

Expect the Unexpected? The Processing of Possibility Hedges in Medical Diagnoses and Medical Advice

Teeranoot Siriwittayakorn

English Department, Faculty of Humanities,
Chiang Mai University
239 Huay Kaew Road, Muang, Chiang Mai
50200 Thailand

teeranoot.s@cmu.ac.th

Abstract

Expectation is known to have an impact on behavioral choices of human beings by intensifying the perception of something or distracting their attention away. However, little is known about how unexpected situations affect information processing. This study investigated the processing of possibility hedges. Specifically, it tested whether participants ignored possibility hedges when processing medical diagnoses and medical advice that contradicted their expectation. The results suggested that possibility hedges used in sentences did not affect the degree of participants' surprise when reading unexpected text continuations. However, it was found that although participants paid attention to possibility hedges when asked to rate credibility of the texts, they paid attention to the hedges to a varying degree depending on whether situations mentioned in the text turned out to be better or worse than expected. The study added to the body of literature in information processing that unexpected text continuations can affect readers' attention to language cues. Moreover, the study also shed some light on the doctor-patient communication topic.

any cues available to them (e.g., Desmet et al., 2006; Fine et al., 2013; Rohde et al., 2011). For example, when processing relative clauses which can modify either of the two noun phrases (e.g., *the daughter of the policeman that...*), up to the point of reading the relative clause marker *that*, readers have already made a prediction as to which nouns the upcoming relative clause modifies (Siriwittayakorn, 2015). Studies have confirmed that continuation of texts that is incompatible with readers' expectation slows down the comprehension process comparing to the compatible one. Such effect of unexpected continuation is reflected in slower reading times (e.g., Fine et al., 2013; Rohde et al., 2011; Siriwittayakorn, 2015), or more eye fixations and regressions (e.g., Desmet et al., 2006). Readers' expectation plays a crucial role not only in online reading but also in offline reading where readers can take their time re-reading sentences as many times as they want before coming to a conclusion how they understand the sentences. For example, Pozniak and colleagues (in press) found that unexpected continuation of a sentence leads to lower acceptability even though such continuation is perfectly grammatical. The examples both from online and offline reading experiments highlight the importance of expectation on text comprehension.

As far as the issue is concerned, the effect of expectation on information processing has often been associated with processing difficulty measured in terms of reading times, eye fixations and regressions, and acceptability of sentences.

1 Introduction

Readers' (or listeners') expectation plays a key role in information processing. In real-time or online processing, it is known that readers do not wait until the next words appear but rather, predict the upcoming part of a text (Traxler, 2014) using

However, according to Sutcliffe and Christianson (2012), expectation affects how people perceive the world and it can bias people's behavioral choices not only by making them pay more attention to some details but also by making them ignore or overlook some details. The studies on the processing of narrative texts also support this view. It was reported that when readers strongly expected that the characters would achieved something, they paid little attention or ignored the context suggesting the characters' failure (see Sparks and Rapp, 2010 and references therein for more details). Therefore, it can be the case that, when processing non-narrative texts, readers ignore some words when there is discrepancy between their expectation and text continuations, and such ignorance may lead to meaning distortion. This study addresses this issue. Specifically, the study investigated the processing of possibility hedges when there was a conflict between readers' expectation on the severity of their illnesses and the medical knowledge (i.e., medical diagnoses and medical advice) given by doctors. This was done by using the degree of surprise and credibility judgment tasks. In this article, the experiment is reported first. Then, discussion on the results and their contributions both to the information processing studies, and to the medical communication topic are given.

2 Experiment

Having good health is what people in general wish for and when it comes to illnesses, doctors are supposed to be ones whom people rely on. When there are signs of illnesses, people normally assess the severity of their illnesses before seeing doctors. Normally, they also strongly expect that doctors will diagnose their illnesses and treat them correctly. However, because of the lack of medical knowledge, severity of illnesses as assessed by patients might be incompatible with doctors' assessment. Medical diagnoses or medical advice that patients do not expect beforehand might lead to a great surprise and this might affect how patients process doctors' diagnoses and advice. Because a great surprise might result from unexpected medical diagnoses and medical advice, an experiment employing medical context is suitable for testing the effect of expectation on text comprehension.

