

Running head: SPECIFICITY OF EMOTIONAL INFERENCES

Inferring Characters' Emotional States: Can Readers Infer Specific Emotions?

Pascal Gygax, Alan Garnham and Jane Oakhill

University of Sussex

PREPRINT

Address for Correspondence:

Alan Garnham

Laboratory of Experimental Psychology

School of Biological Sciences

University of Sussex

Brighton, BN1 9QG, UK

[alang@biols.susx.ac.uk](mailto:alang@biols.susx.ac.uk)

## Abstract

Gygax, Oakhill and Garnham (2003) showed that, contrary to the assumption of earlier research (e.g. Gernsbacher, Goldsmith, & Robertson, 1992), readers do not infer specific emotions such as guilt or boredom. This paper presents evidence for the non-specificity of emotional inferences regardless of the nature of the stories. In Experiment 1 and 2, Gygax et al.'s stories were made longer. In Experiment 1 (off-line), people rated specific emotions as more likely, but in Experiment 2 (on-line), there was no difference between target sentences containing different matching emotions, although participants took longer to read sentences containing emotions mismatching the stories. In Experiment 3 and 4, the stories included a coherence break resolvable by inferring the main character's emotional state. In Experiment 3 (off-line), people rated specific emotions as more likely, but in Experiment 4 (on-line), there was, again, no difference between target sentences containing different matching emotions.

### Inferring Characters' Emotional States: Can Readers Infer Specific Emotions?

In everyday situations, people need to monitor each other's emotions (Ze'ev, 2001). This process enables us to anticipate people's behaviours and reactions. In the same way, understanding the emotions of the main character in a narrative enables readers to understand and anticipate that character's actions and thoughts (Miall, 1989). In essence, inferring the characters' emotions facilitates, or guides, the understanding of the text. Some researchers (DeVega, Leon, & Diaz, 1996; Gernsbacher, Goldsmith, & Robertson, 1992; Gernsbacher & Robertson, 1992; Gernsbacher, Hallada, & Robertson, 1998) have shown that readers infer the main character's emotional response during reading. These researchers introduced the notion of specificity by suggesting that readers' mental representation of the main character's emotional response is as specific as "bored" or "sad". Gygax, Oakhill and Garnham (2003), however, have argued that readers' emotional inferences are not as specific as previously assumed.

In Gygax et al.'s (2003) Experiment 1a, participants were presented with 24 stories (from Gernsbacher et al., 1992; see Table 1 for an example) and were asked to generate emotions appropriate to the characters in the stories. For each story, the participants generated a large number of emotions. The number of distinct answers for each story ranged from 16 to 42.

-----  
Insert Table 1 Here  
-----

In Experiment 1b, participants judged that several of the emotions generated in Experiment 1a were as consistent with the stories as the emotions initially assigned by Gernsbacher et al. (1992). They felt that the main character was very likely to feel any of these emotions. In Gygax et al.'s second experiment, reading times for sentences containing different emotions were measured. It was found that there was no significant difference in

reading times among sentences containing different matching emotions (Matching, Matching Synonym and Matching Similar). Emotions that were very close to the original Gernsbacher et al. emotion (Matching Synonym) and, perhaps more surprisingly, emotions that were only broadly similar (Matching Similar) behaved in the same way as the originally selected emotion (Matching). There was, however, a significant difference between the mismatching condition (Mismatching) and the three matching conditions.

In view of these results, Gygax et al. concluded that the emotional information inferred by readers from the stories is more general than a specific emotion. However, Gernsbacher et al. (1992) provided evidence that the inferred emotional information is not so broad as to merely specify the valence of the emotion (i.e. whether an emotion is positive or negative).

There could be a number of reasons why emotional inferences were not specific in the experiments reported in Gygax et al. (2003). One reason derives from the idea (e.g. Clore and Ortony, 1988; Ortony and Clore, 1989; Ortony and Turner, 1990; Alvarado, 1998) that emotions are constructed from a set of components or subcomponents, such as, for example, valence. In a particular situation, these components are perceived and identified through the process of appraisal, which can be defined as an organism's evaluation of the situation (Scherer, 1997; Smith and Lazarus, 1993). In the case of emotion understanding in text comprehension, readers appraise or evaluate the situation by assessing the information given in the text. Each component identified from the text is then assigned a value. For example, the component of novelty for a situation that elicits anger is characterised by a high value (i.e. highly unexpected). The combination of these components defines a particular emotion. In text comprehension, this emotion is a particular character's emotional response. In the stories used in Gygax et al., the information given to the readers might not have been sufficient to identify the components necessary to infer specific emotions. Alternatively,

readers might have identified the right components, but the values attributed to them may not have been specific enough.

Another reason why the emotional information inferred by participants in Gygax et al.'s experiments might have been relatively unspecific is related to the notion of identification, or more precisely simulation (Oatley, 1994). Simulation is the process by which readers dynamically adopt the goals and actions of the characters. These goals and actions become part of a working model of the situation that allows readers to become actively engaged in the situation portrayed by the text. However, for readers to construct a working model on which to run a simulation, the text needs to provide appropriate contextual information (Oatley, 1999a). In particular, the text needs to convey enough information for readers to adopt the character's goals and actions. Indeed, contextual information can even allow the readers to experience and understand the emotions more clearly than in real life (Oatley, 1999b). More importantly, not only is contextual information essential for simulation, but it also determines the intensity of the readers' identification with the story characters (Oatley, 1999a). In turn, the simulation and the intensity of identification determines the specificity of emotional inferences. The experience and understanding of emotion can thus be seen a continuous process that varies as contextual elements are evaluated or re-evaluated (Storm et al., 1996; Ellsworth & Scherer, in press). Hence, a lack of contextual information may result in the generation of unspecific emotional inferences. Therefore, as more information associated with the emotional state of the characters is presented, readers' representation of characters' emotional responses should become more specific.

Both arguments presented so far suggest that the stories used by Gygax et al. did not convey enough information for readers to infer specific emotions. The first two experiments in this paper examine the hypothesis that people will infer specific emotions if the stories

convey more information. In these experiments, the stories were made twice as long as those used by Gygax et al.. The purpose of Experiment 1 was to create stories that convey enough information for readers to infer specific emotions. It was hypothesised that additional information should make participants consistently choose one emotion as the main character's emotional state. The stories selected in Experiment 1 (an example is shown in Table 2) were subsequently used in Experiment 2. In that experiment reading times for sentences presenting different emotions as the main protagonist's emotional state were compared.

-----  
Insert Table 2 Here  
-----

### Experiment 1

The first aim of this experiment was to create stories that convey more information than the stories used in Gygax et al. (2003). The second aim, more importantly, was to create stories in which the main characters were judged to be experiencing one specific emotion. If this aim could be satisfied, it would suggest that the non-specificity of emotional inferences found in Gygax et al.'s experiments was a result of insufficient information presented in the stories. In terms of identification, it would support the idea that readers need more contextual information to become more engaged in the text and consequently to infer the main character's goals and actions more specifically.

