Running head: EMOTION INFERENCES DURING READING COMPREHENSION

Emotion inferences during reading comprehension: What evidence can the self-pace reading paradigm provide?

Pascal Gygax

University of Fribourg, Switzerland

Isabelle Tapiero

University of Lyon II, France

Emanuelle Carruzzo

University of Fribourg, Switzerland

****PREPRINT****

Address for correspondence:

Pascal Gygax

Département de Psychologie

Université de Fribourg

Rue Faucigny 2

1700 Fribourg, Switzerland

Pascal.Gygax@unifr.ch

Abstract

This paper provides an explanation for the non-specificity of emotion inferences found in previous research (e.g., Gygax, Garnham & Oakhill, 2004). We first demonstrate that behavioral components of emotions, as opposed to emotions per se, are better markers of readers' mental representations of the main character's affective status. We also suggest that in a self-paced reading paradigm, when participants read sentences slower than others, it does not unequivocally provide insight into their mental representations of the text. We show that specific control conditions need to be implemented before such an assumption can be made in order to separate inference and representation processes from contextual integration processes.

Key words: text comprehension, emotion inferences, self-paced reading paradigm

Emotion inferences during reading comprehension: What evidence can the self-pace reading paradigm provide?

The investigation of characters' emotion inferences has emerged in the last thirteen years as an important research area in the field of text comprehension (e.g. Gernsbacher, Goldsmith, & Robertson, 1992; de Vega, Diaz, & Leon, 1997). Theories of text comprehension contend that as readers process text, they form a mental representation of the text (van den Broek, Young, Tzeng, & Linderholm, 1998; Graesser, Singer & Trabasso, 1994). This mental representation includes information pertaining to the people, settings, actions and events either described explicitly or implied by the text (Garnham, 1996). This particular idea is partly rooted in the notion of *mental models*, as advanced by Johnson-Laird (1983). Of prime concern is the fact that when reading a text, we cannot compute all the information presented to us, mainly because of processing limitations. We therefore construct a model of the situation, what can be referred to as a *state of the world* (Garnham & Oakhill, 1994), based on some elements presented to us and based on information stored in long-term memory. Mental models therefore comprise elements *of* the text and elements *derived (or inferred) from* the text as the result of interpretation mediated by long-term memory information. To form a mental representation of the text, readers combine different sources of information (van den Broek et al., 1998; Graesser et al., 1994). To accomplish this, readers activate previously acquired information that is stored in long-term memory, and combine it with information explicitly mentioned in the text (Kintsch, 1988; Gernsbacher, 1997). This means that people go beyond mere linguistic processes (i.e., just reading the words) when comprehending text (Martins, & Le Bouedec, 1998) to construct a relatively complex representation of the text.

The apparent complexity of readers' mental representations has raised several questions, mostly focused on which implied information, or in other words, which inferences, are automatically included in readers' mental representations *during reading* and, naturally, on how to investigate them. In this paper, we are interested in one type of inference:

inferences of emotion. We also examine the *self-paced reading* paradigm.

Inferences of emotions

Several theories have been advanced to account for research focused on the automatic generation of inferences. Only two core theories of text comprehension are presented here, the *constructionist* and the *minimalist* approaches, as both make (fairly) clear, yet different, predictions about emotion inferences, and about their automatic inclusion in readers' mental representations. The core principle underlying the constructionist approach is one of *search (or effort)-after-meaning*. This principle suggests that people naturally attempt to construct meaning from texts, social interactions and perceptual input (Bartlett, 1932, cited in Graesser et al., 1994). A central issue underlying the *search-after-meaning* approach is that of coherence. According to the constructionist approach, readers attempt to construct a mental representation of the text that enables them to maintain both *local* and *global* coherence. Inferences that establish local coherence essentially connect adjacent text constituents, whereas inferences that establish global coherence are deeper features such as general themes, or the moral of a text (Graesser et al., 1994), that connects most constituents of the text together. The constructionist approach considers emotion inferences as general features that guide readers' understanding of the text, thus enabling them to maintain a global coherence. To quote Graesser et al. (1994, p. 382) : "The emotional reactions of characters and the subordinate goals play a prominent role in global plot configurations of stories and are therefore needed for the establishment of global coherence". Therefore, the constructionist approach views emotions as being inferred during reading (p.382): "…the only causal consequences that are generated on-line are…(b) emotional reactions of characters to events and actions…".

According to the minimalist approach of reading, if readers do not adopt a specific goal-directed strategy; the number of inferences generated during reading is limited (McKoon and Ratcliff, 1992). The minimalist approach stipulates that only those inferences needed for

local coherence or those based on information that is readily available are generated during reading (McKoon & Ratcliff, 1992). Such an approach is a little ambiguous regarding emotion inferences. Indeed, if one considers emotion inferences as crucial *only* for global coherence, the minimalist approach would imply that they are not inferred during reading. However, if emotion inferences are considered as necessary for local coherence, the minimalist position would probably predict that they are generated during reading. In essence, the automatic processing of emotion inferences according to the minimalist position largely depends upon coherence issues. In addition to this, one can also assume that, as readers search through their own experiences to understand emotions (Miall, 1989) and as we daily use emotions as a communicative tool (Oatley, 1998), emotional inferences might be based on information easily retrievable from long-term memory. Although we do believe that emotions are easily retrievable from long-term memory, it is likely that emotions are mainly needed for global coherence (as stated by Graesser et al., 1994) and hence we take the minimalist approach as being contrary to the notion that emotions are automatically inferred during reading.