According to the literature, hedges are often used in medical conversations between doctors and patients and in medical academic writing. They are often used for many purposes such as suggesting possibility, building compliance, addressing humility, addressing appropriate level of accuracy, maintaining credibility, and saving medical practitioners from legal charge (Albl-Mikasa et al., 2015; Hyland, 2006). The discussion of Hyland (2006) implied that without appropriate hedges used, criticism on medical claim could be provoked. This discussion highlights the importance of hedges in medical communication. However, despite its importance, hedges were reported to be frequently ignored when interpreters interpreted doctors' words to foreign patients (Albl-Mikasa et al., 2015). Although the cause of omission was unclear, it might be the case that when comprehending doctors' messages, interpreters ignored hedges, and thus did not reproduce them when translating doctors' words to patients. If that is the case, there might also be a tendency for language users to ignore hedges when processing information.

Since hedges showing possibility were frequently used in medical communication (Hanauer et al., 2012), this experiment was conducted to investigate whether readers ignore them when doctors' diagnoses and advice were incompatible with their expectation. The experiment comprised of two tasks. The first task was the degree of surprise judgement task. This was used to examine whether doctors' diagnoses and advice could make participants surprised. The second task was credibility judgement task. This task was used with an assumption that if readers paid attention to possibility hedges and interpreted them when processing texts, the effect of such attention should be reflected in their judgement on the text credibility. The scope of this study was restricted only to medical diagnoses and medical advice given by doctors such that any of the credibility effects observed in the study would be confined only to the use of language not by the credibility of the information source.

2.1 Method

Participants: One hundred and thirty-four native Thai, undergraduate students at Chiang Mai university voluntarily participated in the

experiment. All of them did not study in the medical field.

Stimuli: There were six pairs of test items. All items were written in Thai. Each pair started with a sentence telling what symptoms participants had. Then, there was a statement starting with “The doctor says ...” followed by medical diagnosis and optional medical advice.

There was one within-item factor, namely *possibility*. That is, for each item pair, the use of hedges suggesting possibility for the medical diagnosis and medical advice to be true were manipulated. For the *low-possibility* condition, a hedge showing low possibility (i.e., a word equivalent to *might*) was used. For the *high-possibility* condition, a hedge showing high possibility (i.e., a phrase equivalent to *normally* or *there’s a high possibility that*) was used. Apart from the hedges, words in each pair of items were kept the same.

There was one between-item factor, namely *perceived severity*. The six pairs of test items were categorized into two types based on whether the symptoms indicated in each item pair were considered to be serious by people outside the medical field. To illustrate, for the *perceived-to-be-serious* type, the symptoms indicated in the item pairs were considered serious by general people. However, participants would learn from the doctor’s medical diagnoses and advice that the symptoms did not associated with serious illness. For the *perceived-to-be-unserious* type, the provided symptoms led general people to think that their illness was not serious. However, the doctor’s diagnoses indicated that the symptoms associated with serious illness. Examples of test items are in (1) and (2). The hedges were underlined.

(1) Perceived severity: Serious

a. Low-possibility condition

คุณมีอาการปากเบี้ยวหน้าเบี้ยว หมอบอกว่าเป็นอาการของเส้นประสาทที่ใบหน้าอักเสบ อาการนี้อาจหายได้เองด้วยการพักผ่อนให้เพียงพอ

“One side of your face droops or becomes stiff. The doctor says your facial nerve is damaged. You might recover from the symptom if you take enough rest.”

b. High-possibility condition

คุณมีอาการปากเบี้ยวหน้าเบี้ยว หมอบอกว่าเป็นอาการของเส้นประสาทที่ใบหน้าอักเสบ ส่วนใหญ่ แล้วอาการนี้หายได้เองด้วยการพักผ่อนให้เพียงพอ

“One side of your face droops or becomes stiff. The doctor says your facial nerve is damaged. Normally, you will recover from the symptom if you take enough rest.”

(2) Perceived severity: Unserious

a. Low-possibility condition

คุณมีอาการคัดจมูกมาสักระยะหนึ่ง หมอบอกว่าอาจมีเนื้องอกในโพรงจมูก

“You have got a stuffy nose for a while. The doctor says you might have a nasal tumor.”

b. High-possibility condition

คุณมีอาการคัดจมูกมาสักระยะหนึ่ง หมอบอกว่ามีความเป็นไปได้สูงที่จะมีเนื้องอกในโพรงจมูก

“You have got a stuffy nose for a while. The doctor says there is a high possibility that you have a nasal tumor.”