It was crucial to evaluate the specificity of emotional inferences in the longer stories using an off-line measure. If it were not possible to infer specific emotions with no time or memory constraints, it would be pointless to test for specificity of emotional inferences on-line.

## Method

Participants. Thirty-nine students from the University of Sussex participated in this experiment. They were paid £3 for a session that lasted for about 20 minutes. None of the participants from previous experiments took part in this study.

Materials. The stories used in Gygax et al. (2003) were compared with new extended versions of those stories. To create the new stories, additional sentences were added to the original stories (see Table 2). The number of sentences in each story was doubled. The extra sentences were added immediately before the final sentence of each story. The extra material was intended to enhance the emotional content of the story and bias it towards one specific emotion. Each story was written so as to focus on the emotion identified by Gernsbacher et al. (1992).

Procedure and design. Thirteen participants were presented with the Short versions of the stories, and thirteen participants were presented with the Long versions of the stories. After a first analysis, some of the long stories were further modified, and an additional thirteen participants were presented with the new versions of the long stories. In each group, the participants were presented with 24 stories, each on a separate sheet of paper. The order of the presentation of the stories was random and different for each participant. After each story, the participants were presented with a sentence completion task. The sentence used for this task was:

[The main character] felt.....

The participants had to choose one answer from a list of five possible answers. These possible answers corresponded to the four conditions tested in the previous experiments (i.e. Matching, Matching Synonym, Matching Similar and Mismatching) plus an extra possibility (Other) allowing the participants to write an alternative answer. The participants were told to circle the answer that they thought was the most appropriate to complete the sentence. If they

thought none of the specific answers was appropriate, they could choose Other and write an alternative answer.

### Results and Discussion.

The hypothesis was that there should be a higher consensus following the Long versions of the stories than following the short versions. Furthermore, the consensus should be on the Matching emotions studied by Gernsbacher et al. (1992). This hypothesis can be supplemented by one that stipulates that not only should the consensus for the Matching emotions increase in the Long versions, but the consensus for the Matching Similar emotions should also decrease. Indeed, an increase in the specificity of emotion inference is most likely to create a decrease in the consensus for the latter condition.

The experiment was carried out in two stages. In the first stage, a comparison was made between the results of the Long and the Short versions of the stories (with 13 participants in each condition). Before performing any analysis, a close look at the responses to the Long versions of the individual stories suggested that some of the stories did not show consensus for the chosen specific emotions. It was decided that these stories (first Long versions), would be rewritten and tested with another 13 participants to ensure that all of the stories would imply the character's emotional response in a similar fashion.

For most of the stories, the results from the second Long versions showed an improvement. It was then decided that the Long versions (either first or second long versions) that showed the best consensus would be included in the analysis and considered for future testing. In addition, for story 19 (see Appendix A), it was decided that the emotion in the Matching Synonym condition would be considered as the specific emotion instead of the one originally tested by Gernsbacher et al. (1992). Indeed, in both the Short and (first) Long version of the story there was a consensus on the word stressed (Matching Synonym) as opposed to the expected callous (Matching). As callous is defined as the absence of feelings



(Cambridge English Dictionary), it was decided that stressed would be a more appropriate emotion word to test. Hence, stressed was considered as the Matching emotion for story 19. Table 3 shows the mean number of times (with standard deviations) that the different emotion words were chosen in the Short and in the Long versions of the stories.

-----  
Insert Table 3 Here  
-----

For each Short story, there was a corresponding Long story. Therefore, a matched-sample t-test was performed on the number of times the Matching emotions were chosen in each condition (Short vs. Long). As predicted, the difference was significant,  $t(23) = 3.46$ ,  $p < .01$ . Furthermore, the 63% consensus in the Long Version of the stories was significantly different from chance (25%),  $z = 1.87$ ,  $p < .05$ , whereas the 49% consensus in the Short versions was not,  $z = 1$ ,  $p > .05$ . Even though the consensus for the Matching condition has only increase to 63% (as opposed to 100%), the importance of this result lies in the comparison between the increase in consensus for the Matching condition (from the Short to the Long versions of the stories) and the decrease in consensus for the other matching conditions. Of special interest is the Matching Similar condition, as an increase in the specificity of emotion inference is most likely to create a decrease in the consensus for this condition. As expected, the consensus for the Matching Similar condition was smaller in the Long versions of the stories than in the Short versions. This result is supported by a significant 2X2 interaction ( $F(1;92)=4.9$ ;  $p<.05$ ) showing that the consensus for the Matching condition increases in the Long versions whereas the consensus for the Matching Similar decreases in the Long versions.

As expected, participants more often chose the Matching emotions as consistent with the Long versions of the stories than the Short versions, and participants chose the Matching Similar emotions less often as consistent with the Long versions of the stories than the Short

versions. It was therefore possible to influence the specificity of emotional inferences by adding extra information about the main character's emotional state. However, it is worth noting that, even though the Matching emotions were chosen 63% of the time in the Long versions, this figure is well below 100%. It could be that in some of the stories, the Matching Synonym and the Matching Similar emotions were too closely related to the specific Matching emotion word, hence, participants chose these related emotions as opposed to the Matching ones. Nevertheless, the results suggest that readers are capable of inferring specific emotions. For this process to occur, the stories need to be longer than the stories used in Gygax et al. (2003), and hypothetically more engaging. If the stories are made more engaging, readers can identify with the character more easily and thus understand or experience the character's emotional responses more accurately. Furthermore, the additional information might lead to a stronger activation of the character's emotional state. Hence, readers would generate more specific emotional information. However, the results of this present experiment did not tell us about the processes that occur during reading. It is still possible that while reading, people do not infer specific emotions, even though, given enough time and information, they can determine a specific emotion experienced by a story character. The following experiment assesses the hypothesis that readers can, while reading, infer specific emotions if provided with sufficient information.

## Experiment 2

The purpose of this experiment was to see if people draw specific emotional inferences during reading when provided with the additional information in the Long versions of the stories. If they do, it would support the hypothesis that the non-specificity of emotional inferences in the experiments of Gygax et al. (2003) could be explained by the short and relatively unengaging nature of the stories. The present experiment measured reading times for sentences containing different emotional terms. This paradigm was also used by

Gernsbacher et al. (1992) and Gygax et al. (2003), who assumed that it provided a measure of on-line inference processes. In addition to the three Matching emotions tested in Experiment 1, an incongruent emotion (Mismatching) was included for each story. If people infer emotions while reading, their reading times for sentences containing matching emotions should be faster than those for sentences containing Mismatching emotions. In addition, if people infer specific emotions while reading, their reading times for sentences containing Matching emotions should be faster than those for sentences containing either Matching Synonym or Matching Similar emotions. Moreover, reading times of sentences containing Matching Synonym emotions should be the closest to reading times of sentences containing Matching emotions. If readers do not infer specific emotions, but only a more general impression of how the protagonist is feeling, then there should not be any difference among the reading times for sentences containing Matching, Matching Synonym and Matching Similar emotions.

