**Research Program**

**1. Scientific background**

Effective communication relies on accurate estimation of other people's knowledge, particularly in terms of vocabulary. Overestimating audience’s familiarity with terms can impede the audience's understanding and lead to communication breakdowns (Bullock et al., 2019). This skill is especially crucial for bilingual speakers, who frequently assess their listeners' proficiency to adjust their language use (Gasiorek et al., 2022). However, how bilinguals judge their partners' proficiency and estimate their vocabulary knowledge remains unclear. In a recent study, I examined how 298 Russian-Hebrew bilinguals evaluated other people's interpretations of ambiguous messages and found that they based their evaluations on their own interpretations, even when these interpretations were irrelevant (Segal & Kavé, 2024). Drawing on these findings, I assume that bilinguals may rely on their own proficiency when estimating the proficiency of others. However, people do not have direct access to their own proficiency, and they have to infer it based on various cues (Ackerman, 2019; Koriat, 1997). These cues often lead them to misjudge their own proficiency (Kavé & Halamish, 2015; Zell & Krizan, 2014), and therefore, their evaluations of others' proficiency may also be biased. Moreover, self-proficiency is clearly not the sole determinant of one's estimation of others' proficiency, and other factors may also play a role. These factors may relate to the estimator (e.g., their perceived proficiency), to contextual characteristics (e.g., item familiarity), and to the individual whose proficiency is being estimated (e.g., their linguistic background). The proposed research will explore how these factors affect bilinguals' estimations of both their own and others' language knowledge, with a focus on vocabulary.

Bilinguals must constantly monitor their conversational partners' language proficiency and adjust their language use accordingly (Kapiley & Mishra, 2024). When speaking with other bilinguals, this includes selecting which language to use, and switching languages when necessary (Ariffin & Rafik-Galea, 2009). In conversations with both monolinguals and bilinguals, it involves the selection of words that are not too difficult for less proficient audience, but also not overly simplistic for highly proficient audience (Gasiorek et al., 2022). Importantly, people often misjudge their audience's proficiency, using a language that their listeners do not fully understand (Amano et al., 2021; Babayiğit & Shapiro, 2020; Flores et al., 1998; Wilson et al., 2005). Hence, it is essential that we understand the factors that affect estimation of language proficiency.

Research on estimation of other people’s knowledge and perspective-taking research relies heavily on the Anchoring and Adjustment theory (Tversky & Kahneman, 1974). According to this theory, people judge other people's knowledge by automatically anchoring on their own knowledge and then adjusting their assessment by accounting for possible differences between themselves and others (Damen et al., 2020; Epley & Gilovich, 2001, 2006). These adjustments are often insufficient, leading to biased estimations (Epley et al., 2004; Lau et al., 2022). For example, Keysar and Bly (1995) asked participants to study unfamiliar idioms such as "the goose hangs high" in contexts that supported either the correct meaning (things look good) or the opposite meaning (things look bad). When asked to predict how others would interpret these idioms, participants expected others to choose the meaning that they themselves had studied. Similarly, Epley et al. (2004) found that participants who received information about the identity of two cola drinks (Coca-Cola and Pepsi) believed that others would also be able to tell them apart, and their assessment was higher than the assessment provided by participants who did not know the type of drinks. In my recent study of Russian-Hebrew bilinguals, I asked participants to read message correspondence in Russian and in Hebrew and to judge interpretation of these messages by other recipients (Segal & Kavé, 2024). Half the texts contained information that suggested that the message was sincere, and the other half implied that it was sarcastic (as in Keysar, 1994). Importantly, this information was available to the participants but not to the message recipient, and therefore it should not have affected the recipient’s interpretation. The findings suggested that participants made their decisions according to their own interpretation.

Notably, the notion of automatic anchoring to self-knowledge might conflict with metacognitive research that shows that people do not have full access to their own knowledge and that they must make inferences about their knowledge on the basis of various cues (Ackerman, 2019; Koriat, 1997; Nisbett & Wilson, 1977; Reder & Ritter, 1992). Such cues may include self-perceptions (e.g., "I am good at this task"), momentary experiences during the task (e.g., the perceived ease of processing, called fluency), as well as other broader cues that include beliefs about the task (e.g., "multiple-choice questions are easier than open-ended ones"). Cues can lead to accurate estimates of one's actual knowledge, but they may also introduce bias, resulting in underconfidence or in overconfidence. If people base their estimations of other people's knowledge on their estimation of self-knowledge, then biased estimation of self-knowledge can in turn bias the estimation of other people’s knowledge. For instance, Tullis (2018) asked participants to answer trivia questions either before or after predicting the knowledge of others. Participants’ predictions of other people's knowledge were more strongly related to their own processing (i.e., response fluency and accuracy) when they answered the questions before estimating other people's knowledge. In a later study, participants answered trivia questions, received feedback on whether their answers were correct, and then estimated the likelihood that other participants would know the correct answer, on a 0% to 100% scale (Tullis & Feder, 2023). After three rounds of this process, participants' knowledge of the trivia answers increased, and they began to significantly overestimate the knowledge of others. Thus, the assessment of other people’s knowledge depends on the cues that are available to the judges.

When bilinguals estimate others' proficiency, they may use their confidence in their own language skills as a cue. This confidence can be influenced by their perception of their own language proficiency. In a study that looked at estimation of one’s vocabulary knowledge, Kavé and Halamish (2015) asked younger, middle-aged, and older adults to complete a vocabulary test and to rate their confidence in each answer. Then, they compared the actual performance (vocabulary score) to the perceived performance (the mean confidence ratings across items) to form a calibration score (Nelson & Narens, 1990). A score of zero reflects perfect calibration, while a negative score indicates underconfidence and a positive score indicates over-confidence. Results showed that younger adults knew the meaning of fewer words than did the other groups and were underconfident in their performance. Middle-aged and older adults knew more words and their confidence aligned with their actual performance, indicating better calibration. This study demonstrates that the assessment of knowledge in one’s native language involves imperfect calibration, at least among younger adults. However, it is unclear whether the differences in confidence levels across age groups are due to differences in vocabulary knowledge or simply because of age itself. This question can be addressed by comparing confidence levels in languages with varying proficiency levels within the same individuals.

