentation conditions</th>
<th>Massed</th>
<th>SD</th>
<th>Spaced</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.18</td>
<td>0.20</td>
<td>0.49</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.30</td>
<td>0.22</td>
<td>0.52</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.40</td>
<td>0.25</td>
<td>0.53</td>
<td>0.28</td>
<td></td>
</tr>
</tbody>
</table>
introduction Psychology. It is likely that the use of nonsense labels for the categories discouraged them from trying to apply their prior knowledge to the task. In terms of the spacing effect, Experiment 1 provided results that are consistent with past studies (e.g., Kornell & Bjork, 2008; Kornell et al., 2010). Participants’ performance in the spaced condition outperformed their performance in the massed condition. Performance in the massed condition was greatly improved over test blocks and this could be due to the accuracy feedback they received after each test trial. Performance in the spaced condition demonstrated high accuracy from the first test block. On the post-experimental questionnaire the majority of participants appeared to believe that massed presentation made it easier to recognize the nature of each individual psychopathological category during the presentation or study phase.

3. Experiment 2

To further investigate the generality of the spacing effect over texts, Experiment 2 was performed to determine if the spacing effect can also be generalized to aurally presented texts. Using the auditory modality is a primary method of teaching in the school classroom and in university lectures, thus it would be interesting to find out whether or not spacing the textual material presented aurally also enhances learning.

3.1. Method

3.1.1. Participants and design

Participants were 40 students (27 females, 13 males) from an introductory psychology class who received course credit to participate in the experiment. All major aspects of the design of Experiment 2 were identical to those in Experiment 1 except that the case studies were aurally presented during the presentation phase.

3.1.2. Materials

The materials (i.e., case studies) used in this experiment were identical to those used in Experiment 1 (see Table 1). Each case study was recorded onto the computer by an Australian speaker, with a maximum duration of 40 s.

3.1.3. Procedure

The procedure for Experiment 2 was identical to that used in the previous experiment, except that in the presentation phase, each case study was aurally presented to the participants for approximately 40 s. While participants listened to each of the case studies, the novel label of the of the corresponding category was displayed on the computer screen.

3.2. Results

A two-way repeated measures ANOVA was conducted on the data for Experiment 2. With respect to Hypothesis 2, we found that participants’ performance in spaced study was significantly better than their performance in massed study, $F(1, 39) = 11.72$, $p = .001$, $\eta_p^2 = .23$, and participants’ performance had also increased significantly across test blocks, $F(2, 78) = 4.34$, $p = .016$, $\eta_p^2 = .10$ (as illustrated in Fig. 3), which was consistent with the findings of the first experiment. However, the interaction between presentation condition and test blocks was not significant, $F(2, 78) = 1.62$, $p = .204$. Table 3 shows the means and standard deviations for all test blocks in each presentation condition (i.e., massed vs. spaced) for Experiment 2.

Interestingly, with regard to Hypothesis 3, participants’ responses to the questionnaire administered after the test revealed consistency with the previous experiment. A one-way Chi-square analysis showed a significant difference between the three judgement options, $\chi^2 (2, N = 40) = 10.55$, $p = .005$. Of a total of 40 participants, a majority of 23 (57.5%) claimed massed presentation was best, eight (20.0%) claimed spaced and nine (22.5%) judged that both massed and spaced conditions contributed equally in helping them to learn during the learning phase, regardless of their performance in the two conditions (i.e., massed and spaced), as depicted in Fig. 4. With respect to categorisation performance, 23 (57.5%) of the participants performed better in spaced condition, 10 (25%) performed better in massed condition while another seven (17.5%) participants performed equally effectively in both massed and spaced conditions.

### Table 3

<table>
<thead>
<tr>
<th>Test block</th>
<th>Presentation conditions</th>
<th>Massed</th>
<th>SD</th>
<th>Spaced</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>0.31</td>
<td>0.40</td>
<td>0.55</td>
<td>0.33</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>0.43</td>
<td>0.31</td>
<td>0.57</td>
<td>0.26</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.48</td>
<td>0.34</td>
<td>0.58</td>
<td>0.29</td>
</tr>
</tbody>
</table>
different concepts in juxtaposition would facilitate discrimination. It is low, it might be expected that the placing of exemplars from
styles, interleaving is the primary causal factor in the benefits of the spaced condition, the categories of the learning materials used in the present study were likely
to have a low degree of discriminability. When the differences between categories are easy to detect, massed presentations may
facilitate discrimination learning (cf. Kornell & Bjork, 2008). In the spaced condition, the categories of the textual materials used in the present study were likely
to be the primary factor in facilitating discrimination learning (cf. Kornell & Bjork, 2008). In the spaced condition, the categories of the learning materials (i.e.,
psychopathological disorder names used in Experiment 1 and Experiment 2) were interleaved, which might have enhanced
discrimination learning, allowing the participants to compare and
contrast consecutive exemplars to identify the different categories.