#### Method.

Participants. Twenty-four students from the University of Sussex participated in this experiment. They were paid £4 for a session that lasted for about 30 minutes. None of the participants from previous experiments took part in this study.

Materials. The Long versions of the stories from Experiment 1 were used for this experiment. As in the experimental stories, the number of sentences in each of the filler stories (also taken from Gygax et al., 2003) was doubled.

Apparatus. The stories were presented on a PC fitted with an Advantech PCLabCard, using a version of the TSCOP program (Norris, 1984). Responses were collected using response buttons attached to the PCLabCard, which permits millisecond accuracy.

Procedure and design. The participants were instructed to read each story at a normal reading speed, as though they were reading a magazine. To make sure that participants read

the stories carefully, some stories ( $N = 16$ ) were followed by a question related to the text. Participants had to answer the question by pressing a button labelled either “yes” or “no”. Each story was presented in four parts (of one or more sentences), with the last part being the target sentence. Participants were instructed to press the “yes” button when they finished reading each part.

Reading times for the target sentences were recorded. Different carrier sentences were created for the target words to prevent participants becoming too accustomed to a particular sentence structure at the end of each story. These carriers are shown in Table 4. All the target sentences were approximately the same length. In each list, each of the different target carrier sentences was randomly selected for six stories.

-----  
Insert Table 4 Here  
-----

Before the main part of the experiment, the participants read two practice stories, both of which were followed by questions, to familiarise them with the procedure and with the kinds of passages that they would be reading.

Each experimental story appeared in four conditions, defined by the target emotion words (Matching, Matching Synonym, Matching Similar, Mismatching). Since there were four conditions per story, four different lists of stories were constructed, and six participants were assigned to each list. Each list had six stories in each of the four conditions, and each story appeared the same number of times, across the experiment, in each of the four conditions. The filler stories were the same in each list. The order of presentation of the filler and experimental stories was random and different for each of the four lists.

### Results and Discussion.

The aim of this experiment was to see whether people infer specific emotions while reading if the stories provide more information relevant to the character’s emotional state. If

they do, reading times for sentences containing Matching emotions should be faster than for Matching Synonym emotions, which should be in turn faster than Matching Similar emotions. If they do not infer specific emotions, and instead only infer general emotional information, reading times for sentences containing different matching emotions should not differ. In addition, according to either hypothesis, sentences containing Mismatching emotions should be read more slowly than sentences containing matching emotions.

Table 5 shows the mean reading times for the target sentences in the four conditions. A one-way ANOVA showed a significant difference among the four conditions both by-subjects,  $F_1(3, 69) = 18.60, p < 0.001$ , and by-items,  $F_2(3, 69) = 8.37, p < 0.001$ .

-----  
Insert Table 5 Here  
-----

However, when the Mismatching condition was excluded, the analysis showed no significant differences among the Matching, Matching Synonym and Matching Similar conditions,  $F_1(2, 46) = 0.26, p > 0.05$ ,  $F_2(2, 46) = 0.17, p > 0.05$ . A series of t-tests (with Bonferroni corrections) was performed to see if the difference between the Mismatching condition and each of the matching conditions was significant. The Mismatching condition was significantly different from the Matching condition,  $t_1(23) = 6.26, p < 0.001$  and  $t_2(23) = 4.24, p < 0.001$ , from the Matching Synonym condition,  $t_1(23) = 5.08, p < 0.001$  and  $t_2(23) = 3.80, p < 0.001$ , and from the Matching Similar condition,  $t_1(23) = 5.18, p < 0.001$  and  $t_2(23) = 3.78, p < 0.001$ .

Experiment 1 suggested that readers do infer specific emotions when provided with sufficient information. However, in Experiment 1, an off-line task was used. The results of this present experiment, using an on-line task, did not support the conclusion from Experiment 1. In this present experiment, as in the experiments reported in Gygax et al. (2003), there was no difference in sentence reading times between sentences containing

different matching emotions. Nonetheless, the Mismatching condition was significantly different from each of the matching conditions.

We also performed an analysis on the 12 stories that showed the highest consensus for the Matching condition in Experiment 1 (Mean=10.33). If, as suggested by the results above, readers do not infer specific emotions while reading (i.e. on-line), the extent of the consensus found in Experiment 1 should make no difference. This was indeed the case. Even though we only took the stories (N=12) that showed the highest consensus, a one-way ANOVA showed a significant difference among the four conditions ( $F_2(3, 44) = 7.55, p < 0.001$ ), but no difference between the three matching conditions ( $F_2(2,33) = 0.72, p > 0.05$ ).

This result implies that readers do infer emotions while reading, but even if the stories are made longer and more relevant to the character's emotional state, readers do not infer specific emotions. It might be the case that readers need a significant amount of time to reflect on the character's emotional state and to determine a specific emotion. This extra time could facilitate the process by which readers identify with the character. In Experiment 1, the participants had as much time as they wanted to think about the character's emotional state. Hence, they might have had sufficient time to adopt the character's viewpoint and to make specific emotional inferences. As mentioned earlier, another reason for the non-specificity of emotional inferences might be that the stories did not compel readers to infer the main character's emotional state. Therefore, the resources allocated to assessing the main character's emotional state during reading might have been minimal, resulting in a non-specific emotional inference. The following experiments investigate this issue.

These experiments (Experiments 3 & 4), test the hypothesis that people infer specific emotions if the text compels them to infer emotional information. This idea is derived from the constructionist approach to reading comprehension, and in particular the idea that readers attempt to make sense of the text by asking why-questions. The inferences generated in

response to such questions enable readers to maintain a coherent representation of the text. Such inferences may establish either local or global coherence. Inferences that establish or maintain local coherence connect adjacent text constituents. For example, to understand the sentences “John looked at the windowsill. The family photograph was covered in dust”, the reader needs to make the inference that the photograph is on the windowsill. This inference enables the reader to link the two sentences together. On the other hand, inferences that establish or maintain global coherence, connect most constituents of a text. These inferences represent deeper features, such as the general theme, the main point, or the moral of the text (Graesser et al., 1994). In the case of literary narratives, and in the stories used in the experiments reported in Gygax et al. (2003) and in Gernsbacher et al., 1992, emotional inferences are needed to establish global coherence. The reason is that most of the events and characters’ actions in the text are related to the main character’s emotional state.

It is possible that the stories used in Gygax et al. (2003) were coherent enough not to require effort that would lead readers to infer specific emotional information. A clearer lack of coherence might, therefore, compel readers to engage in a more effortful search for explanation or, in constructionist terms, a more effortful search-after-meaning. In turn, this effortful process could result in a more specific emotional representation.