Contemporary research on emotion inferences can be traced to a study by Gernsbacher et al. (1992). They investigated the ability of readers to represent the situation described in a text, and especially their ability to represent the emotional status of the (main) characters. In their three experiments, they used two different paradigms: the first two experiments used the self-paced reading and the third experiment used a naming task. In this paper, we employed the self-paced reading paradigm to investigate emotion inferences during text comprehension. The assumption underlying such a task is that sentences that are congruent with readers' current mental model of the situation should be read faster than sentences that are incongruent. Thus, reading time mirrors the ease with which a sentence is incorporated into the situation model. In their first experiment, Gernsbacher et al. (1992) compared reading times of target sentences in short stories containing different emotion terms. For each story,

they first matched pairs of emotions that shared intensity, duration and relevance to self (Frijda, 1986) but that had opposite affective valence (e.g., Guilty-Proud, Bored-Curious, Sad-Joyful, Shy-Confident). They therefore compared reading times of target sentences containing either *matching* or *mismatching* target emotion words. Across the experiment, each emotion term was once tested as a matching emotion and once as a mismatching emotion. Gernsbacher et al. (1992) found that sentences were read significantly faster when they contained matching emotion terms than when they contained mismatching emotion terms. In their second experiment, Gernsbacher et al. (1992) investigated whether emotional inferences were valence driven (i.e. do readers merely incorporate "global outcomes, for instance, whether the outcome is negative or positive" (p.100)). They compared reading times of target sentences containing matching (e.g., bored) and mismatching emotion words (e.g., angry) that had the same valence. The difference between reading times of sentences containing the matching and the mismatching emotion terms was smaller in this experiment than in the first experiment, but it was still significant, implying that emotional inferences were not mainly driven by valence. Gernsbacher et al. (1992) stressed the importance of the mismatching condition as a marker of readers' mental representations. They stated that "the more the emotions words conflict with the implied emotional state, the more slowly their target sentences should be read". The notion of slow reading processes as markers of readers' mental representations is closely investigated in the present paper.

From those experiments, Gernsbacher et al. (1992) suggested that readers do not integrate merely valence in their mental representation (i.e. the main character is feeling either positive or negative), but rather integrate specific emotions (i.e. sad or happy). Such specificity has been suggested not only in research on emotion inferences (e.g. de Vega, Leon, & Diaz, 1996; de Vega, Diaz, & Leon, 1997; Gernsbacher, Goldsmith, & Robertson, 1992; Gernsbacher, & Robertson, 1992; Gernsbacher, Hallada, & Robertson, 1998) but also in research on characters' trait inferences (Rapp, Gerrig, & Prentice, 2001), which indicated that

readers form a rather specific trait-based model of narrative protagonists, without particular effort.

The notion of specificity was questioned by Gygax, Oakhill, and Garnham (2003) and Gygax, Garnham, and Oakhill (2004). In Gygax et al. (2003), using stories based on Gernsbacher et al. (1992), participants read target sentences containing either the initial *matching* emotion (e.g., *sad* from Gernsbacher et al., 1992), a *matching synonym* emotion (e.g., *depressed*), a *matching similar* emotion (e.g., *useless*) or a *mismatching* emotion (e.g., *happy*). Matching emotions were always read faster than the mismatching emotion, but there was no difference in reading times between sentences containing the different matching emotions. Gygax et al. (2004) extended these findings to stories that were made either longer (Experiments 1 & 2) or more ambiguous (Experiments 3 & 4). Both story manipulations were intended to make participants focus or concentrate on specific emotions. In the longer versions, Gygax et al. (2004) intended to provide sufficient information for readers to reach a specific representation of the main character's emotional status. In the ambiguous versions, they hoped to force participants into more effortful processing of the main character's emotional status, which was needed to resolve the ambiguity presented in the stories. Two different paradigms were used in Gygax et al. (2004): a forced-choice task and a self-paced reading task. In the forced-choice task experiments (Experiments 1 & 3), participants were explicitly asked to choose for each story the most appropriate emotion that described the main character's emotional status among several emotions (*initial*, *synonym, similar* or *mismatching*). In the self-paced reading task experiments (Experiments 2 & 4), participants' reading time for sentences containing different emotions were measured. For both story manipulations, when forced to make a choice among initial, synonym, similar or mismatching emotions, participants were more likely to select the initial emotion than the others. However, the self-pace reading task did not differentiate between the three matching emotions. The results of the self-paced reading experiments mirrored the ones found in Gygax et al. (2003),

leading the authors to conclude that readers do not represent specific emotions, at least not during reading (i.e., automatically). However, Gygax and colleagues suggested but never truly explained why readers' mental representations of emotion would be non-specific. In this paper, we attempt to give an explanation to account for the non-specificity of the mental representation of the main character's emotional status based on both theoretical and empirical foundations.