Many bilinguals have unbalanced knowledge of their languages. They tend to be more proficient in their dominant language and perceive their proficiency as higher in that language compared to their non-dominant language (Gollan & Ferreira, 2009). Low proficiency and perceived proficiency might contribute to decreased confidence in self-knowledge (Botes et al., 2020; Liu, 2006; MacIntyre et al., 1997). However, this decreased confidence may not accurately reflect true knowledge; instead, it could lead to underconfidence in self-assessments. If bilinguals base their estimations of other people’s proficiency on their self-assessed proficiency, and if they are more confident in their dominant language, they may judge the proficiency of others as higher in the dominant language relative to the non-dominant language, especially when they perceive these others as having similar language knowledge to their own. In the current study, I plan to examine this hypothesis by comparing bilinguals' confidence in selecting the correct meanings of words in their dominant and non-dominant languages. I will explore how these confidence levels align with actual performance (i.e., calibration score) and test the association between their estimation of self-knowledge and their predictions of other people's knowledge in both languages.

Confidence judgments are influenced not only by participants' proficiency and perceived proficiency but also by their ongoing task experiences (Ackerman, 2019) and by the familiarity of task items (Koriat, 2012). Familiar items are typically processed more easily than unfamiliar ones (familiarity effect; Geng et al., 2023; Segal & Gollan, 2018), and therefore participants are more confident in their knowledge of these items (Fitzsimmons et al., 2020; Markovits et al., 2015). Ease of processing (fluency) may not only affect one's confidence in self-knowledge, but also the estimation of other people’s knowledge. According to the fluency misattribution account, we often believe that ease of processing reflects objective ease (Birch et al., 2017). For example, quickly recalling the meaning of the word "chaos" may lead someone to assume that this word is easy and familiar to everyone. The relationship between word familiarity and ease of processing may be modulated by exposure, as suggested for frequency effects (i.e., faster recognition of high frequency words compared to low frequency words; Diependaele et al., 2013; Monaghan et al., 2017; Mor & Prior, 2020). Beyond a certain point, additional exposure no longer enhances familiarity, and therefore greater proficiency should lead to a reduction in the familiarity effect. In a previous study, I examined how Spanish-English bilinguals and English-speaking monolinguals processed literal and metaphorical English expressions of varying familiarity (Segal & Gollan, 2018). Bilinguals with the lowest English proficiency processed less familiar expressions more slowly and less accurately than familiar ones, whereas the most proficient English speakers showed no familiarity effect. These findings suggest that proficiency levels interact with item familiarity, and it is possible that this interaction will also affect estimation of one’s own and other people's knowledge. I thus plan to examine whether individuals with different proficiency levels take word familiarity into consideration when judging their own vocabulary as well as the vocabulary knowledge of others.

It is commonly assumed that when bilinguals encounter a word in one language, they automatically activate similar words in both languages (Dijkstra & van Heuven, 2002). This dual activation is especially pronounced in words that share semantic and phonological similarities across languages, such as cognates (e.g., /balon/ in Hebrew and "balloon" in English), and loanwords (e.g., the Hebrew word /gravitazia/ which is equivalent to the English word "gravity"). These overlapping words can create a false sense of familiarity (Costa et al., 2000; Masson, 2013), thus leading to biases in the assessment of knowledge, especially when estimating the knowledge of monolingual speakers who have only one representation of the word (Birch et al., 2017). Consequently, bilinguals may rely on cross-linguistic similarities as an additional cue when inferring the meaning of unknown words, leading to biases in their estimation of other people’s knowledge. I will explore this hypothesis by testing how bilinguals judge their knowledge of native words and loanwords, as well as by asking them to estimate monolinguals' knowledge of these items.

One's experiences during a task can vary not only by item familiarity, but also by the nature of the task (Ackerman, 2019). For example, Ibabe and Sporer (2004) found that undergraduate students were more confident in their memory of a robbery film when answering multiple-choice questions compared to open-ended questions. Zell and Krizan (2014) conducted a meta-analysis to examine the correspondence between self-evaluations of ability and objective performance, and found that the correspondence was higher for low complexity and familiar tasks compared to less familiar and more complex tasks. Vocabulary knowledge is often assessed with multiple-choice questions (Segal, 2023), but some authors use additional tests (Kavé & Yafé, 2014). Multiple-choice questions require recognition of both the target word and the response options, and participants can rely on partial recognition, elimination of incorrect answers, or guessing (Gyllstad et al., 2015; Kumar et al., 2023). In contrast, a word definition task requires true knowledge of a word’s meaning (Kavé & Yafé, 2014). If the nature of the task affects one’s confidence in one’s own knowledge (Ackerman, 2019), it may also influence the estimation of other people's knowledge. To investigate these issues, I plan to compare estimation of knowledge on a multiple-choice task with estimation that involves open-ended tasks.

Finally, estimation of the vocabulary knowledge of others may also depend on the presumed similarity between the estimator and the person whose knowledge is being estimated (Dragojevic et al., 2016; Stell & Dragojevic, 2017). A native speaker who encounters an unfamiliar word might assume that another native speaker will not recognize it, but a more proficient speaker (e.g., a language editor, a language teacher) will know its meaning. These assumptions could be correct or incorrect, reflecting over-confidence when predicting the knowledge of more proficient speakers or underconfidence when predicting the knowledge of less proficient speakers. Hence, I plan to examine how the linguistic background of different others affects one's estimation of their knowledge.