At this point, it is worth mentioning that the two types of items did not differ from each other only in terms of perceived severity but also in terms of the presence of medical advice. To illustrate, for the *perceived-to-be-serious* type, medical advice was always present together with medical diagnoses. However, medical advice was absent in the *perceived-to-be-unserious* type. This is because for the *perceived-to-be-unserious* type, only the medical diagnoses alone were enough to let participants know that their expectation was wrong (i.e., the symptoms were actually signs of serious illnesses). However, in the *perceived-to-be-serious* type, only the diagnoses alone were not enough to do so. Medical advice was added into the items of this type to let participants know that the symptoms they expected to be serious was in fact unserious. The hedges suggesting high possibility used in the two types of items were also different but the use was based on naturalness and was expected not to affect the results of the study as their meanings in Thai and in this experimental context were similar. All items were confirmed by

a doctor that the symptoms, medical diagnoses and medical advice were related. The use of high-possibility hedges in the *perceived-to-be-unserious* type was reported by a doctor to be rare and heavily rely on context because the symptoms indicated in the *perceived-to-be-unserious* type were uncommon presentation of the diseases.

To confirm that the symptoms described in each item pair would make participants perceive the severity as intended, a stimulus norming was conducted. A group of eleven native Thai speakers who did not participate in the experiment and did not work in the medical field were asked to rate the severity of ten sets of symptoms on a three-point scale where one indicated the lowest degree of severity and three indicated the highest degree of severity. Based on the norming results, three sets of symptoms were included in the *perceived-to-be-serious* type. The perceived severity of these items were between 54.55% - 90.91%. Another three sets of symptoms were included in the *perceived-to-be-unserious* type. The perceived severity of these items were between 0% to 18.18%. For the other four items, the majority of ratings was in between. That is, the perceived severity of the symptoms was not clear cut. Therefore, they were not included in the experiment.

Each test items were followed by two questions. The first question asked participants to rate the degree of their surprise when taking the symptoms and the doctor's saying into consideration. The second question asked participants to rate how much they believe in what the doctor said. The rating was done on a five-point scale with one indicating *not surprised* or *not believe* and five indicating *very surprised* or *strongly believe*.

Apart from the six pairs of test items, 14 fillers which were unrelated to medical contents were created. The structure of the fillers was the same as that of the test items. Six fillers discussed situations in which participants did something and teachers commented on their actions. The other six fillers discussed police's comments in relation to what participants did. Another two fillers were unrelated to all the situations mentioned so far. Some fillers contained possibility hedges. The possibility for the situations in the fillers to happen in reality was high for half of the items and low for the other half. All fillers were followed by two questions similar to those of test items. To monitor participants' attention, there were three fillers in

which the rating to one end of the scale was impossible.

Procedures: Test items were distributed into two lists according to a Latin Square design. The test items were interspersed with 14 fillers and were shown in a pseudo-random order such that participants would not read two test items consecutively. Each participant saw only one list. After reading each item, participants rated how great their surprise was and how credible the saying was.

For test items, when low-possibility hedges were used, it implied that doctors were not sure about his/her diagnoses or advice. This was regardless of whether the items belonged to the *perceived-to-be-serious* or *perceived-to-be-unserious* types. The doctor's low certainty left room for participants to think that their perceived severity had not yet been totally ruled out. However, when doctors used hedges indicating high possibility, it implied that the doctors were very certain about his/her diagnoses or advice. The use of high-possibility hedges thus, showed a strong sign that participants' perceived severity was wrong. In other words, the doctors' saying strongly contradicted participants' expectation. It was hypothesized that regardless of the types of items, the use of high-possibility hedges would cause a greater surprise to participants than did the use of low-possibility hedges. Moreover, the use of low-possibility hedges would make the credibility of the doctors' saying lower than did the use of high-possibility hedges. This was also regardless of the item types. However, if the information given by doctors strongly contradicted participants' expectation (i.e., perceived severity) and affected their comprehension, participants might ignore the hedges used. If this was the case, none of the effects should be observed.

Analyses: Twenty-seven participants who rated any two or three of the three fillers with either end where such rating was impossible were excluded from the analyses. Therefore, the analyses included data from 107 participants. For these 107 participants, 75.70% of them did not make mistake in rating the three fillers and 24.30% of them made only one mistake.