The core theoretical idea behind our explanation of the non-specificity of emotion inferences is that readers include some information in their mental representations that is shared by several emotions. We suggest that readers include some *physical behavioral response* to their mental representations of the situations. By *physical behavioral response*, we refer to any movement, or lack of movement in response to the situation. For example, someone might *clinch their fist* in response to a frustrating situation or *freeze* in response to a scary situation. Essentially, we refer to such information as *what the main character physically does* in response to the situation. We believe that such a response to a situation is relatively easy to infer, at least easier than a fairly complex and abstract emotion representation. We further suggest that such information is connected to what we refer to as *an emotion construct*. In essence, readers construct a representation foundation, formed of stereotypical, or common information (i.e., behavioral reaction). As reading progresses, readers might elaborate that foundation to reach a more complex representation (i.e., specific emotions). In terms of Gernsbacher's (e.g., 1997) Structure Building framework, one could say that the structure foundation is composed of some behavioral elements, and that new information is mapped onto that developing structure.

Such an idea somehow distances us from the too simplistic constructionist vs. minimalist dichotomy (an inference *is* or *is not* drawn). If readers infer some behavioral components, which, as will be explained in detail later, can be considered as part of an emotion construct, one can argue that emotion inferences are neither specific (as consistent

with the constructionist approach) nor absent (as consistent with the minimalist approach). Such an argument follows the idea that inferences are not encoded in an all-or-none fashion (van den Broek et al., 1998; Graesser et al., 1994; McKoon & Ratliff, 1990, 1992). This means that the emotion construct attached to the main protagonist is not entirely built and represented. The specificity and strength of the representation of a particular inference varies from situation to situation (van den Broek et al, 1998), and can change as new information is provided by the text. We further propose that the comprehension of a text does not need a complex and cognitively demanding representation of the main characters' emotional status, and that semantic parts of the complex representation may well be sufficient to assist reading comprehension.

The self-paced reading paradigm

As previously discussed, Gygax et al. (2004) suggested that readers make specific emotion inferences only when forced to (i.e., in a forced-choice task). In their self-paced reading task, participants were equally fast at reading sentences containing different matching emotions (e.g., sad, depressed and useless), but were always slower at reading mismatching emotions (e.g., happy). The *match/mismatch* effect was interpreted as evidence that readers do make emotion inferences, but the *same match* effect was interpreted as evidence that these inferences are not specific. Even though such a match/mismatch effect has often been interpreted as an indication of inference processing in emotion inference studies, researchers may have been a little hasty in their interpretation. More specifically, we believe that the studies that have used the self-paced reading paradigm have not been conducted in a way that would allow to distinguish between two alternative interpretations for faster reading of emotion terms. Faster reading times of matching information relative to mismatching information could be because it either matches inferences that have been already incorporated into readers' mental model and therefore activates a previously made inference (i.e., inference activation), or it is simply consistent (or at least not inconsistent) with but does not match

information already included in the representation. The second alternative reflects what we refer to as *context integration* as compared to actual *inference activation*. What we mean by context integration is that information not previously incorporated in a mental representation might be read at a seemingly fast pace merely due to the fact that the information is not directly opposed to the context (e.g. somebody eating sweets when the it's raining). The term integration is important here, as it mirrors the process by which new information may be integrated in readers' mental representations. The two main questions therefore, at least for our purpose, are essentially (a) whether a match/mismatch effect, found in all previous studies on emotion inferences, does necessarily mean that the matching information has been included in readers' mental representations and (b) whether the self-paced reading time paradigm can differentiate between inference activation and context integration.

In the present experiment, we added a *context integration* manipulation to contrast it with our *inference processing* manipulation[2] so that we could evaluate variation in differences between matching and mismatching conditions. Basically, an inference processing match/mismatch effect should be greater than a mere context integration match/mismatch effect. That greater difference mirrors a closer association between the information in the inference processing manipulation and the information in readers' mental representations.

To investigate the *context integration vs. inference processing* issue, we added a control condition (using filler stories) in which we contrasted reading times of non-emotional sentences that were *neutral* to sentences that were *contextually irrelevant*. By *neutral* information, we refer to any sentence that can be easily integrated in the context. For example, the sentence *I take the train* is a neutral continuation of the sentence *The weather is nice*. Although the neutral sentence does not contradict the context established by the first sentence, it is very unlikely that it has been included in readers' mental representations as readers processed the first sentence. It is important to note here that we do not see *context integration* and *inference activation* as a dichotomy per se (see Figure 1). Figure 1 represents

a continuum of integration ease, echoed by sentence reading speed. The faster a sentence is read, the closer it resembles the information previously processed and incorporated in readers' mental representations. However, a fast reading time, as shown in the figure, can but does not unequivocally mean that the information has been previously incorporated in readers' mental representation; likewise, slow reading time does not necessarily mean that the information contradicts information included in readers' mental representations. As shown in Figure 1, *I take my bicycle* may be more easily integrated in readers' mental representation after *The weather is nice* than *I take the train*, but it does still not mean (although it's possible) that the information has been already incorporated in readers' mental model. When constructing the *neutral* sentences in the *context integration* manipulation, we made sure that the information tested had truly no chance of being incorporated in readers' mental representations[3] (in Figure 1, the information needed as a control condition is in bold).

By *contextually irrelevant*, we refer to any sentence that is difficult to integrate in the preceding context. For example, the sentence *I don't like public transport* is an incongruent continuation of the sentence *I take the train*. Given the right explanation (i.e., *My car is being repaired*), the information that *I take the train* might not be so difficult to integrate in the context (it might even be inferred) but without that information, the sentence should normally be difficult to integrate. Using such a control condition permitted us to differentiate between integration and inference processes. To be interpreted as an inference processing indication, a *match/mismatch* difference in an *emotion* condition, or in a *behavior* condition, according to our aforementioned hypothesis, should be different than a *match/mismatch* difference in a *neutral context integration* condition as defined above. A similar match/mismatch difference might imply that our target information is easily integrated in readers' mental representations, but not that it is inferred during reading.