**2. Research Objectives and Expected Significance**

Previous research shows that people misjudge the proficiency of their listeners in both monolingual (Gotlieb et al., 2022; Sharon & Baram-Tsabari, 2014) and bilingual contexts (Flores et al., 1998; Wilson et al., 2005), and often use vocabulary that the listeners do not fully understand. However, the source of this bias is unclear. The proposed research explores how bilinguals assess their own vocabulary knowledge in languages of varying proficiency, and how these self-assessments are influenced by factors such as word familiarity and task type. Additionally, the study investigates how self-assessments shape the judgement of other people's vocabulary knowledge, particularly when their linguistic backgrounds are similar to or different from the bilinguals' own background. The project aims to provide an overarching account of the factors that affect knowledge estimation in bilingual contexts for both self and others, and to shed light on related topics in bilingualism, perspective taking, and metacognition.

**The significance of this research** lies in its exploration of novel theoretical questions with broad implications both within and beyond the field of bilingualism, specifically regarding how individuals assess their own and other people's knowledge. Accurate estimation of knowledge is crucial for effective word selection in communication, for comprehension monitoring, and for learning (Shulman et al., 2020). The results of the current project can also account for the frequent discrepancies between subjective and objective proficiency measures in bilinguals (Marian et al., 2007; Tomoschuk et al., 2019; Zell & Krizan, 2014). It can also help explain why misjudgments of the proficiency of others are so common, and it can help reduce this bias in the future. The research is designed to achieve the following objectives:

***Objective 1:* To advance and refine existing models of assessment of self-knowledge by pinpointing the role of level of knowledge, knowledge familiarity and the type of task that examines this knowledge in the assessment process.**

***Objective 2:* To propose a new conceptual framework that defines the dynamic interplay between the estimation of self-knowledge and confidence in that knowledge, and the judgement of others' knowledge.**

***Objective 3:* To define the role of contextual cues (knowledge familiarity) and the similarity between the estimator and the person whose knowledge is being estimated in the estimation process.**

***Objective 4:* To deepen the understanding of how bilinguals estimate their proficiency in both their dominant and non-dominant languages, uncovering the factors that lead to frequent misjudgments of self-proficiency. Additionally, the study aims to explore the cognitive processes involved in judging others' proficiency and identify the existence and sources of biases in these judgments.**

**3. Detailed Description of the Proposed Research**

**3.1 Working Hypotheses**

I hypothesize that: (a) participants will exhibit underconfidence in their vocabulary knowledge in both languages, particularly for less familiar words; (b) participants who demonstrate greater knowledge on a vocabulary task will also show higher confidence in their knowledge (i.e., a more attenuated underconfidence); (c) confidence ratings for self-knowledge will correlate with the estimation of others' knowledge, but only when judging individuals with a similar linguistic background to one’s own; (d) confidence ratings will be higher for familiar words and for loanwords compared to less familiar words and to native words; (e) confidence ratings will be higher in multiple-choice tasks compared to definition tasks, leading to a stronger correspondence between self-evaluations and actual performance; and (f)participants will be overconfident in judging the proficiency of individuals they perceive as highly proficient and underconfident when judging those they perceive as less proficient

**3.2 Research Design and Methods**

The proposed research will consist of five interrelated but independent studies (see Table 1). All studies will receive ethics approval prior to administration.

**Table 1**

*Summary of Research Design by Study*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Main variable** | **Task** | **Language of test** | **Linguistic background of the assessed group** |
| 1 | Self-proficiency | Multiple-choice | Hebrew + English | Hebrew speakers |
| 2 | Word familiarity | Multiple-choice | English | Hebrew speakers |
| 3 | Cross-language similarities | Multiple-choice | English | Monolingual English speakers |
| 4 | Task type | 1. Multiple-choice  2. Word definition  3. No specific task | Hebrew | Hebrew speakers |
| 5 | Similarity in linguistic background | Multiple-choice | Hebrew | 1. Hebrew speakers  2. Language editors (higher proficiency)  3. Immigrants (lower proficiency) |

*Participants*: A power analysis using the G\*Power3 program (Faul et al., 2007) indicated that for a regression analysis with four predicting variables (years of education, objective proficiency, self-rating of proficiency, and confidence in self-vocabulary), a sample size of 87 participants would suffice to detect a medium effect size (f2 = .15) with an alpha of .05, leading to a power of .80. Therefore, studies with this design will include 90 participants. All participants will be Hebrew-English bilingual students aged 20-30, who are either Hebrew-dominant or balanced bilinguals, with Hebrew as their native language, and English acquired after age 8. Israeli children typically begin learning English in 3rd grade (around ages 8-9) and continue through high school, yet English proficiency varies significantly among adult Hebrew speakers. Participants with reading disorders or those reporting poor or very poor English proficiency (rating their production, comprehension, reading, or writing skills as 1 or 2 on a 1-7 scale) will be excluded. Additionally, participants who acquired English from birth or spent more than six months in an English-speaking country will be excluded. This procedure ensures sufficient variation in proficiency while minimizing alternative explanations that involve disorders or complete lack of knowledge. Participants will be recruited through the Open University lab, online survey companies, or snowball sampling. These methods will also serve to recruit raters of stimuli, as well as comparison groups (see below).

*Background tests*: All participants will complete a demographic background and language history questionnaire (as in Segal & Kavé, 2024), in which they will provide demographic information, indicate when they began studying English, rate their proficiency in both Hebrew and English on a scale of 1 (not at all) to 7 (like a native speaker), report their daily and weekly use of each language, and declare their reading habits. Participants' proficiency in Hebrew and English will also be assessed with the time-limited version of the Multilingual Naming Test (MINT Sprint), developed by Garcia and Gollan (2022). The MINT Sprint, based on the original MINT (Gollan et al., 2012), measures language proficiency in bilingual speakers of English, Hebrew, Spanish, and Mandarin. Participants will be given three minutes to orally name as many pictures as possible in each language, with the order of languages counterbalanced.