Analyses were run on R version 3.5.0 (R Core Team, 2018) with Bayesian cumulative link mixed-effects model (see Pozniak et al., in press, and references therein for advantages in using

Bayesian analyses) using the *brms* package version 2.3.1 (Bürkner, 2017; in press).

The categorical variables were coded following Pozniak et al. (in press). That is, the variable namely *possibility* was coded as -1 for *low possibility* and 1 for *high possibility*, and the variable namely *perceived severity* was coded as -1 for *perceived-to-be-unserious* type and 1 for *perceived-to-be-serious* type. According to Pozniak et al. (in press), by coding levels of a variable as either 1 or -1, zero is between the two levels and represents the lack of difference. If zero is included in the 95% credible intervals (CrI) of the posterior distribution of the coefficient parameter (β), it means the difference is not reliable. Otherwise, the difference is considered reliable.

For the analysis of the degree of surprise, the rating scores were analyzed as a function of possibility, perceived severity and their interaction. Even though the test items were categorized as either *perceived-to-be-serious* type or *perceived-to-be-unserious* type, the items in each type were still different from one another in terms of the degree of perceived severity. Therefore, to capture the effect of such difference, the degree of perceived-severity which were obtained from the stimulus norming was also included as a fixed factor. Random intercept, random slope and their correlation were included for participants and for items. The model was run with four chains, each with 2000 iterations and the first 1000 iterations were served as warm-up. The fact that all the chains converged and the \hat{R} statistics of every parameter coefficient was equal to 1 verified the model convergence. For the analysis of the credibility scores, the model was run in the same fashion as that of the degree of surprise.

2.2 Results

Degree of surprise: For the *perceived-to-be-serious* type, when hedges showing low possibility was used, the median of the degree of surprise was 3. The median was 3.5 when hedges showing high possibility was used. For the *perceived-to-be-unserious* type, the medians of both sentences with low-possibility hedges and sentences with high-possibility hedges went up to 5.

However, the Bayesian cumulative link mixed-effects model indicated that none of the main effects nor the interaction was observed

(possibility, $\beta = -0.18$, 95% CrI = [-0.51, 0.14]; perceived severity, $\beta = -4.62$, 95% CrI = [-15.06, 7.24]; degree of perceived-severity, $\beta = 1.49$, 95% CrI = [-4.16, 6.56]; interaction between certainty and perceived severity, $\beta = -0.23$, 95% CrI = [-0.89, 0.42]).

Credibility: For the *perceived-to-be-serious* type, the medians of the credibility scores were 4 regardless of whether hedges showing low or high possibility were used. For the *perceived-to-be-unserious* type, the medians dropped to 3 regardless of the hedges used.

According to the Bayesian cumulative link mixed-effects model, there was neither main effect of possibility ($\beta = 0.18$, 95% CrI = [-0.12, 0.46]), main effect of perceived severity ($\beta = 3.31$, 95% CrI = [-4.60, 10.39]), nor main effect of the degree of perceived severity ($\beta = -1.22$, 95% CrI = [-4.87, 2.55]). However, there was an interaction between possibility and perceived severity ($\beta = -0.85$, 95% CrI = [-1.49, -0.23]).

Since there was an interaction between possibility and perceived severity, additional analyses were run for the items in the *perceived-to-be-serious* and *perceived-to-be-unserious* types separately. The models included possibility and the degree of perceived severity as fixed factors. Otherwise, they were the same as those indicated in the Analyses section. It was found that for the *perceived-to-be-serious* type, none of the main effects was observed (possibility, $\beta = -0.20$, 95% CrI = [-0.63, 0.23]; degree of perceived severity, $\beta = -1.61$, 95% CrI = [-16.36, 11.05]). Importantly, the result showed that for the *perceived-to-be-unserious* type, there was a main effect of possibility ($\beta = 0.72$, 95% CrI = [0.28, 1.18]) such that when hedges showing high possibility was used, the credibility scores increased. The main effect of the degree of perceived severity was not observed ($\beta = -0.21$, 95% CrI = [-21.91, 22.91]).

2.3 Discussion

The medians of the degree of surprise ranging from 3 to 5 suggested that the experimental stimuli worked as planned. That is, the stimuli could make participants surprised when reading the continuation of the texts. The fact that the degree of surprise of the *perceived-to-be-unserious* type was higher, albeit not reliable, than that of the *perceived-to-be-serious* type was not surprising. When a situation turned out to be better than

expected, there could be a sign of relief but when the situation turned out to be worse than expected, a great surprise could be resulted.

