Participants and design

Undergraduate and graduate phontecis and psychology students (80.8% female, median age = 21, IQR = 3, range = [18, 31], total $N = 207$) from the University of Zagreb participated in the study in exchange for course credit. Participants were randomly assigned to one of five groups which differed in the type of activity they engaged in between parts of the text they have read and in whether they received feedback on their intermittent test achievement or not.

Materials and procedure

Materials

Participants read a text on the evolution, ecological and biological characteristics of weeds. The text was taken from a chapter in a Croatian university-level textbook. Some sentences and passages were slightly modified, so as to avoid odd language constructions; Latin plant names were translated to Croatian, and some plants were removed from the text to make it less difficult for the target participant population. The text was divided into three parts of 874, 754, and 835 words, respectively. Additionaly, there was a practice text taken from the same chapter, but unrelated to any of the other three parts of the text (768 words).

Forty-four content related questions with four response options were generated from the presented texts. Four questions were presented after the practice text, ten after each of the first two parts (only to the participatns in the content related test condition), and twenty after the third part of the text (to all participants). Starting from the second ten question set, the distractor options were chosen so that (a) two distractors were plausible, but unrelated to the text, and (b) one distractor was a term or concept mentioned in the previous part of the text — this was considered to be the "intrusive" option.

An example question is:

> Compared to younger weeds, older weeds:
> (a) have a stronger allelopathic effect
> (b) contain more phytotoxins
> (c) **contain less inhibitory matter**
> (d) *show greater plasticity.*

Option (c) is the correct answer, and option (d) is the intrusive distractor.

Further, twenty-four general knowledge questions were generated. These questions were presented to participants in the general knowledge test condition, after the first two parts of the text. An example general knowledge question is:

> The name of Kurt Vonnegut's famous anti-war novel is:
> (a) **Slaughterhouse Five**
> (b) All Quiet on the Western Front
> (c) A Farewell to Arms
> (d) Journey to the End of the Night.

At the beginning of the session, participants' ID, age and sex information was collected. At the end of the session, participants were asked to estimate how much of each text they have read. The texts and questions

were presented on a personal computer, in an application constructed using the open source `oTree` framework (version 2.1.35, **?**) for the Python programming language (version 3.6.4, October 20, 2018).

Procedure

Participants were first given a brief introduction to the study, and were encouraged to carefully read and follow the written instructions. Then, they were led to one of six compartments containing a computer, which was running a fullscreen instance of the `oTree` application with a randomly chosen experimental condition. There, participants read the informed consent form and, in case there were no questions, started the experiment.

After entering their personal information, participants were presented with instructions for their first task, which was to read the practice text at a speed that comes naturally to them. They were to click a button at the bottom of the text when they have finished reading it. Unbeknownst to the participants, the time they took to read the practice text was recorded, and used as the basis for determining the reading time limits for the remaining texts. Results of a pilot study using different participants have shown that most participants found 4 minutes to be too short, and 9 minutes too long, so we have set the lowest possible limit to 5 minutes, and the longest to 8 minutes.

Next, participants were familiarised with the interpolated activity they were going to perform during the main part of the procedure. The rereading group reread the practice text (this time with the time limit applied), the general knowledge test group answered four general knowledge questions, and the content related test group answered four questions based on the practice text.

Participants assigned to the feedback condition also received feedback on their practice test achievement. Feedback was presented on a separate screen, which listed the questions, the participant's answers, and the correct answers in tabular format. Incorrectly answered questions were highlighted in red. After 40 seconds elapsed, a "Next" button appeared, allowing participants to proceed to the next text. This way, we wanted to prevent participants from simply clicking through the feedback, hoping that they will spend their time examining it. The feedback was presented for maximally 60 seconds, after which the application proceeded to the next text.

Subjects in the rereading and general knowledge conditions also answered the four questions related to the practice text, so as to familiarise them with the scope and specificity level of questions that they will receive after reading the final text. All participants were told that there would be a cumulative test after the final text, examining their knowledge of the three texts following the practice text. In reality, the final test examined only the knowledge of the final text.

After the practice round, participants proceeded to read the main texts, engaging in the interpolated activities they were assigned. After the third text, all participatns were presented with twenty questions examining their knowledge of the third text. The computer recorded whether a participant correctly answered a question and whether the participant chose an intrusive distractor. This allowed us to compute our dependent variables — the total number of correct answers and the total number of intrusive distractors chosen.