Method

Participants

Twenty-four students from the University of Fribourg took part in this experiment as part of a course accreditation.

Materials

Experimental stories. Twenty-four stories from Gygax and Tapiero (2003) were used in this experiment. In Gygax and Tapiero (2003), thirty students were explicitly asked to write twenty-four stories each focused on specific given emotions. Gygax and Tapiero (2003), by instructing the participants to focus their stories on specific emotions tried to compel the participants to write sufficient context information to allow the inference of specific emotions. A categorical analysis of participants' responses enabled them to construct twenty-four stereotypical stories using the most recurrent categories. These stories were generated by the participants and therefore represented the most salient stories to the chosen emotions (see Table 1 for an example story).

Each story was present in four different versions, depending on the target sentence. The target sentence contained either a *matching emotion*, a *mismatching emotion*, a *matching behavioral description* or a *mismatching behavioral description.* Four lists were therefore constructed, to ensure that each participant would see all conditions and that each item was represented in each condition across the experiment. For each story, the matching behavior was matched with the matching emotion, and the mismatching behavior matched with the mismatching emotion.

Filler stories. Twenty-four filler stories, with the same structure as the experimental stories, but emotionally neutral were also constructed. Out of the twenty-four filler stories, twelve were used to test the *neutral match/mismatch* question mentioned earlier. Six of these stories had a matching second sentence and six had a mismatching second sentence. In the latter case, the rest of the story was written so as to clarify the ambiguity raised by the presence of mismatching information (see Table 2 for examples of filler stories). We used the second sentence as the target sentence to ensure that the *matching neutral* information was

unlikely to have been incorporated in readers' mental representations of the text (i.e. after just one sentence). A pilot study involving twenty participants ensured that the matching neutral sentences were indeed neutral and that the mismatching sentences were considered as mismatching the context. The participants in the pilot study were presented with two sentences and had to decide on a 7 pts scale whether the second sentence was an obvious continuation (score of 7), a possible continuation (score of 4) or an impossible continuation (score of 1). The results of the pilot study showed a significant difference between the sentences that we wanted to use as *neutral* (M = 4.28; SD = 0.6) and as *contextually irrelevant* (M = 1.55; SD = 0.22) (t (10) = -10.47; $p < .0001$). The filler stories were the same for each participant.

In total, forty-eight stories were presented to the participants: twenty-four experimental stories, twelve experimental filler stories and twelve normal filler stories. For each participant, these were presented in a random order.

Apparatus

The stories were presented on a Power Macintosh 4400 using Psyscope Software (Cohen, MacWhinney, Flatt, & Provost, 1993). Responses to the target sentences of each story (final sentence in the experimental stories and second sentence in the filler stories) were collected using a response button box attached to the computer, which permits millisecond accuracy. Note that before analysing the results, the reading times were transformed to account for sentence length, readers' particularities and position of the sentence within the text. The exact transformation is explained later.

Procedure

The participants were instructed to read the stories at a normal reading speed, as if they were reading a magazine. Each story was presented in three or four parts of one or two sentences and participants had to press the *yes* button when they finished reading each part. To make sure that participants read the stories carefully, some stories (N=16) were followed

by a question related to the text. Participants were asked to answer the questions by pressing a button labelled either *yes* or *no*. After the presentation of a story (including questions), the participants were prompted with the message *Are you ready?* followed by the story. Before the main part of the experiment, the participants read three practice stories, two of which were followed by a question, to familiarize them with the procedure and with the kinds of passages that they would be reading.

Results and discussion

The aim of this experiment was twofold. First, it was to demonstrate that behavioral information (as part of emotion) is a better marker of readers' mental representation than emotions per se. Second, it was to assess whether the usual *match/mismatch* effect could be considered as a true indication of readers' mental representations of the main protagonist's emotional status and not merely as a mirror of basic integration processes.

Before conducting the statistical analysis, we transformed the data principally to account for the fact that the target sentences were not all of the same length and were not at the same position in the text. The transformation procedure was inspired by Trueswell, Tanenhaus and Garnsey (1994) and was conducted as followed: For each participant and for each manipulation, we produced a regression equation of time (i.e. reading time) against length (i.e. number of characters in the target sentence) using all 2nd sentences for the *context integration* manipulation and all 5th sentences for the *inference processing* manipulation. We separated the 2nd and 5th sentences in the regression analysis to account for different reading times due to different positions in the texts. The actual calculation is fairly simple. For the *inference processing conditions*, the reading time of each story's 5th sentence was taken. The 5th sentences that were not manipulated (i.e., the 5th sentences in the filler stories) were also considered for the regression. One could view these filler 5th sentences as control sentences, or baseline sentences, incorporated in the regression transformation. Then, for each participant, a *time* by *sentence length* regression was calculated by computing the slope and

the intercept of the regression. Residual reading times for each participant were then calculated by subtracting the actual reading times from the reading times predicted by the regression equation. Statistical analyses were conducted on the residuals. Negative residual times mean that reading times were longer than expected. Although relatively complex, this data transformation was needed to address differences in item length and position, as well as variations between participants. In this sense, a baseline reading measure is not needed for each participant, as it is already accounted for by the participant's regression line. Mean residual times are shown in Figure 2.