In the minimalist approach to text comprehension, inferences such as the main character’s emotions are not usually drawn while reading, unless they are based on quickly and easily available information or are needed for local coherence. In the stories used in Gygax et al. (2003), the emotional inferences were not needed for local coherence. It could, however, be argued that the emotional inferences were made because readers did possess easily retrievable emotional knowledge. The situations portrayed in the text might have been strongly associated with emotional concepts. Hence, the main character’s emotional state was inferred while reading.

Both the minimalist and constructionist arguments presented above assume that in the stories used in Gygax et al. (2003), the emotional inferences were not needed for local coherence. In this sense, readers were not forced to make emotional inferences to link different parts of the stories together. A possible way to compel readers to infer specific emotional information would be to alter the stories to render the emotional inference necessary for local coherence. Therefore, in Experiments 3 and 4, it was hypothesised that if an emotional inference is needed to link different parts of the stories, the participants should be forced to make this inference. It was expected that, as the participants were forced to make the emotional inference, they would allocate more effort to the generation of the inference, resulting in a more specific representation of the main character's emotional state. If people are forced to represent the main character's emotional response, they might increasingly focus their attention and comprehension resources on the main character. As a result, readers might identify more components of the character's emotion. The combination of these components could result in a more specific representation of emotion.

The stories used in Experiments 3 and 4 were divided into two parts. The first part of each story was made incoherent enough to induce the need for an explanation for the events it described (an example story is shown in Table 6). To link the two parts of each story and derive a coherent interpretation, readers had to infer the main character's emotional state. The incoherence of the first part could be resolved by inferring emotional information from the second part. We hypothesised that as readers are forced to infer emotional information to establish local coherence, this information should be specific. For example, one story started with the sentences "Don and his wife moved in a year ago. Tonight, Don was at home. He was in the living room, sitting on the floor". These sentences provide no obvious explanation for the fact that Don is sitting on the floor. The rest of the passage conveys information that should lead readers to infer that Don feels depressed. We hypothesised that to understand



why Don is sitting on the floor, and to link the first part of the story to the second, readers need to concentrate on Don's emotional state.

-----  
Insert Table 6 Here  
-----

A pilot-study was performed to ensure that the first part of each story induced the need for readers to find an adequate explanation of the events it described. In the following experiments, Experiment 3, like Experiment 1, assessed the specificity of emotional inferences in an off-line experiment. In Experiment 4, as in Experiment 2, reading times of sentences containing different target emotion words were compared. If readers infer specific emotions when the text compels them to infer the main character's emotional response, reading times for sentences containing Matching emotions should be shorter than sentences containing Matching Synonym emotions, which in turn should be shorter than those for sentences containing Matching Similar emotions. In any case, reading times for sentences containing matching information should be shorter than sentences containing Mismatching information.

### Experiment 3 - pilot study

In Experiment 1 and 2, the stories used by Gygax et al. (2003) were made longer, so as to convey more information relevant to the character's emotional state. The results of these experiments suggested that people can infer specific emotions, but only given enough time. The stories in the used in this and the following experiment were further changed in an attempt to compel readers to infer specific emotions.

As mentioned earlier, the stories in this experiment were divided into two parts. The first part of each story was made incoherent so as to encourage participants to find an adequate explanation for the events it portrayed. The second part of the story provided the necessary explanation, which was related to the main character's emotional state.

## Method.

Participants. A total of 27 participants took part in the pilot study. None of the participants from previous experiments took part.

Design and procedure. In each of the stories used in previous experiments, the main character's emotional response was suggested by several sentences followed by a final statement (composed of one or two sentences) that reinforced the main character's emotional response. In the stories to be used in Experiments 3 and 4, the final statement was placed at the beginning. The purpose of this manipulation was to create for each story an ambiguous first part that needed extra information to be fully understood. The rest of the story was intended to allow readers to resolve the ambiguity by inferring the main character's emotional response, thus establishing local coherence. Such a manipulation raises two important issues. The first is whether the beginnings of the stories were genuinely ambiguous. The pilot study was intended to ensure that more information was needed to understand the first part of the stories. The second issue is what emotional content, if any, the first part of the stories had. The first part of each story was constructed so that it did not convey information about the specific emotion associated with the story. Thus, to understand the first part of a story, the participants had to infer the main character's emotion from the second part of the story. Table 6 shows an example of a story in both the original and modified version. For each story, the first part was made ambiguous, and its content as emotionally neutral as possible. The participants in the Pilot Study were presented with the first (ambiguous) part of each story, composed of one or more sentences. Their primary task was to rate its ambiguity on a 7-point scale by answering the question: "Do you think more information is needed to clarify [the main character]'s behaviour?" (1 = "No, it is clear" to 7 = "Yes, it is not clear"). Their second task was to rate on a 7-point scale (1 = "not very likely" to 7 = "very likely") the likelihood

of the main character feeling the specific emotion tested in previous experiments. For each participant, the order of the presentation of the passages was random.

### Results.

The pilot was conducted in three stages. In each stage, some passages were modified to achieve scores that indicated greater ambiguity. Table 7 shows the mean and standard deviation of the Ambiguity and Emotional Content responses to the passages in the final stage of the pilot.

-----  
Insert Table 7 Here  
-----

Even though the mean for the Ambiguity responses was fairly high, it is likely that if the question related to the Emotional Content of the passages had not been presented, it would have been even higher. The question on emotion was always presented after the question on ambiguity, but the participants knew that there would be an emotion question. As a result, the participants might have primed to infer some emotional information that could resolve the ambiguity. Hence, some passages might not have seemed ambiguous. In addition, participants who gave a score of "4" on the Emotional Content scale (35% of the participants) wrote that the main character might feel any emotion. Overall, these results suggest that the beginnings of the passages did not imply the specific emotions tested in previous experiments. As the Ambiguity scores were high and the Emotional Content scores low, it was decided that the passages were appropriate for further testing.

### Experiment 3

The aim of this experiment was to see if participants would choose specific emotions as consistent with the Ambiguous versions of the stories more often than with the Short versions. It was hypothesised that the participants should allocate more effort to generating emotional inferences when reading the Ambiguous versions of the stories, as the emotional

inferences were needed for coherence. This extra effort was expected to result in a more specific representation of emotion. This experiment was the similar to Experiment 1, except that the Ambiguous versions of the stories were tested, rather than the Long versions. Therefore, the results from the Short versions of the stories from Experiment 1 were compared to the results of the Ambiguous stories tested in this experiment.

### Method.

Participants. A total of twenty-six students from the University of Sussex participated in this experiment. They were paid £3 for a session that lasted for about 20 minutes. None of the participants from previous experiments took part in this study.

Materials. The 24 stories generated in the Pilot study were used in this experiment.