3.3. Discussion

The results of Experiment 2 replicated the results of Experiment 1, in that performance was significantly better in the spaced
condition and significantly increased over the test block. Furthermore, when using the aurally presented cases as the learning material, participants demonstrated a similar pattern of judge-
ments favouring massed presentation, as in Experiment 1 (which used visually presented case studies).

4. General discussion

Two experiments were conducted in the present study, which aimed to examine the effect of presentation style (massing versus
spacing) on inductive learning. Both experiments were extensions to Kornell and Bjork’s (2008) study. The results of the two experi-
ments revealed that induction profited from spacing. Interestingly, both experiments provided experimental evidence confirming the generalisability of the spacing effect to texts, regardless of whether
they were visually or aurally presented (Hypotheses 1 and 2).

The better performance in the spaced condition could be attributed to the effects of spacing on the discrimination process
that is involved in induction. Furthermore, because the present study provides no evidence of the effect of temporal spacing
without interleaving in inductive learning, there is a possibility that
interleaving, and not spacing itself, may be the primary factor in
facilitating discrimination learning (cf. Kornell & Bjork, 2008). In
the spaced condition, the categories of the learning materials (i.e.,
psychopathological disorder names used in Experiment 1 and
Experiment 2) were interleaved, which might have enhanced
discrimination learning, allowing the participants to compare and
contrast consecutive exemplars to identify the different categories
of the psychopathological disorders illustrated by the case studies.
This key question is the focus of ongoing research which examines
interleaving and the nature of the discrimination between cate-
gories. Additionally, a recent report by Kang and Pashler (in press)
indicates that, in their task which used painter names and painting
styles, interleaving is the primary causal factor in the benefits of
spacing for induction.

Kornell and Bjork (2008) also endorsed the argument by Kurtz
and Hovland (1956), that stated, “When the degree of discrimina-
bility is low it might be expected that the placing of exemplars from
different concepts in juxtaposition would facilitate discrimination learning” (p. 242). It is likely that the paintings used in Kornell and
Bjork (2008) had a low degree of discriminability, probably due to the
fact that the dimensions of the paintings go beyond superficial aspects (e.g., colour). Thus it was more difficult for participants to
discriminate the paintings and this would provide an advantage to
the interleaving of exemplars in the spaced condition. In the present study, the case studies were based upon four factors of
symptoms (Cognitive, Behavioural, Physical and Physiological). In
light of the argument put forth by Kurtz and Hovland (1956), the
question that may be asked here is, why did induction profit from
spacing in this situation when the four factors of symptoms seem
definable? Understanding and distinguishing between the symp-
toms should have been easier than understanding the style of each
artist in Kornell and Bjork’s (2008) experiment. In answering this
question, we first note that even though the dimensions (symptom
factors) used in describing the cases in Experiment 1 and Experi-
ment 2 seem definable and distinctive, participants were not
informed at any stage of the experiment that they were to study the
cases that incorporated these four descriptions of symptoms
factors. In other words, participants did not know in advance what
the critical dimensions were that differentiated the categories.
Exemplars of cases from all six categories were presented to participants and they needed to use induction to extract the four
general factors of symptoms from the set of presented case studies.
As participants read (Experiment 1) or listened to (Experiment 2)
each presented case, they had to understand the nature of the
different symptoms of each case based upon the descriptions given,
and simultaneously, they had to learn to categorise the cases into
different psychopathological disorders.