A general 3 (Nature: Emotion vs. Behavior vs. Neutral) X 2 (Match: Matching vs. Mismatching) repeated measures ANOVA showed a main effect of *Match* (F_1 (1, 23) = 29.76 and F_2 (1,23) = 24.72; p < .001[4]), but most importantly, it showed an interaction effect for *Match* X *Nature* (F_1 (2, 46) = 7.85; p < .001 and F_2 (2,46) = 4.22; p < .05). We explored the source of the interaction in separate follow-up ANOVAs that addressed the two main questions of this research. In the first follow-up we looked at whether emotion and behavior conditions produced different patterns in the matching and mismatching conditions. For this purpose, we examined the interaction of Match with respect to two of the three values of the Nature variable. Thus we conducted a 2 (Nature: Emotion vs. Behavior) X 2 (Match: Matching vs. Mismatching) repeated measure ANOVA. The ANOVA indicated a main effect of *Match* (F_1 (1, 23) = 25.66; p < .001 and F_2 (1,23) = 22.70; p < .01) and an interaction effect (F_1 (1, 23) = 8.68; p < .01 and F_2 (2,46) = 4.48; p < .05). As the means in Figure 2 show, the *match/mismatch* difference was the bigger in the behavior condition than in the emotion condition, implying that the behavior information is the closest information to the content of the participants' mental representations. Such results support our first hypothesis, namely that behavioral information is a better marker of readers' mental representations of the main character's emotional status than emotions per se. This follows the idea that, as suggested by Gygax et al. (2004), readers might only infer part(s) (i.e., behavioral components) of a

complex emotion construct in their mental representations. In essence, readers might attribute some intuitive or subjective feeling(s) to the character as they read through the text, but as this or these feelings are very difficult to incorporate into their mental representations, they include only some components of their construct in their representations. We do not suggest that readers are unable to infer specific and complex emotions, but as a result of limited processing capacity, we believe that emotions might be too complex to be fully integrated in readers' mental representations in a normal reading process. If one were to see the construction of emotion inferences as a sequential process, we believe that given sufficient time, readers could infer more components of emotions, resulting in a more accurate representation of the main character's emotional status. An alternative, yet related explanation of our *behavior* vs. *emotion information* could be rooted in the idea, introduced earlier, that readers build a structure foundation composed of different explicit and implicit elements, such as behaviors. As readers gradually accumulate these elements, they gather adequate information to encode a complex or specific emotion representation. Essentially, readers may need several particular elements to generate specific emotion representations. Gernsbacher et al. (1992) hinted that *valence* was one of the element, and we believe that *behavioral response* is another. Gygax and Tapiero (2003) suggested that other important elements to the generation of emotion representation include, for example, internal *physical states* (e.g., heart beating fast) or *appraisal processes* (e.g., *she thought it was difficult*).

Two additional follow-up ANOVAs were necessary to examine the inference processing versus contextual integration explanation of a slow down in the mismatch condition. In the second follow-up of the interaction of Match with the three-level Nature variable, we looked at inference processing versus contextual integration by analyzing the Emotion and Neutral levels of the Nature variable and the Match (Matching vs. Mismatching) variable in a repeated measure ANOVA. This analysis only showed a main effect of *Match* (F_1 (1, 23) = 28.85; p < .001 and F_2 (1,23) = 16.72; p < .001), but no interaction effect. As

suggested earlier, a similar match/mismatch difference between these two conditions might imply that emotions are easily integrated in readers' mental representations, but not that they were inferred during reading. In other words, because the mismatch cost for emotion words is similar to the cost for mismatch on neutral, we cannot distinguish between inference processes and context integration. The third follow-up ANOVA indicated that there was a significant interaction when behavior and neutral were examined with respect to the match/mismatch interaction. A 2 (Nature: Behavior vs. Neutral) X 2 (Match: Matching vs. Mismatching) repeated measure ANOVA showed a main effect of *Match* (F_1 (1, 23) = 23.96; $p < .001$ and F_2 (1,23) = 15.98; $p < .01$) as well as an interaction effect (F_1 (1, 23) = 11.65; $p < .01$ and F_2 (2,46) = 5.07; $p < .05$). As the means in Figure 2 show, the mismatch effect was significantly larger in the Behavior condition than in the Neutral condition. This implies that inference processes can be distinguished from context integration when the semantic element is a behavioral component. The pattern of data in the Nature x Match interaction indicate a significantly greater match/mismatch effect for behavior versus emotion and behavior versus neutral along with an equivalent match/mismatch effect for emotion versus neutral. This data pattern indicates that the self-paced reading paradigm *can* differentiate between context integration and inference processes but only for certain kinds of emotion-related semantic information, in this case behavioural information. For specific emotion words, self-paced reading did not distinguish between inference processes and contextual integration. These results are consistent with our proposals that emotion words might not be part of readers' representations even though a significant match/mismatch effect is observed in self-paced reading time data. Thus, we have demonstrated that a *match/mismatch* effect during self-paced reading does not necessarily mean that the matching information has been included in readers' mental representations of the text.