*Experimental tasks*: In all studies, participants will complete a vocabulary test in either English, Hebrew, or both. In each test, they will select the correct meaning of words from four response alternatives, rate their confidence in each response, and predict the likelihood that another participant will choose the correct answer. In Study 4, participants will also be asked to define words and to estimate their knowledge of each word without a specific task.

*Planned analyses*: To assess the accuracy of participants' self-knowledge estimations, their mean confidence ratings across items will be compared to their actual performance (as in Kavé & Halamish, 2015; Sidi et al., 2017). Similarly, the accuracy of participants' estimations of other people's knowledge will be evaluated by comparing their mean predictions for others to the actual performance of a relevant sample. For example, estimation of the knowledge of Hebrew speakers whose parents are English speakers will be compared to the actual performance of such a sample. The difference scores will serve as calibration measures, with positive values reflecting over-confidence and negative values reflecting underconfidence. Response times for selecting the correct answer will serve as a measure of ease of processing (fluency). Correlation and regression analyses will examine the associations between proficiency measures, response times, confidence ratings, and calibration scores for both self and others.

**Preliminary results**

A preliminary study included two experiments. In Experiment 1, Hebrew English bilinguals (*N* = 89, *M*age= 29.12), who met all the inclusion criteria described above, except for their age that ranged between 20 and 40, completed a multiple-choice vocabulary task consisting of 35 low-frequency words in their non-dominant language, English. They chose the correct meaning of each word, out of 4 response alternatives, rated their confidence in each response on a 0-100 scale, and rated the chances that other Hebrew speakers would select the correct meaning of each word on a similar scale. Experiment 2 was identical, only that participants (*N* = 106, *M*age= 30.11) predicted the chance that other Hebrew speakers, with English speaking parents, will answer each question correctly. The order of confidence rating for self-knowledge and prediction ratings for others' knowledge was counterbalanced among participants. All participants were recruited through a survey company and each was included in only one experiment. The study received an ethical approval from the Ethics Committee of the Open University.

For each participant, five measures were calculated (see Table 2). Confidence in self-knowledge was measured as the average of confidence ratings across items. The prediction of others' knowledge was assessed as the average prediction ratings for others' knowledge across items. The vocabulary score was determined by the number of correct responses out of 35 questions. Self-calibration was calculated by subtracting the participant's mean confidence in self-knowledge from their vocabulary score. In Experiment 1, an other-calibration measure was calculated by subtracting the mean prediction of others' knowledge from the overall vocabulary score of the entire sample, which consisted of Hebrew speakers. In Experiment 2, the other-calibration measure was determined by subtracting participants' mean prediction of others' knowledge from the vocabulary scores of a separate group of native Hebrew speakers whose parents were English speakers (*N* = 42, *M*age = 29.74), who were tested using the same vocabulary task. Table 2 presents the mean confidence ratings for self-knowledge (on a 0-100 scale), the mean predictions of other people’s knowledge (on a 0-100 scale), the actual vocabulary scores (%), and the calibration scores for both self and others.

**Table 2**

Confidence Ratings, Prediction for Other People's Performance and Accuracy Measures, by Experiment

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Experiment 1 | | | Experiment 2 | |
|  | Mean | | SD | Mean | SD |
| Confidence in self-knowledge | | 47.64 | 24.41 | 45.03 | 24.66 |
| Prediction of other people’s knowledge | | 44.67 | 19.13 | 76.91 | 17.78 |
| Vocabulary score (%) | | 58.01 | 22.33 | 53.29 | 19.43 |
| Vocabulary others (%) | | 58.01 | 22.33 | 72.91 | 27.31 |
| Self-calibration | | -10.37\*\* | 14.69 | -8.25\*\* | 16.56 |
| Other-calibration | | -13.34\*\* | 19.13 | 5.27 \*a | 17.80 a |

\* < .001, \* < .01; a Based on a different sample of participants with English-speaking parents.

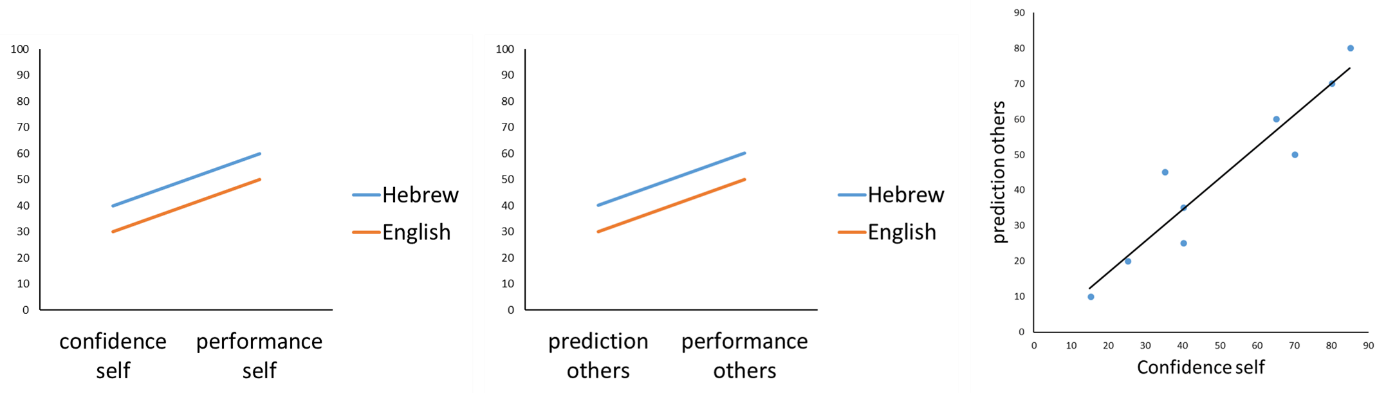
In Experiment 1, participants were less confident in their responses compared to their actual performance, resulting in a negative calibration score that differed from zero, *t* (88) = -6.66, *p* < .001, Cohen's *d* = .706. They were also underconfident regarding the performance of other Hebrew speakers, *t* (88) = -6.58, *p* < .001, Cohen's *d* = .697. These calibration scores were similar, *t* (88) = 1.46, *p* = .147, and associated (*r* = .379, *p* < .001). In Experiment 2, participants were also underconfident in their own performance, *t* (105) = -5.13, *p* < .001, Cohen's *d* = .499, but they were overconfident in the performance of other Hebrew speakers whose parents were English speakers, predicting they would perform the task better than they did. This overconfidence was reflected in a positive calibration score for others, which differed from zero as well, *t* (105) = 3.05, *p* = .003, Cohen's *d* = .297. In this case, self and others' calibration scores were different, *t* (105) = -5.394, *p* < .001, Cohen's *d* = .524, and they were not associated (*r* = -.13, *p* = .185).