Procedure and design. The participants were presented with the 24 stories, each on a separate sheet of paper. The order of the presentation of the stories was random and different for each person. The participants were asked to read each story carefully and to carry out a sentence completion task following it. The sentence used for this task was:

[The main character] felt.....

As in Experiment 1, the participants had to choose one answer from a list of five possible answers. These possible answers represented the four conditions tested in the previous experiments (i.e. Matching, Matching Synonym, Matching Similar and Mismatching) plus a fifth option (Other) allowing the participants to write an alternative answer. The participants were told to circle the answer that they thought was the most appropriate to complete the sentence. If they thought none of the specific answers was appropriate, they could choose Other and write an alternative answer.

### Results and Discussion.

The hypothesis was that there should be a higher consensus on the sentence completion task following the Ambiguous versions of the stories than following the Short

versions. Furthermore, as in Experiment 1, the consensus should be on the Matching emotions originally studied by Gernsbacher et al. (1992). As in Experiment 1, this hypothesis can be supplemented by one that stipulates that not only should the consensus on the Matching emotions increase in the Long versions, but the consensus on the Matching similar emotions should also decrease. Such a result would clearly indicate an increase in the specificity of emotional inference.

The experiment was carried out in two stages. In the first stage, thirteen participants were tested on the Ambiguous versions (first versions). Before performing any analysis, a close look at the responses to the individual stories suggested that some of the stories did not show consensus on the chosen specific emotions. It was decided that these stories would be rewritten and tested with another 13 participants to ensure that all of the stories would imply the character's emotional response in a similar fashion. For most of the stories, the results from the second Ambiguous versions showed an improvement. It was then decided that the results from the Ambiguous versions (either first or second Ambiguous versions) that showed the best consensus would be compared with the results for the Short versions of the stories (tested in Experiment 1).

For stories 13 and 23 (see Appendix B), it was decided that the emotion in the Matching Synonym condition would be considered as the specific emotion instead of the one tested by Gernsbacher et al. (1992). Indeed, in the Ambiguous version of story 13 (as in the Long version in Experiment 1), there was a consensus on miserable as opposed to the expected depressed. Depressed might have been considered more as a mood than as an emotion. In story 23, the consensus was on furious as opposed to the expected angry. Angry might have lacked the intensity of furious, which participants felt was more appropriate for story 23. It was decided that these emotion words would be taken as the primary emotions for subsequent analysis and experiments. Table 8 shows the mean number of times (with

standard deviations) that the different emotion words were chosen in the Short and in the Ambiguous versions of the stories.

-----  
Insert Table 8 Here  
-----

A matched-sample t-test was performed on the number of times the Matching emotions were chosen in each condition. As predicted, the difference was significant,  $t(23) = 3.32$ ,  $p < .01$ ). Furthermore, the 62% consensus in the Ambiguous Version of the stories was significantly different from chance (25%),  $z = 2.13$ ,  $p < .05$ , whereas the 49% response in the Short versions was not,  $z = 1$ ,  $p > .05$ . As in the first experiment reported in this paper, even though the consensus for the Matching condition has only increased to 62% (as opposed to 100%), the importance of this result lies in the comparison between the increase in consensus for the Matching condition (from the Short to the Long versions of the stories) and the decrease in consensus for the Matching Similar condition. Indeed, the consensus for the Matching Similar condition was smaller in the Long versions of the stories than in the Short versions. This result is supported by a significant 2X2 interaction ( $F(1;92)=5.6$ ;  $p<.05$ ) showing that the consensus for the Matching condition increases in the Long versions whereas the consensus for the Matching Similar decreases in the Long versions. As expected, participants chose the Matching emotions as more consistent with the Ambiguous versions of the stories. These results suggested that when readers need to make emotional inferences to establish coherence, they are more likely to build a specific representation of the main character's emotional state. Perhaps readers allocated more effort to the generation of emotional inferences in Ambiguous stories than in Short stories. The stories used in this experiment caused readers to generate specific emotions when they had as much time as they wanted. However, nothing follows about the influence of the ambiguity of the stories on the specificity of emotional inferences while reading. The next experiment assesses the

hypothesis that readers do infer specific emotions while reading, but only if they need to in order to establish local coherence.

#### Experiment 4

The hypothesis in Experiment 4 was that if the information presented in the stories forces the readers to make emotional inferences to establish coherence, the representation of the main character's emotional response created during reading will be specific.

Except for the stories, this experiment was the same as Experiment 2. We measured reading times for sentences containing different emotional terms. If people infer specific emotions while reading, then reading times for sentences containing Matching emotions should be faster than those for sentences containing either Matching Synonym or Matching Similar emotions. Moreover, reading times for sentences containing Matching Synonym emotions should be closer than times for Matching Similar emotions to reading times for sentences containing Matching emotions. If readers do not infer specific emotions, but only a more general impression of how the protagonist is feeling, even when the stories force readers to infer emotional information, then there should not be any differences among the reading times for sentences containing Matching, Matching Synonym and Matching Similar emotions. Both hypotheses predict that reading times for sentences containing Mismatching emotions should be longer than for sentences containing Matching emotional information, on the assumption that people infer (some) emotional information while reading.

#### Method.

Participants. Twenty-eight students from the University of Sussex participated in this experiment. They were paid £4 for a session that lasted for about 30 minutes. None of the participants from previous experiments took part in this study.

Materials. The Ambiguous versions of the stories generated in the Pilot study were used in this experiment. In addition to the 24 experimental stories, 24 filler stories were used. The 24 filler stories were written in the same general style as the experimental stories.

Apparatus. The stories were presented on a PC fitted with an Advantech PCLabCard, using a version of the TSCOP program (Norris, 1984). Responses were collected using response buttons attached to the PCLabCard, which permits millisecond accuracy.

Procedure and design. As in Experiment 2, the participants were instructed to read each story at their normal reading speed, as though they were reading a magazine. To make sure that participants read the stories carefully, some stories ( $N = 16$ ) were followed by a question related to the text. Participants had to answer the question by pressing a button labelled either “yes” or “no”. Each story was presented in four parts (of one or more sentences), with the last part being the target sentence. Participants were instructed to press the “yes” button when they finished reading each part.

Reading times for the target sentences were recorded. Different carrier sentences were created for the target words to prevent participants becoming too accustomed to a particular sentence structure at the end of each story. These carriers are shown in Table 4. All the target sentences were approximately the same length. In each list, each of the different target carrier sentences was randomly selected for six stories.

As in Experiment 2, each experimental story appeared in four conditions, defined by the target emotion words (Matching, Matching Synonym, Matching Similar, Mismatching). Since there were four conditions per story, four different lists of stories were constructed, and six participants were assigned to each list. Each list had six stories in each of the four conditions, and each story appeared the same number of times, across the experiment, in each of the four conditions. The filler stories were the same in each list. The order of presentation of the filler and experimental stories was random and different for each of the four lists.