Thus, it is likely that the discrimination among categories was
difficult, although the case studies outlined the key symptoms that
were needed for the classification task. Furthermore, text pro-
cessing may have added attentional demands that were not present
for the paintings. Research on reading has made clear the central
role that attention plays on the comprehension process (Kintsch
& Van Dijk, 1978). According to Kintsch’s model of text comprehen-
sion, at the beginning of reading, readers form a mental repre-
sentation of the text, and during this process, word by word, new
propositions and relations are constructed and added to this
representation. Later in the integration phase, readers use their
knowledge and experience related to the current mental repre-
sentation to build the most consistent and coherent ideas about the
meaning of text. Comprehension is involved in reading as well as in
listening, and Kintsch’s model applies to both (Kintsch, 1998, pp.
93–120).

In summary, as suggested by Kurtz and Hovland (1956), inter-
leaving may provide an opportunity to compare and contrast
exemplars from different categories to help participants solve
difficult category discriminations. We have argued that the cate-
gories of the textual materials used in the present study were likely
to have a low degree of discriminability. When the differences
between categories are easy to detect, massed presentations may
help participants integrate the members of a category. A question
for future research is whether the spacing/interleaving benefit is
observed only when category discrimination is difficult.

Spacing may also affect the attention paid to each exemplar. In
the context of a memory study, Zechmeister and Shaughnessy
(1980) found that subjects tend to overestimate the degree to
which they would remember massed items since the same item
appears again right after its presentation, causing them to pay little
attention to the second and subsequent presentations of the item. In
contrast, subjects might not feel overconfident when repetitions
of an item are spaced apart and thus might be more likely to pay full
attention to the second and next occurrence of a repeated stimulus.
Applying this to the inductive learning context, it is likely that the
massed condition may produce inattention. For example in the
massed condition, participants may experience a sense of famil-

ity with an exemplar because of its similarity to the prior
exemplar, and therefore pay little attention to the second and

subsequent exemplars. On the other hand, the spacing of exemplars may encourage more attention to or deeper processing of the second and next occurrence of an exemplar, resulting in better learning.

The spacing effect found in the present study may also rely on the interaction between induction and abstraction. Vlach et al. (2008) argued that spacing allows time for forgetting, and forgetting promotes abstraction. For the spaced exemplars in the present experiments, the gap between studying the first exemplar and the next exemplar of a particular category allowed time for participants to forget the surface characteristics of each exemplar. Additionally, it allowed the participants to come out with a more general and abstract representation of the central features of those exemplars from the same category. For example, a possible central feature of all cases in schizophrenia disorders is hallucination which can involve any of the five senses (e.g., auditory hallucination and visual hallucination). To provide a rich experiential structure to the case studies, a variety of different types of hallucinations were included. Abstract memories tend to be more durable than concrete memories, thus at testing, the more abstract representations engendered by spacing would be beneficial (Brainerd & Reyna, 2002).

Finally, it has been argued that memory plays an important role in categorisation (Vlach et al., 2008). Vlach et al. claimed that memory plays a critical factor in categorisation in two ways. “First the formation of categories depends on one’s ability to remember previous category exemplars” (Vlach et al., 2008, p. 166). As participants attempted to form an idea of the nature of each different psychopathological disorder (as in Experiment 1 and Experiment 2), they needed to remember the nature of the psychopathological disorders in the previous presented exemplars. In other words, it is likely that success in remembering those central features of a particular psychopathological category will facilitate category formation. In the current experiments, a memory benefit for spacing most probably would arise from improved encoding because of the attentional engagement promoted by the category changes from trial to trial, as discussed above. It is possible that memory effects might be more evident over a longer retention interval between presentation and test phases. In all previous research on the spacing effect in inductive learning, as well as in the present study, the spacing effect was measured over a brief term retention as the test was given shortly after the study phase. In future research it would be interesting to examine the effect of spacing on longer term retention, which obviously is also important for educational practice.

Most interesting in the present findings is that the majority of the participants judged massing as more effective than spacing, when their actual performance revealed the opposite (Hypothesis 3). One possible factor for explaining these results is the sense of familiarity towards the similar exemplars that were presented consecutively in the massed condition. Familiarity may not only reduce attentional engagement, as argued previously, but also affect the participants’ perception of the effort that they have made. If participants viewed massed presentation as involving or requiring less effort, they may have inferred that the learning task was easier in the massed condition, and thus that their learning outcomes were better than in the spaced condition. It would be interesting to further examine this familiarity explanation in future research.