Our results on the behavior component of emotion confirm the idea that the constructionist vs. minimalist debate is not so informative here, as emotion inferences are

neither specific (constructionist position) nor entirely absent (our interpretation of the minimalist position). Instead, at least under normal conditions[5], readers might only infer some components of the semantic construct of emotions. In our experiment, we tested behavioral elements, assuming that the content of readers' mental representations is *closely related* to such information. Of course, our results do not necessarily imply that readers infer specific behaviors. By using a self-paced reading paradigm, we cannot differentiate between forward and backward processes (Keenan, Potts, Golding & Jennings, 1990). When readers encounter the target behavioral information, the time it takes them to read the sentence mirrors the ease by which that information can be mapped onto the information in their mental representation. If the information in readers' mental representations is exactly the same as the one presented in the target sentence, the latter should be easily mapped onto the mental representation, as shown in Figure 2, and thus should be faster to read. However, if the information in readers' mental representations *resembles* or is *closely related to* the one presented in the target sentence, it should also be read faster. Our point is that the bigger the *match/mismatch* difference, the closer that information is to the content of readers' mental representations. It could be argued that faster reading times merely mirror the ease by which sentences containing behaviors are interpreted, or integrated in the context. Consequently, we may only have addressed different *degrees* of context integration. If this interpretation seems plausible for interpreting faster reading times, it seems more laborious to apply it to slower reading times of incongruent information. In a sense, incongruent sentences containing behaviors should also be faster interpreted as being incongruent, if the aforementioned explanation held true. The data however do not seem to support such an explanation.

In this paper, we argue that a component *likely* to be inferred is the main character's behavioral reactions to the situation. This idea stems from the strong link between emotions and behaviors (Consedine, Strongman, & Magai, 2003). In addition to this, we suggest that anticipating the main character's physical reaction might enable readers to keep an open

representation of the main character's emotional experience. A less complex representation such as a *common behavior* might be easier to modify than a more complex one, such as *an emotion*. It is likely, in narratives, that the main character undergoes emotional changes as the story evolves. Therefore, readers might have to modify their representation to account for different situations eliciting emotional reactions. It would be more effective and less effortful for readers to keep a non-specific representation of the main character's emotional status. Any changes in the representation would thereafter be relatively fast and incur little cognitive cost.

In contrast to this latter explanation, it is possible that readers might simply not have sufficient time (or cognitive ability) to construct a complex representation of emotions. Readers may just maintain a sufficient representation, or a more concrete one maybe, for comprehension purposes. If our results suggested that only semantic constituents of emotion might be integrated in readers' mental representations, it is not to say that readers' mental representations could not, at some particular time in the reading process, include specific emotions. Such an idea follows work from Calvo and Castillo (1996) who demonstrated that elaborative inferences do take time to build. In terms of emotion inferences, this means that the elaboration of a complex representation of emotion may well follow a specific sequence of which the first step is to generate a behavioral component to be included in the representation of the situation[6].

At this point, we would like to relate Gygax and colleagues' results of non-specificity to our present results. In their experiments, they found that reading times of sentences including different matching emotions did not differ. For example, for a story in which a man lost his job, had a car accident and found out that his wife wanted a divorce, there was no reading time differences between sentences containing *depressed*, *miserable* or *useless*. There was no reading time differences because all three emotional states share some emotion component(s) that was (were) likely in readers' mental representation. We believe that as some common behaviors (e.g. *lower your shoulders* or *cry*) are tied in with several emotions,

it is not surprising that several emotions *seemed* to be equally inferred in Gygax et al. (2003) and Gygax et al. (2004). In fact, they all easily mapped onto the existing structure that encompassed some behavioral reaction.

As a final note, we would like to stress that although the experiment presented here hinted that behavioral elements might be integrated in readers' mental representations, we did not evaluate the life span of such elements, nor did we evaluate the influence that such information may play in processing subsequent information. Subsequent information might direct readers towards more specific or explicit emotions, resulting in a refinement of readers' mental representations.

References

Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*(2), 257-271.

Consedine, N. S., Strongman, K. T., & Magai, C. (2003). Emotions and behavior: Data from a cross-cultural recognition study. *Cognition and Emotion, 17*(6), 881-902.

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

de Vega, 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.

Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.

Garnham, A. (1992). Minimalism versus constructionism: A false dichotomy in theories of inference during reading. *Psycholoquy, 3*(63).

Garnham, A. (1996). The other side of mental models: Theories of language comprehension. In J. Oakhill & A. Garnham (Eds.), *Mental models in cognitive science: Essays in honour of phil johnson-laird* (pp. 35-52). Hove: Psychology Press.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. W. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion, 6*(2), 89-111.

Gernsbacher, M. A., & Robertson, R. W. (1992). Knowledge activation versus sentence mapping when representing fictional characters' emotional states. *Language and Cognitive Processes, 7*(3/4), 353-371.

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Processes, 23*, 265-304.

Gygax, P., Oakhill, J., & Garnham, A. (2003). The representation of characters' emotional responses: Do readers infer specific emotions? *Cognition and Emotion, 17*(3), 413-

428.

Gygax, P., Garnham, A., & Oakhill, J. (2004). Inferring characters' emotional states: Can readers infer specific emotions? *Language and Cognitive Processes, 19*(5), 613-638.

Gygax, P., & Tapiero, I. (2003). *Divide and conquer: A study of the semantic structure of emotional inferences in reading comprehension.* Paper presented at the Annual Conference of Swiss Psychology, Bern, Switzerland.

Graesser, A. C., Singer, M., & Tabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395.

Herrmann, T. (1982). Language and situation: The pars pro toto principle. In C. Frazer & K. Scherer (Eds.), *Advances in the social psychology of language* (pp.123-158). Cambridge, England: Cambridge University Press.