## Results and Discussion.

The aim of this experiment was to see if readers infer specific emotions when provided with information that compels them to represent the main character's emotional state. More specifically, the hypothesis was that if emotional inferences are needed for local coherence, then readers should infer specific emotional information. If this hypothesis is correct, reading times for sentences containing Matching emotions should be faster than for Matching Synonym emotions, which should be in turn faster than Matching Similar emotions. If readers do not infer specific emotional information, and instead only general emotional information, reading times for sentences containing different matching emotions should not be different.

Reading times that were more than 2.5 standard deviations above or below the participant's mean were discarded from the analysis. They represented 3% of the data. Table 9 shows the mean reading times for the target sentences in the four conditions. A one-way ANOVA showed the expected difference among the four conditions, which was significant both by-subjects,  $F_1(3, 81) = 24.69$ ,  $p < 0.001$ , and by-items,  $F_2(3, 69) = 15.41$ ,  $p < 0.001$ .

-----  
Insert Table 9 Here  
-----

When the Mismatching condition was excluded, the analysis showed no significant differences between the Matching, Matching Synonym and Matching Similar conditions either by-subjects,  $F_1(2, 54) = 1.03$ ,  $p > .05$ , or by-items,  $F_2(2, 46) = 0.79$ ,  $p > .05$ . A series of t-tests (with Bonferroni corrections) was performed to see if the difference between the Mismatching condition and each of the matching conditions was significant. The Mismatching condition was significantly different from the Matching condition,  $t_1(23) = 5.53$ ,  $p < 0.001$  and  $t_2(23) = 4.51$ ,  $p < 0.001$ , from the Matching Synonym condition,  $t_1(23) = 6.46$ ,  $p < 0.001$  and  $t_2(23) = 4.93$ ,  $p < 0.001$  and from the Matching Similar condition,  $t_1(23) =$

= 6.25,  $p < 0.001$  and  $t_2(23) = 4.78$ ,  $p < 0.001$ . It should be noted first that the reading times and the standard deviation in the Matching Similar condition were numerically different from those in the other conditions and second that the reading times in this experiment were generally longer than those in Experiment 2. These two points raise one important issue regarding the distribution of the times in the different matching conditions. So far, the different distributions have been assumed to be normal (and hence uni-modal). However, it could be the case that average reaction times reflect two type of response: rapid responses to the inferred emotion and slower responses to a related emotion. Differences between the conditions (although not significant) could arise from a change in the balance between faster and slower responses. As a result, one would expect the standard deviations to rise. In addition, the conclusion that readers do not infer specific emotions would have to be revisited. However, the distributions of times in the different matching conditions in Experiment 4 (see Figure 1) shows no evidence of bi- or multi-modality.

-----  
Insert Figure 1 Here  
-----

In addition, we also conducted an analysis on the standard deviations of the matching conditions. Participants may have reacted differently to the Matching Similar condition than they did to the other conditions. A one-way ANOVA showed no difference in the standard deviations ( $F_1(2, 46) = 0.73$ ,  $p > 0.05$ , and  $F_2(2, 46) = 0.47$ ,  $p > 0.05$ ), which suggest that readers were reacting similarly to all three matching conditions.

These results suggest that the story manipulation used in this experiment had no effect on the specificity of emotional inferences made during reading. Indeed, the results showed that, even though there was a significant difference between reading times for sentences containing Mismatching emotions and sentences containing matching emotions, there was no difference among the three matching conditions. Once more, the emotional information

inferred was too general to determine specific emotions. This was the case even though the story manipulation should have compelled the readers to infer emotional information.

### General discussion

Gygax et al. (2003) challenged the assumption that readers infer specific emotions. Their hypothesis was that people do not infer specific emotions while reading, but instead infer a more general feeling, composed of different emotional components shared by several emotional terms.

In the first two experiments of this paper, it was hypothesised that if readers are provided with longer stories conveying sufficient information pertinent to the main character's emotional response, their emotional inferences should be specific. More specifically, it was hypothesised that creating longer stories would lead to specific emotions for two main reasons. First, the information conveyed in the (original) short stories was not sufficient for the readers to recognise the necessary components that lead to specific emotions. Second, longer stories should enable the readers to run a more comprehensive simulation of the situation. This simulation refers to the process by which readers identify with the story characters. By becoming more engaged in the stories and thus identifying with the characters, readers embrace their views, actions and plans. It is believed that the understanding of an emotion is enhanced, as the context facilitates the process of identification. Some authors (e.g. Oatley, 1999b) suggest that, when provided with an appropriate context, readers can experience and understand emotions more clearly from text than in real life. The stories used in these two experiments were twice as long as the stories used in Gygax et al. (2003). In Experiment 1, which was off-line, the hypothesis that longer stories lead to specific inferences was supported. In this experiment, participants presented with the long versions of the stories showed a higher consensus towards the specific emotions associated with the stories by Gernsbacher et al.'s (1992) than those presented with the short

versions (used by Gygax et al., and by Gernsbacher et al.). In Experiment 2, which used a self-paced reading task, there was no significant difference in reading times between sentences containing Matching, Matching Synonym or Matching Similar target emotion words. However, there was a difference in reading times between sentences containing Mismatching and matching target emotion words. This result suggests that people infer emotions while reading, but even when the stories are made longer and contain more information relevant to the character's emotional state, the emotional inferences are not specific.

In Experiments 3 and 4, the stories were further modified to compel readers to draw emotional inferences. It was hypothesised that if readers need to infer emotional information to establish coherence, their representation of the main character's emotional response should be specific. The stories were modified, so as to convey an initial ambiguity that could be resolved by inferring the main character's emotional response. In Experiment 3, which was off-line, the participants chose Gernsbacher et al.'s (1992) original emotions as being consistent with the character's emotional state more often in the ambiguous versions of the stories than in the original (short) versions of the stories. This result implied that readers infer specific emotions when presented with stories that compel them to infer the main character's emotional response. This result, however, was not reflected in the findings of Experiment 4, which used a self-paced reading task.

In summary, the experiments presented in this paper suggest that people do not infer specific emotions during reading. If, however, the participants have enough time to reflect on the main character's emotional response, as they do in off-line experiments, they do infer specific emotions. This finding was obtained both when the participants were presented with longer stories (Experiment 1) or when the stories compelled them to infer the main character's emotional response (Experiment 3).

One of the main assumptions of this paper is that emotions are composed of different components. When a situation is appraised, these components are selected and evaluated. Combinations of these components represent particular emotions (O'Rorke & Ortony, 1994). For example, the emotion sadness has a low-novelty level component, as the situations leading to the emotional reaction of sadness are often expected. Its valence component is negative, and its intensity is relatively mild. Another emotional component is the physiological response. When somebody feels sad they might cry. A stronger intensity might lead to the emotion depression (which could be considered as a mood when it is appraised as an enduring emotional state). The identification of these components, though an important aspect of emotion research, is not the focus of this paper. The main assumption of Gygax et al., which is supported by the results presented in this paper, was that readers do not infer specific emotions, but they infer several emotion components. These components are shared by similar emotions, and might themselves be specific. For example, the component of valence has two specific values, positive and negative. In one sense, if readers infer the valence of the situation, they have made a specific inference, though not an inference to a specific emotion.