These findings suggest that bilinguals are underconfident in their vocabulary knowledge in their non-dominant language, possibly due to their proficiency level but likely also because of the low familiarity of task items. The findings also indicate that the estimation of other people's knowledge may not depend solely on one's knowledge, as suggested by the Anchoring and Adjustment Theory, but also on the confidence in that knowledge and on the linguistic background of the other people. When participants perceive others as similar to themselves, they rely on their confidence in self-knowledge and are underconfident in other people's knowledge, just as they are in their own knowledge. However, when others are perceived as more proficient (e.g., having English-speaking parents), they disregard their confidence in self-knowledge and tend to be overconfident in those individuals' knowledge. The findings support the hypothesis that individuals rely on their confidence in their own language knowledge when predicting the knowledge of others, revealing biases in evaluating both their own and others' proficiency. This holds true regardless of whether the others have a similar or different linguistic background, highlighting the intricate dynamics at play in the process of proficiency estimation. However, it remains unclear whether these findings are limited to the non-dominant language and low-familiarity words or if they extend to the dominant language and to more familiar items. Additionally, it is also unclear whether employing a more complex task would yield similar results. Investigating these questions is crucial to deepen our understanding of the mechanisms underlying knowledge estimation in general and proficiency judgment in particular.

**Study 1: Effect of self-proficiency on self-confidence**

Study 1 will investigate how objective and perceived proficiency influence participants' confidence in their vocabulary knowledge, the alignment between confidence levels and actual knowledge, and how self-confidence impacts the estimation of others' knowledge in both the dominant and non-dominant languages. The study will include vocabulary tasks in Hebrew and in English, each consisting of 30 target words. The words will be selected from a larger pool of words, rated for familiarity and concreteness by a different sample of Hebrew-English bilinguals with similar demographic and linguistic backgrounds. To avoid floor and ceiling effects, the selected words in both languages will be relatively unfamiliar to native Hebrew speakers, and matched in terms of familiarity, concreteness, and length. A pretest will ensure that both tasks yield comparable vocabulary scores, with items replaced as needed.

Ninety participants will complete the vocabulary tests in both Hebrew and English. The order of languages will be counterbalanced across participants. For each item, participants will rate their confidence in their response on a 0-100 scale. They will also rate the chances that another Hebrew speaker will select the correct answer on a similar scale. The order of self and other ratings will be counterbalanced across participants. After completing the vocabulary task in each language, participants will perform the MINT-sprint in that language.

**** Figure 1 illustrates some of the expected results. Given the low familiarity of task items, and based on the preliminary findings, I expect that participants will demonstrate underconfidence in their own vocabulary knowledge (left panel) and in the knowledge of other Hebrew speakers (middle panel) in both languages. In addition, I expect to find a positive association between confidence in self-knowledge and prediction of other people's knowledge (right panel) in both languages. These findings would suggest that bilinguals rely more on their confidence in their own knowledge than on language dominance as a cue for estimating others' proficiency.

**Figure 1.** Confidence and actual performance (vocabulary scores) for self (left), prediction of others' performance and their actual performance (middle), by language, and association between self-confidence and prediction of others' knowledge (right).

**Study 2: Word familiarity**

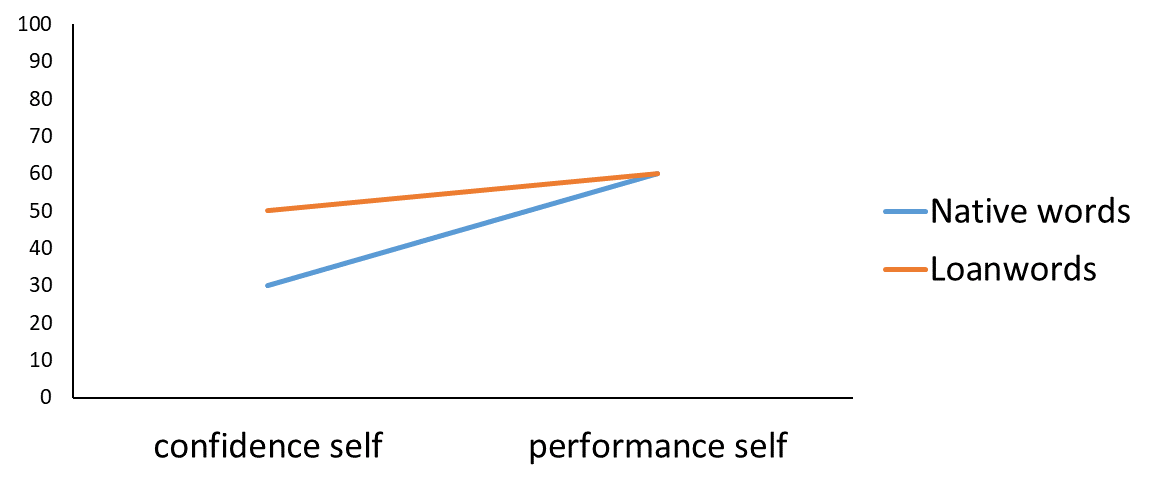
Study 2 will investigate whether word familiarity affects ease of processing (i.e., fluency) as well as confidence in self-knowledge and estimation of the knowledge of others. The design will be similar to the design of Study 1, but all items will be in English (i.e., the less dominant language). A new vocabulary task will be created, consisting of 45 English words: 15 highly familiar, 15 moderately familiar, and 15 unfamiliar. These words will be selected from a larger pool of words, rated for familiarity and concreteness by a separate group of Hebrew-English bilinguals. The selected words will be matched for concreteness and length to ensure that familiarity drives the observed effects rather than other lexical characteristics. Ninety participants will complete the task, estimating knowledge of other Hebrew speakers. The data will be analyzed similarly to Study 1, with familiarity treated as a repeated measure.