Keenan, J. M., Potts, G. R., Golding, J. M., & Jennings, T. M. (1990). Which elaborative inferences are drawn during reading? A question of methodologies. In D. A. Balota, F. d'Arcais & K. Rayner (Eds.), *Comprehension processes in reading*. Hove: LEA.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163-182.

McKoon, G., & Ratcliff, R. (1990). Dimensions of inference. In A. C. Graesser & G. H. Bower (Eds.), *Inferences and text comprehension*. London: Academic Press.

McKoon, G., & Ratcliff, R. (1992). Inferences during reading. *Psychological Review, 99*, 440-446.

Martins & Le Bouedec (1998). La production d'inférences lors de la comprehension de textes chez les adultes: Une analyse de la literature. *L'Année Psychologique, 98*, 511-543.

Miall, D. S. (1989). Beyond the schema given: Affective comprehension of literary narratives. *Cognition and Emotion, 3*(1), 55-78.

Oatley, K. (1998). Emotion. *The Psychologist, 11*(6), 285-288.

Rapp, D. N., Gerrig, R. J., & Prentice, D. A. (2001). Readers' trait-based models of characters

in narrative comprehension. *Journal of Memory and Language, 45*(4), 737-750.

Trabasso, T., van den broek, P., & Suh, S. Y. (1989). Logical Necessity and Transitivity of Causal Relations in Stories. *Discourse Processes, 12*(1), 1-25.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic Influences on Parsing - Use of Thematic Role Information in Syntactic Ambiguity Resolution. *Journal of Memory and Language, 33*(3), 285-318.

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1998). The landscape model of reading: Inferences and the on-line construction of a memoty representation. In S. R. Goldman & H. van Oostendorp (Eds.), *The construction of mental representation during reading* (pp. 71-98). Mahwah: N.J. Erlbaum.

Author Note

We would like to thank Julien-François Gerber and Isabelle Reverte for their work on the project and Ben Parris for his comments on the manuscript. We are also indebted to Susan Goldman for her suggestions and would like to thank the two reviewers for their constructive criticisms.

Footnotes

[1]We define such a concept as the behavioral action (e.g. movement or lack of movement) associated to emotions, a kind of physical attitude reaction. We do not consider this concept of behavioral action to be the same as physiological reaction (e.g. heart beat,..), which could be tested in future research.

[2]The terms *manipulations* and *conditions* are not to be confused here. In the *context integration manipulation*, we had to conditions, so did we in the *inference processing manipulation*. These conditions are explained in detail in the text.

[3]In the example given above, there might be occasions in which the sentence *I take my bicycle* will be automatically inferred by readers (good weather might mean good weather for a bicycle ride for some people).

[4]In order to perform a by Items analysis, we decided to compute each list as if they had different filler items. That way, we managed to get 24 items, facilitating the by-items analyses.

[5]*Under normal conditions* means that the participants were not asked to adopt particular reading strategies. We are currently testing the idea that specific strategies (simulating the characters' emotion for example) influence the complexity of the representation of emotions.

[6]We are currently testing the hypothesis that as the time between sentence presentation increases, giving more time for participants to compute a complex mental model, the match/mismatch difference increases in the emotion condition (more complex emotion representation) whereas the match/mismatch difference in the behavior condition decreases. Preliminary results seems to indicate that this is the case.

Table 1

Example of a story used in the experiment.

Story:
Suzanne came back from her regular visit to the nursing home. She walked slowly from the nursing home to her place. She thought of the days with her grandmother with a heavy heart. She had trouble retaining her tears when thinking of her grandmother alone in her room.
Target sentences:
Matching emotion: As you could expect, Suzanne was feeling sad.
Mismatching emotion : As you could expect, Suzanne was feeling happy.
Matching behavior: She sat on her settee, wrapped in a blanket.
Mismatching behavior: She danced all night, as she was always the one to show others how to party

Table 2

Examples of filler stories. In the first story, the second sentence is the matching neutral target sentence, and in the second story, the second sentence is the mismatching neutral target sentence .

Story 1: Neutral matching
Jean was writing and preparing for a conference in the East. *He wanted to make the most out of this trip*. He therefore planned several visits to his friends and to the people he knew on the way. He was taking his time to prepare for the trip and had been preparing since spring.
Story 2: Neutral mismatching
Cindy had just finished work and was going to her gym (*au fitness* in French). *She thought that going to the gym was individualistic and preferred volleyball*. But after a knee injury, she could not play volley-ball any more. After changing, she entered the gym which was empty and dark. “That’s strange” she thought as she was warming up.

Figures Caption

Figure 1

Conceptual representation of reading times in relation to the extent to which a sentence matches its preceding context (*context integration* vs. *inference processes*). In this paper, we suggest that the self-paced reading paradigm can be useful, but appropriate control conditions (normal integration processes) need to compared to experimental conditions (inference processing conditions). In the example, we suggest that reading a sentence such as *The whether is nice* is more likely to active some information such as *The sun is shining*, hence decreasing the time to read the information, than information such as *I take the train*. However, we do believe the boundaries are not precise, and the continuation sentences may fluctuate along the line, depending on readers and reading conditions.

Figure 2.

Mean residual times in all the conditions. Negative residual times mean slower times.