With respect to the use of feedback during the test phase in the present study, we provided accuracy feedback after each test trial. As mentioned previously, we pilot-tested Kornell and Bjork’s (2008) experiment with paintings, and the accuracy feedback during the test stage facilitated the comparison with their findings. In order to maintain consistency in the approach over materials, we maintained the feedback in the present experiments with textual material. Thus we were able to make a fair comparison of the participants’ performance across all experiments with paintings and texts (visual and auditory). Apparently, in both experiments of the present study, the difference between massed and spaced was observed in the very first block of testing (and in fact was the largest there — see Figs. 1 and 3), when presumably, feedback had minimal influence. This indicates that the inclusion of feedback did not compromise the results of the present study. However, the interaction effect found in Experiment 1 might be due to providing feedback during testing.

5. Conclusions

In summary, this study provides evidence that the spacing superiority effect can be generalised to the learning of categories from descriptive texts and that, extending Kornell and colleagues’ finding (i.e., Kornell & Bjork, 2008; Kornell et al., 2010), the majority of the learners incorrectly believed that massed learning was superior to spaced learning. A consistent finding on the spacing effect obtained in both experiments of the present study seems to suggest that there is a good generality of the finding of the spacing effect in category learning.

In terms of the practical implications, the present study provides experimental evidence regarding the importance of spacing in category learning. It is undeniable that the way the learning materials (e.g., pictures and texts) are presented during the learning process is crucial to ensure effective learning, as in a lecture or speech. The present study provides evidence that in category learning, regardless of the type of presentation (either by visually presented texts or aurally presented texts), induction benefits from the spaced presentation of exemplars. Inductive learning, or in particular, category learning is generally important at all levels of education, be it at school, college or university. Given the evidence from the present study, as well as from several earlier studies (e.g., Kornell & Bjork, 2008; Kornell et al., 2010; Vlach et al., 2008) which confirmed that induction profited more from spacing (and not massing), a more informed approach in category learning could be planned by educators in order to achieve the optimal benefits of the spacing effect in induction of the categories taught to the students.

A second practical implication for university and other educational contexts concerns the dissociation between performance measures and student judgement. In each of the two studies, students believed they had benefited more from massed presentation, when the opposite was shown by their performances, to be true. Reliance on student judgements about conditions which assist their learning would seem questionable in light of these findings.

Appendix

Examples of case studies from Experiment 1 and Experiment 2

Category TEM (Schizophrenia disorder)

Sample 1: Wills, 35 years old, is a successful businessman but lately, his behavioural changes seemed to affect his relationship with clients. Since 6 months ago, he had begun to hear voices that tell him he is not a good man. He has begun to talk to himself about how bad he is during meetings with clients. This has affected his relationship with clients. At the office, his workers were shocked by his very rapid changing mood, from happy to sad to angry, for no apparent reason. When he talked, it seemed that he was having thought disturbances, as he mixed up unrelated issues and could not connect his thoughts logically. He also keeps rolling up his tongue and that is somewhat annoying to his workers.

Sample 2: John, 42 years old, is a self-employed businessman but lately, his work performance has degenerated. He has been having thought disturbances, as he mixed up unrelated issues and could not connect his thoughts logically. He also keeps rolling up his tongue and that is somewhat annoying to his workers.
Sample 2: Melinda, a 40-year-old lady complained to her neighbours that she was fearful, depressed, and couldn’t get off to sleep at night. She said she had been seeing her late mother lately, and her mother told her that her husband was going to hurt her badly. Melinda’s husband was confused with Melinda’s unusual behaviour, such as staring at him and locking herself in another room at night to avoid him. Two weeks later, Melinda ran away and stayed with her friend. While there, she wrote a letter to her husband saying that she was protected by a superpower and can never be hurt by anybody.

Category PLIQ (Attention Deficit Disorder-hyperactive and impulsive type)
Sample 1: Ben, 11 years old, is a cheerful child, who often has problems in concentrating and following instructions by his teacher at school. When he does his schoolwork, he will make one or two scribbles on it and then he will start to giggle and whisper with his classmates. At home, Ben often fails to complete the house work assigned to him by his parents. For exemplar, when asked to clean up his room, he does it for a minute and then does something else which will also left unfinished. Lately, Ben also complaints that he often feels hot and he drinks more water than he usually does.

References

Zechmeister, E. B., & Slaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don’t. Bulletin of the Psychonomic Society, 15, 41–44.