 I expect to find a familiarity effect, with faster and more accurate responses and higher confidence ratings for familiar words compared to moderately familiar, and unfamiliar words (see Figure 2). This effect will likely vary with language proficiency; such that lower proficiency will be associated with a stronger familiarity effect. Additionally, I anticipate that participants will be relatively well-calibrated for familiar words but underconfident in their performance for the other words, as observed in the preliminary study. If participants rely on their confidence in their own knowledge when estimating the knowledge of others, a similar pattern will emerge in their estimate of other people’s knowledge. These findings will suggest that familiarity with words interact with language proficiency to determine one's knowledge and confidence in that knowledge, which in turn affects the estimation of other people's vocabulary knowledge. Such results will highlight the interplay between individual differences in knowledge and contextual cues in estimating others' proficiency.

**Figure 2.** Familiarity effects in confidence ratings and actual performance (vocabulary score) for self-knowledge.

**Study 3: Cross-language similarities**

Study 3 will examine the effect of cross-language similarities on confidence in self-knowledge and on estimation of monolingual English speakers' knowledge. Ninety Hebrew-English bilinguals will choose the correct meaning of 15 English loanwords and 15 native words matched for familiarity, length, and concreteness. Participants will then rate their confidence in each response and estimate how likely monolingual English speakers are to select the correct response for each word. Words will be selected from a larger pool, rated for familiarity and concreteness by a separate group of Hebrew-English bilinguals. Loanwords will also be rated by another sample of Hebrew-English bilinguals for their likelihood of use in Hebrew (e.g., "Rate the chances of using the word 'gravitatsia' versus its Hebrew equivalent 'kvida' in Hebrew). Only loan words with a high probability of use (>70%) will be included. To assess prediction accuracy, a sample of monolingual English speakers with no knowledge of Hebrew will be asked to select the correct meaning of each word. Additionally, these participants will write down whether they know how to say each word in a different language to ensure they have only one representation of each word. The monolingual sample will be recruited through an American survey company.

**** If cross-language similarity creates a sense of familiarity (Costa et al., 2000; Masson, 2013), responses to loan words should be faster and more accurate than responses to native words, and should lead to higher levels of reported confidence (see Figure 3). If fluency is incorrectly attributed to objective ease when estimating the knowledge of monolingual English speakers, who have only one representation of these loan words (Birch et al., 2017), participants will demonstrate an estimation bias. This will suggest that bilinguals use cross-linguistic similarities as an additional cue when estimating the knowledge of other people. If no differences emerge between word types, it will suggest that bilinguals base their estimation on a single representation.

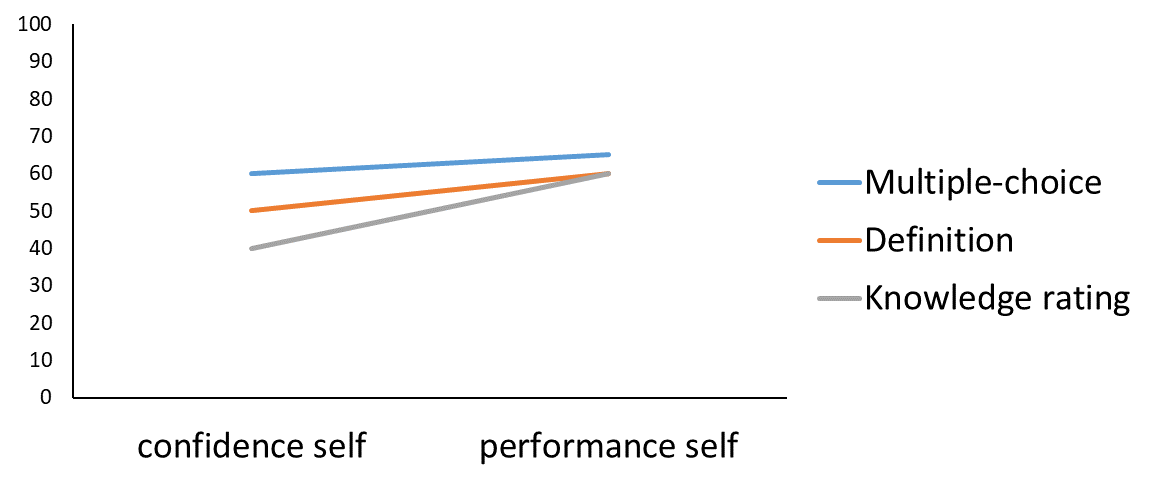
**Figure 3.** Confidence ratings and actual performance (vocabulary score) for self-knowledge of native words and loanwords.