The lack of main effects and an interaction on the degree of surprise indicated that the use of low-possibility and high-possibility hedges could not decrease or intensify the degree of surprise no matter how large the discrepancy between what participants expected (i.e., perceived severity) and the continuation of the text were. However, whether or not the lack of such effects resulted from participants' ignorance of the hedges when reading unexpected text continuations could not be determined by the analyses of the degree of their surprise alone.

As for the credibility, the medians credibility scores for both types of items was moderate to high (i.e., 3 to 4). This might be because it was stated clearly in the text that the medical diagnoses and advice were given by doctors, an often-regarded-as credible information source.

Importantly, the presence of the possibility effect on the credibility of the items in *perceived-to-be-unserious* type suggested that participants paid attention to hedges when processing information. However, the lack of the possibility effect on credibility of the *perceived-to-be-serious* type suggested that attention participants paid to hedges could be in a varying degree based on the situations.

The lack of the possibility effect on the credibility of items in *perceived-to-be-serious* type might be because no matter how highly possible the illness situation was, when the unexpected situations turned out to be better than expected, all was good. Therefore, there was no need to rely on possibility cues. However, for the *perceived-to-be-unserious* type, participants might have paid more attention to possibility hedges because when the situation turned out to be worse than expected, there might have been a need to seek for language cues that could help estimate how true the situation could be.

Another possible explanation for the dissimilar effect of possibility on the two types of items is that, the lack of possibility effect on the *perceived-to-be-serious* type resulted from the difference in the stimuli used. That is, the medical advice present in the items of *perceived-to-be-serious* type might have enhanced credibility of the doctors' saying. Therefore, there was no need to

rely on hedges when interpreting credibility of doctors' saying. However, further experiment that provides medical advice to the items in the *perceived-to-be-unserious* type is needed in order to prove whether this is true.

Because of the absence of a control condition in which the perceived severity matches medical diagnoses, the exact direction of the ratings cannot be pinpointed. For example, whether the degree of surprise and credibility are lower or higher than usual cannot be concluded from this study. Further studies with a controlled condition as a baseline are needed.

3 General Discussion

The results of the present study were in line with previous literature in information processing suggesting that when processing sentences, readers used any cues available to them to guide their interpretation. This present study added to the body of literature in information processing that in unexpected situations, the degree of participants' attention to the language cues can be varied.

Sparks and Rapp (2010) found that in the processing of narrative texts, readers ignored language cues when the story context contradicted their expectation. This present study was conducted in a non-narrative context. Moreover, in this study there were two types of unexpected situations including the ones that turned out to be better than expected (i.e., *perceived-to-be-serious* item type) and those that turned out to be worse (i.e., *perceived-to-be-unserious* item type). By comparing the results of both types of situations, it can be seen that participants paid less attention to possibility hedges when the situations turned out to be better than expected. Thus, first this present study showed that not only in the narrative context but also in the non-narrative one that the effect of expectation on the processing of unexpected texts can be observed. Second, different types of unexpected situations can cause different effects on information processing.

Sparks and Rapp (2010) also suggested that emotional factor played a role in processing. Given that in real life, reading or listening to medical diagnoses can sometimes involve feelings such as anxiety and fear, it is worth to further explore whether the signal of ignorance to language cues will be stronger if the emotional factor is involved.

The results of this study also shed some light on the topic of doctor-patient communication. Firstly, studies on the use of hedges in medical context largely focused on medical academic writing (e.g., Hyland, 2006; Csongor and Rébék-Nagy, 2013). Some focused on how health-care providers and patients interpreted possibility hedges by asking participants to assign percentages to those hedges (Biehl and Halpern-Felsher, 2001; Foppa et al., 2011). However, research on how patients' expectation affects the processing of possibility hedges has not yet been found. The present study added to a body of literature that hedges showing high possibility can make doctors' diagnoses sound more credible to patients, especially when patients do not expect severe illnesses. Van Boekel et al. (2017), stated that when readers read texts that contradict their knowledge, credibility of the texts play an important role in helping them revise their knowledge. When discussing on severe illness with patients who do not expect to have it, making the diagnoses sound more credible might help the patients correct their false belief about their illness, accept the truth more easily and give more cooperation to the therapeutic process.