Figure 1

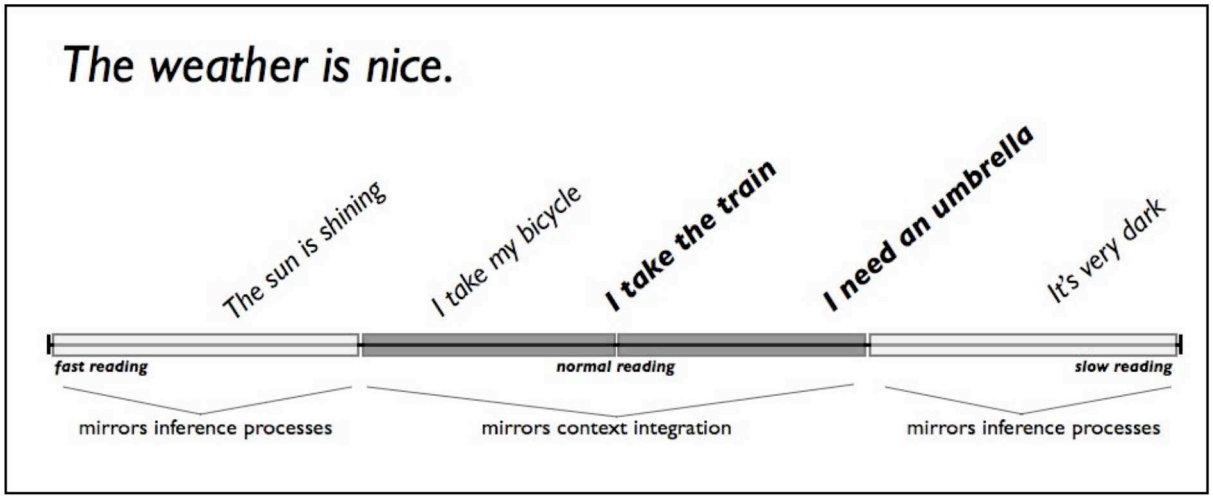
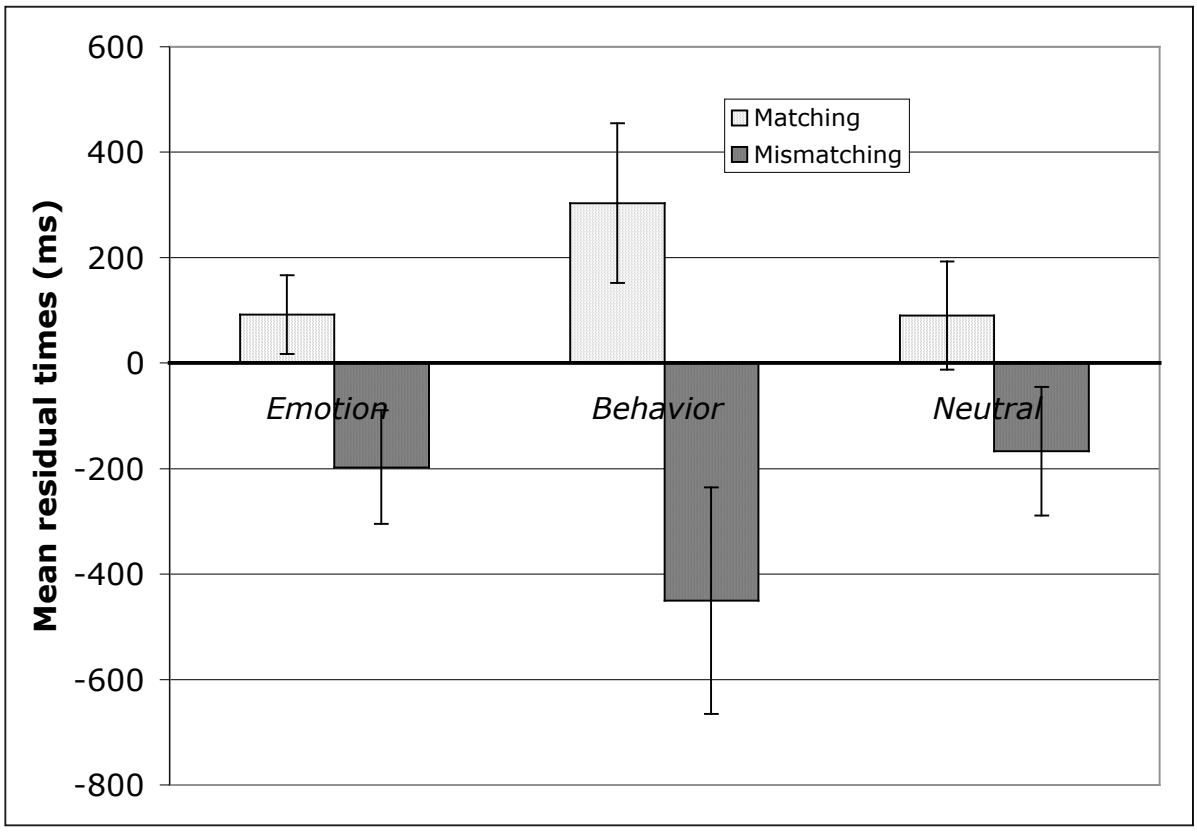

The weather is nice.
The sun is shining
I take my bicycle
I take the train
I need an umbrella
It's very dark
fast reading
normal reading
slow reading
mirrors inference processes
mirrors context integration
mirrors inference processes

Figure 2


600
400
200
0
-200
-400
-600
-800
Mean residual times (ms)
Matching
Mismatching
Emotion
Behavior
Neutral