Although the interpretation presented so far represents the authors' favoured account of the results, an alternative interpretation could be considered. Readers might infer specific emotions, but different readers might infer different specific emotions. In this sense, readers might have integrated a specific emotion in their mental representation of a story. Differences between readers' (in reading ability, emotional intelligence, etc.) might lead to the generation of different specific emotions, which in turn would result no reading time difference between the conditions. Even though this explanation seems plausible, an examination of the reading time distributions do not support it.

In conclusion, the experiments presented in this paper suggest that the combination of the different emotional components inferred by readers does not lead, at least during the reading of the passages, to inferences about specific emotions. Two manipulations, that did allow the inference of more specific emotions when people had as much time as they needed to make a judgement, did not produce specific emotional inferences during reading. However, it is worth reiterating that more than just the component of valence, i.e. whether the emotion is positive or negative, is inferred during reading, as the results of Gernsbacher et al. (1992) showed. Therefore, the emotional information inferred by readers is not composed of the components necessary to lead to inferences about specific emotions, but it is not merely composed of the valence component either.

One possibility, which could provide the focus for future research, would be to examine the semantic elements needed for the readers to move from a superficial representation of the main character's emotional response to a more specific one. Such research should first attempt to identify these elements and then assess their impact on readers' mental representation of the main character's emotional responses.

## References

- Alvarado, N. (1998). A reconsideration of the structure of the emotion lexicon. Motivation and Emotion, *22*, 329-344.
- Clore, G. L., & Ortony, A. (1988). The semantics of the affective lexicon. In V. Hamilton & G. H. Bower & N. H. Frijda (Eds.), Cognitive perspectives on emotion and motivation.
- DeVega, M., Diaz, J. M., & Leon, I. (1997). To know or not to know: comprehending protagonists' beliefs and their emotional consequences. Discourse Processes, *23*, 169-192.

DeVega, M., Leon, I., & Diaz, J. M. (1996). The representation of changing emotions in reading comprehension. Cognition and Emotion, 10, 303-321.

Ellsworth, P. C., & Scherer, K. R. (in press). Appraisal processes in Emotion. In R. J. Davidson & H. H. Goldsmith & K. R. Scherer (Eds.), Handbook of the affective sciences. New York and Oxford: Oxford University Press.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. W. (1992). Do readers mentally represent characters' emotional states? Cognition and Emotion, 6, 89-111.

Gernsbacher, M. A., Hallada, B.M. and Robertson, R.R.W. (1998). How automatically do readers infer fictional characters' emotional states? Scientific Studies of Reading, 2, 271-300.

Gernsbacher, M. A., & Robertson, R. W. (1992). Knowledge activation versus sentence mapping when representing fictional characters' emotional states. Language and Cognitive Processes, 7, 353-371.

Graesser, A. C., Singer, M., & Tabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101, 371-395.

Gygax, P., Oakhill, J., & Garnham, A. (2003). The representation of characters' emotional responses: Do readers infer specific emotions? Cognition and Emotion.

Miall, D. S. (1989). Beyond the schema given: Affective comprehension of literary narratives. Cognition and Emotion, 3, 55-78.

Norris, D. (1984). A computer-based programmable tachistoscope for nonprogrammers. Behavior Research Methods Instrument & Computers, 16, 25-27.

Oatley, K. (1994). A taxonomy of the emotions of literary response and a theory of identification in fictional narrative. Poetics, 23, 53-74.

Oatley, K. (1999a). Meetings of minds: Dialogue, sympathy, and identification, in reading fiction. Poetics, 26, 439-454.

Oatley, K. (1999b). Why fiction is twice as true as fact: Fiction as cognitive and emotional simulation. Review of General Psychology, *3*, 101-117.

O'Rourke, P., & Ortony, A. (1994). Explaining emotion. Cognitive Science, *18*, 283-323.

Ortony, A., & Clore, G. L. (1989). Emotions, moods and conscious awareness. Cognition and Emotion, *3*, 125-137.

Ortony, A., & Turner, T. J. (1990). What's basic about basic emotions? Psychological Review, *97*, 315-329.

Scherer, K. R. (1997). The role of culture in emotion-antecedent appraisal. Journal of Personality and Social Psychology, *73*, 902-922.

Smith, C. A., & Lazarus, R. S. (1993). Appraisal components, core relational themes, and the emotions. Cognition and Emotion, *7*, 233-269.

Storm, C., Storm, T., & Jones, C. (1996). Aspects of meaning in words related to happiness. Cognition and Emotion, *10*, 279-302.

Ze'ev, A. (2001). The subtlety of emotions. PSYCOLOQUY, *12*(007).



## Appendix A

Story where the consensus in Experiment 1 was clearly on the Matching Synonym emotion

---

Story:

The man was lying face down, probably unconscious, on the busy pavement. Other men and women bustled by on their way to work. Mark, who was late again, almost tripped over the man. "Why doesn't someone move this guy so people can get through", Mark yelled. He grabbed the man by his jumper and pulled him to the side of the path. While pulling, Mark tore the man's jumper. The man was still unconscious. Mark didn't take any notice either of the jumper or of the man's condition. He just left him, on the side of the pavement, unconscious. When Mark left, he jabbed the man with his foot and then continued on his way.

Conditions:

Matching: callous

Matching synonym: stressed

Matching similar: angry

Mismatching: caring

---

## Appendix B

Stories where the consensus in Experiment 3 was clearly on the Matching Synonym emotionStory13:

Don and his wife moved in a year ago. Tonight, Don was at home. He was in the living room, sitting on the floor. "How many things like this can happen in one day?" Don had asked himself. First, he'd been beaten out of a new job by a younger man. If that hadn't been enough, on the way home, he'd wrecked his car. Then, when he'd got home, he'd found out his wife wanted a divorce. She had already started to pack her things. When his wife had left, Don had slowly walked through the empty flat. He'd sat on the floor and thought about his life. He'd never had much luck in his life. Don hadn't wanted to do anything.

Conditions:

Matching: depressed

Matching synonym: miserable

Matching similar: useless

Mismatching: happy

Story23:

Tracy and Patty were having a conversation in the dorms, at University. They were alone in the room. At the end of the conversation, Tracy, who didn't know what to say, left the room. Tracy and Patty had been sleeping in the same dorm. Tracy now considered Patty to be an ex-friend. She had trusted Patty with her deepest, most private secrets, and it had seemed that everyone in the dorm knew of them. Tracy had confronted Patty with her suspicions. "But they were just too funny to keep secret", Patty had replied. Tracy had wanted to slap Patty in the face. What Patty had done was wrong. Tracy really had not wanted anybody else to know. She had gone to Patty thinking that she was somebody she could trust. Tracy hadn't known

what to do. Patty had even been laughing. "Tracy, you probably don't realise how nerdy you are", Patty had said.