Secondly, with a quick survey on breaking bad news guidelines, it was found that useful language devices receive little attention. Although the experimental stimuli are totally or almost totally different from what will be used in real doctor-patient communication, there is a good sign showing that language devices that are useful for doctor-patient communication, can alleviate patients' emotional distress, or promote more collaboration in therapeutic process can be identified by using the method in information processing.

4 Conclusion

The study on the processing of possibility hedges in unexpected text continuations showed that the difference in the possibility hedges used did not affect the degree of participants' surprise. However, for the text credibility, it was found that participants paid more attention to possibility hedges when text continuations turned out to be worse than expected than when the continuations turned out to be better than expected. Altogether, the results suggested that even though participants paid attention to language cues when processing

sentences, the degree to which they paid attention to a particular cue varied depending on their expectation.

Acknowledgments

I would like to express my gratitude to Chatmanee Siri Wittayakorn, M.D., for invaluable discussion on medical diagnoses and advice and Pasoot Lasuka, Ph.D., for inspiring discussion.

References

- Albl-Mikasa, M., Glatz, E., Hofer, G., & Sleptsova, M. (2015). Caution and compliance in medical encounters: Non-interpretation of hedges and phatic tokens. *Translation & Interpreting Research*, 7(3), 76-89.
- Biehl, M., & Halpern-Felsher, B. L. (2001). Adolescents' and Adults' Understanding of Probability Expressions. *Journal of Adolescent Health*, 28, 30-35.
- Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28. doi:10.18637/jss.v080.i01
- Bürkner, P. (in press). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*.
- Csongor, A., & Rébék-Nagy, G. (2013). Hedging in popular scientific articles on medicine. *Acta Medica Marisicensis*, 59(2), 97-99.
- Desmet, T., De Baecke, C., Drieghe, D., Brysbaert, M., & Vonk, W. (2006). Relative clause attachment in Dutch: On-line comprehension corresponds to corpus frequencies when lexical variables are taken into account. *Language and Cognitive Processes*, 21(4), 453-485.
- Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T. (2013). Rapid expectation adaptation during syntactic comprehension. *PloS one*, 8(10), e77661.
- Foppa, M., Araujo, B. S. d., Macari, A., Reichert, R., & Goldim, J. R. (2011). Limitations in the use of qualitative terms to inform diagnoses. *ARCH INTERN MED*, 171(14), 1291-1292.
- Hanauer, D. A., Liu, Y., Mei, Q., Manion, F. J., Balis, U. J., & Zheng, K. (2012). Hedging their bets: the use of uncertainty terms in clinical documents and its potential implications when sharing the documents with patients *AMIA Annual Symposium Proceedings* (pp. 321-330).
- Hyland, K. (2006). Medical Discourse: Hedges. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (pp. 694-697): Elsevier.
- Pozniak, C., Hemforth, B., Haendler, Y., Santi, A., & Grillo, N. (in press). Seeing events vs. entities: The processing advantage of Pseudo Relatives over Relative Clauses.

- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rohde, H., Levy, R., & Kehler, A. (2011). Anticipating explanations in relative clause processing. *Cognition*, 118(3), 339-358.
- Siriwittayakorn, T. (2015). *The processing of relative clause attachment in Thai and the role of experience in sentence processing*. (Doctoral), Chulalongkorn University, Bangkok.
- Sparks, J. R., & Rapp, D. N. (2010). Discourse processing - examining our everyday language experiences. *WIREs Cognitive Science*, 371-381. doi:10.1002/wcs.11
- Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.17.3. <http://mc-stan.org/>.
- Sutcliffe, K. M., & Christianson, M. K. (2012). Managing the unexpected. In K. Camerron & G. M. Spreitzer (Eds.), *The Oxford Handbook of Positive Organizational Scholarship* (pp. 843-854). USA.: Oxford University Press.
- Traxler, M.J. (2014). Trends in syntactic parsing: anticipation, Bayesian estimation, and good enough parsing. *Trends in cognitive sciences*, 18(11), 605-611.
- Van Boekel, M., Lassonde, K. A., O'Brien, E. J., & Kendeou, P. (2017). Source credibility and the processing of refutation texts. *Memory & Cognition*, 45, 168-181.
- Wickham, H., Hester, J., & Chang W. (2018). devtools: Tools to Make Developing R Packages Easier. R package version 1.13.6. <https://CRAN.R-project.org/package=devtools>