**Study 4: Task type**

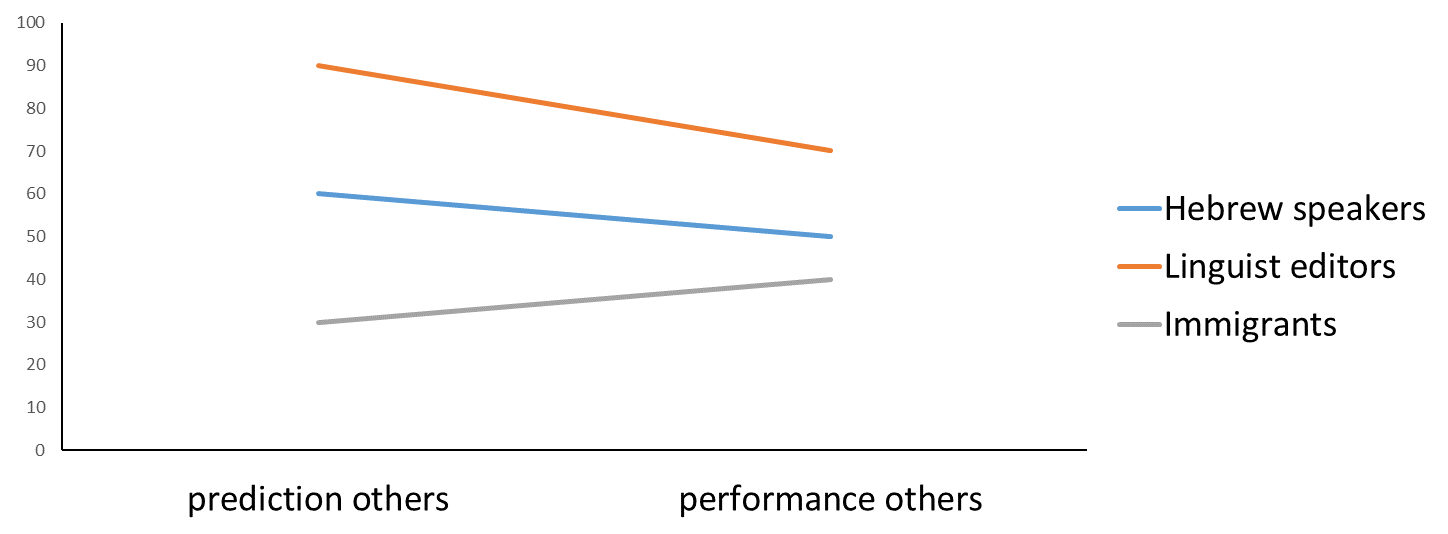
Study 4 will investigate whether task type affects the estimation of vocabulary knowledge. Two-hundred and seventy Hebrew-English bilinguals will be randomly assigned to one of three conditions. In the Multiple-choice condition, participants will complete the same Hebrew vocabulary task as in Study 1. In the Word-definition condition, participants will be asked to define the words, rate their confidence in their definitions, and predict the knowledge of others. In the third condition, participants will see a list of words and will be asked to rate their confidence in knowing each word's meaning and to predict the knowledge of other Hebrew speakers. After completing the task, participants in this condition will also define the words to measure self and other's actual score. The use of Hebrew vocabulary tests aims to ensure that any incorrect definitions reflect a lack of knowledge about the word, rather than a difficulty in writing the definition due to low English proficiency. Vocabulary scores will be determined by three proficient native Hebrew speakers, who will judge the definitions as correct (1 point) or incorrect (0 points).

If the nature of the task affects the estimation of self-knowledge, as suggested by Ackerman (2019) and Kavé and Yafé (2014), I expect the findings to differ across the three conditions. Multiple-choice tasks have been shown to yield higher confidence in performance (Ibabe & Sporer, 2004) and greater correspondence between self-evaluations of ability and objective performance (Zell & Krizan, 2014) than more open-ended tasks. Therefore, as shown in Figure 4, I expect confidence ratings for self-knowledge to be higher, and the self-calibration score to be closer to zero in the Multiple-Choice condition compared to the Word Definition condition. I also expect the third condition to yield the smallest correspondence between self-evaluations of ability, as participants will have the fewest cues reflecting their knowledge. I anticipate finding similar patterns in the estimation of other people's knowledge. These findings will suggest that task type can affect the estimation of both self- and others' knowledge. Understanding this effect is crucial for advancing research in this field, as it highlights how different assessment methods can yield varying results in knowledge estimation.

**Figure 4.** Confidence ratings and actual performance (vocabulary score) for self-knowledge across task types.

**Study 5: The linguistic background of other people**

Study 5 will examine how the linguistic background of the people whose knowledge is being estimated affects estimation. The study will be similar to the other studies in design, but it will include three experiments. In Experiment 1, participants will complete a vocabulary test comprised of 45 Hebrew words, varying in familiarity. After choosing the correct meaning of each word and rating confidence in each response, participants will be asked to predict the chances that other Hebrew speakers will select the correct meaning. This experiment will serve as a baseline. In Experiment 2, after rating self-confidence, participants will be asked to estimate the knowledge of language editors, reflecting higher level of proficiency. To measure estimation accuracy, data will be collected from a group of language editors. In Experiment 3, participants will estimate the knowledge of other Hebrew speakers who immigrated to Israel after age 15, reflecting lower level of proficiency. To measure estimation accuracy, data will be collected from participants who immigrated to Israel after age 15.

In Experiment 1, I expect participants to be less biased for familiar words but underconfident for unfamiliar words. Additionally, confidence in self-knowledge will correlate with confidence in the knowledge of others, consistent with findings from the preliminary study. In Experiment 2, participants will be overconfident when predicting the knowledge of language editors. In Experiment 3, participants will be overconfident on familiar words and either well-calibrated or underconfident in their estimation of the knowledge of others concerning unfamiliar words. Figure 5 shows the general pattern of expected results, with the understanding that the slope between prediction of others' performance and the actual performance will vary depending on word familiarity. The findings of these three experiments will contribute to our understanding of the way in which perceptions of other people’s proficiency shape estimations of their knowledge.