Conditions:

Matching: angry

Matching synonym: furious

Matching similar: sad

Mismatching: grateful

---

Author Note

Pascal Gygax, Alan Garnham and Jane Oakhill, Laboratory of Experimental Psychology. Pascal Gygax is now at the University of Fribourg, Switzerland (pascal.gygax@unifr.ch).

This research was supported by an EPSRC research studentship to the first author. Correspondence and reprint requests should be addressed to Alan Garnham, Laboratory of Experimental Psychology, School of Biological Sciences, University of Sussex, Brighton, BN1 9QG, UK. Email: [alang@biols.susx.ac.uk](mailto:alang@biols.susx.ac.uk).

Table 1

Example of a story used in Gygax et al. (2003) with the emotion words in the four conditions

---

Story:

“How many things like this can happen in one day?” Don asked himself. First, he was beaten out of a new job by a younger man. If that wasn't enough, on the way home, he wrecked his car. Then, when he got home, he found out his wife wanted a divorce. All he could do was sit in his living room and stare into space. It was a moment when [Don] felt really [emotion]

Conditions:

Matching: depressed

Matching synonym: miserable

Matching similar: useless

Mismatching: happy

---

Table 2

Example of an extended story used in Experiment 1 & 2Short version of the story:

For Trevor, this had to be the best week of his 18-year life. Tonight he would be graduating first in his high school class. Just yesterday he received a formal acceptance letter from Harvard. And he had just hung up the phone after talking with someone very special who had said that she'd go with him to the graduation party.

Extended version of the story

It was Friday afternoon. For Trevor, this had to be the best week of his 18-year life. Tonight he would be graduating first in his high school class. Just yesterday he received a formal acceptance letter from Harvard. Harvard had been Trevor's first choice. He had worked very hard for it, and finally he'd been accepted. The acceptance letter also mentioned that he could apply for a scholarship. That scholarship would help him a lot. And he had just hung up the phone after talking with someone very special who had said that she'd go with him to the graduation party.

Conditions:

Matching: happy

Matching synonym: pleased

Matching similar: proud

Mismatching: depressed

Table 3

Means\*, percentages and standard deviations of the number of times the Matching emotion was chosen in the Short and the Long versions of the stories.

		Mean	Standard Deviation
Short versions	Matching	<b>6.38 (49%)</b>	<b>3.13</b>
	Matching Synonym	3.63 (28%)	2.18
	Matching Similar	1.75 (13%)	1.87
	Other	1.25 (10%)	1.78
Long versions	Matching	<b>8.25 (63%)</b>	<b>2.67</b>
	Matching Synonym	2.62 (20%)	2.30
	Matching Similar	1.42 (11%)	1.84
	Other	0.71 (5%)	1.04

\*Maximum value = 13

Table 4.

Example of a story used in Experiment 2 with the emotion words in the four conditions and the four target carrier sentences.

---

Story:

"How many things like this can happen in one day?" Don asked himself. First, he was beaten out of a new job by a younger man. If that wasn't enough, on the way home, he wrecked his car. Then, when he got home, he found out his wife wanted a divorce. She had already started to pack her things. When his wife left, Don slowly walked through the empty flat. He sat on the floor and thought about his life. He'd never had much luck in his life. Don didn't want to do anything. All he could do was sit in his living room and stare into space.

Conditions:

Matching: depressed

Matching Synonym: miserable

Matching similar: useless

Mismatching: happy

Target carrier sentences:

It was a moment when [character] felt really [emotion]

[character] could not believe how [emotion] s/he felt

There could be no doubt that [character] felt[emotion]

It was not surprising that [character] felt[emotion]

---



Table 5

Mean Reading Times (in ms) and Standard Deviations in Experiment 2.

Condition	Mean	Standard Deviation
Matching	1770	328
Matching Synonym	1809	301
Matching Similar	1805	331
Mismatching	2208	461

Table 6

Example of a story where the emotion is illustrated by several sentences leading to a final reinforcing statement (in italics) along with its modified version.

---

Story used in Gygax et al. focused on the emotion of sad:

Pam had just returned from her regular Tuesday visit to the nursing home. Today, there had been several problems. One elderly patient had died. Another had fallen and broken her hip. And all the faces had looked wrinkled, withered, and neglected. The sheer magnitude of the problems simply overcame Pam. When she had wanted to see the patient with the broken hip, she was told that the patient had been transferred to the city hospital. The injury was more serious than initially thought. Pam entered the empty room. She knew the patient well. Pam used to bring that patient fresh flowers every Tuesday morning. She always enjoyed having a chat with her. *Now Pam was sitting on the patient's bed. A tear ran slowly down her cheek.*

Modified version of the story (the ambiguous part is in italics):

*Pam was working for an organisation helping the elderly. The organisation was in close contact with different nursing homes. Today, Pam was visiting one of the nursing homes. She was sitting on an empty bed. For a while, she looked at her hands.* It was Pam's regular Tuesday visit to the nursing home. Today, there had been several problems. One elderly patient had died. Another had fallen and broken her hip. And all the faces had looked wrinkled, withered, and neglected. The sheer magnitude of the problems had simply overcome Pam. When she had wanted to see the patient with the broken hip, she was told that the patient had been transferred to the city hospital. The injury was more serious than initially thought. Pam had entered the empty room. She knew the patient well. Pam used to bring that patient fresh flowers every Tuesday morning. She always enjoyed having a chat with her.

---

Table 7

Mean\* and standard deviation of the ambiguity and emotional content responses to the passages in the final stage of the pilot.

	Mean	Standard Deviation
Ambiguity	5.54	1.82
Emotional content	3.92	1.75

\*Maximum value = 7

Table 8

Means\*, percentages and standard deviations of the number of times the Matching emotion was chosen in the Short and the Long versions of the stories.

		Mean	Standard Deviation
Short versions	Matching	<b>6.38 (49%)</b>	<b>3.13</b>
	Matching Synonym	3.63	2.18
	Matching Similar	1.75	1.87
	Other	1.25	1.78
Long versions	Matching	<b>8.04 (62%)</b>	<b>2.27</b>
	Matching Synonym	3.00 (23%)	1.98
	Matching Similar	1.25 (10%)	1.29
	Other	0.7 (5%)	1.27

\*Maximum value = 13

Table 9

Mean Reading Times (in ms) and Standard Deviations in Experiment 4

Condition	Mean	Standard Deviation
Matching	1920	389
Matching Synonym	1883	333
Matching Similar	1815	488
Mismatching	2496	685

Figure 1

The uni-modal distributions in the three matching conditions in Experiment 4