**Figure 5.** Prediction of other people's knowledge and their actual knowledge (vocabulary score) when the others are Hebrew speakers (similar to the participants), linguistic editors (representing higher proficiency), and immigrants (representing lower proficiency).

**3.4 Resources for Conducting the Research**

The resources required for this project include relevant software, research assistants, and access to Hebrew-English bilingual participants, Native Hebrew speakers whose parents are native English speakers, monolingual English speakers, language editors, and English-Hebrew bilinguals who immigrated to Israel after age 15. The Open University’s Psychology lab is equipped with all necessary software for conducting the experiments. Proficient Hebrew-English bilingual research assistants will prepare the tests and all rating scales, and will program and administer the experiments under the supervision of a postdoctoral fellow. Most participants will be Hebrew-English bilingual students who will receive course credit for their participation. Recruiting bilinguals with varying levels of English proficiency should be relatively easy, as most adults in Israel study English from elementary school onward. Participants with specific characteristics, such as bilinguals who immigrated to Israel after age 15, will be recruited through snowball sampling and will be compensated for their participation. Monolingual English speakers will be recruited through an American survey company.

**3.5 Expected results and pitfalls**

I expect to find biases in vocabulary estimations for both self and others, regardless of language dominance or task type. These biases will likely vary with language proficiency, word familiarity, cross-language similarity, and the linguistic background of the others. A potential pitfall is the focus on a single aspect of proficiency, namely vocabulary knowledge, although other factors, such as syntactic knowledge, may affect judgment of proficiency. However, this narrow focus will allow a more in-depth exploration of this crucial component of language proficiency. A further risk is the potential difficulty in recruiting enough participants with specific characteristics, such as language editors, although snowball sampling and monetary incentives should increase the likelihood of securing the necessary sample.

Ackerman, R. (2019). Heuristic cues for meta-reasoning judgments: Review and methodology. *Psihologijske Teme, 28*(1), 1-20.

Amano, T., Rios Rojas, C., Boum II, Y., Calvo, M., & Misra, B. B. (2021). Ten tips for overcoming language barriers in science. *Nature Human Behaviour, 5*(9), 1119-1122.

Ariffin, K., & Rafik-Galea, S. (2009). Code-switching as a communication device in conversation. *Language and Society Newsletter, 5*(9), 1-19.

Babayiğit, S., & Shapiro, L. (2020). Component skills that underpin listening comprehension and reading comprehension in learners with English as first and additional language. *Journal of Research in Reading, 43*(1), 78-97.

Birch S. A. J., Brosseau-Liard P. E., Haddock T., Ghrear S. E. (2017). A ‘curse of knowledge’ in the absence of knowledge? People misattribute fluency when judging how common knowledge is among their peers. *Cognition, 166*, 447–458.

Botes, E., Dewaele, J.-M. & Greiff, S. (2020). The power to improve: Effects of multilingualism and perceived proficiency on enjoyment and anxiety in foreign language learning. *European Journal of Applied Linguistics, 8*(2), 1–28.

Bullock, O. M., Colón Amill, D., Shulman, H. C., & Dixon, G. N. (2019). Jargon as a barrier to effective science communication: Evidence from metacognition. *Public Understanding of Science, 28*(7), 845-853.

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283-1296.

Damen, D., van der Wijst, P., van Amelsvoort, M., & Krahmer, E. (2020). Can the curse of knowing be lifted? The influence of explicit perspective-focus instructions on readers’ perspective-taking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(8), 1407-1423.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first-and secondlanguage word recognition: A lexical entrenchment account. *Quarterly Journal of Experimental Psychology, 66*(5), 843–863.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*(3), 175–197.

Dragojevic, M., Gasiorek, J., & Giles, H. (2016). Accommodation strategies as core of the theory. In H. Giles (Ed.), Communication accommodation theory: Negotiating personal and social identities across contexts (pp. 36–59). Cambridge University Press.

Epley, N., & Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self generated and experimenter-provided anchors. *Psychological Science, 12*(5), 391-396.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science, 17*(4), 311-318.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology, 87*(3), 327-339.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191.

Flores, G., Abreu, M., Olivar, M. A., & Kastner, B. (1998). Access barriers to health care for Latino children. *Archives of Pediatrics and Adolescent Medicine, 152*(11), 1119-1125.

Fitzsimmons, C. J., Thompson, C. A., & Sidney, P. G. (2020). Confident or familiar? The role of familiarity ratings in adults’ confidence judgments when estimating fraction magnitudes. *Metacognition and Learning, 15*, 215-231.

Garcia, D. L., & Gollan, T. H. (2022). The MINT Sprint: Exploring a fast administration procedure with an expanded multilingual naming test. *Journal of the International Neuropsychological Society, 28*(8), 845-861.

Gasiorek, J., Dragojevic, M., & Vincze, L. (2022). Perspective-taking and language competence as predictors of language accommodation by adolescents from monolingual and bilingual households. *International Journal of Bilingual Education and Bilingualism, 25*(1), 148-155.

Geng, Y., Song, Q., & Fei, X. (2023). Factors in cognitive processing of Japanese loanwords by advanced Chinese Japanese-as-a-foreign-language learners. *Frontiers in Psychology, 14*, 1224830.

Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost–benefit analysis of voluntary language switching in young and aging bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 640-665.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: language and cognition, 15*(3), 594-615.

Gotlieb, R., Praska, C., Hendrickson, M. A., Marmet, J., Charpentier, V., Hause, E., Allen, K. A. Lunos, S. & Pitt, M. B. (2022). Accuracy in patient understanding of common medical phrases. *JAMA Network Open, 5*(11), e2242972-e2242972.

Gyllstad, H., Vilkaitė, L., & Schmitt, N. (2015). Assessing vocabulary size through multiple-choice formats: Issues with guessing and sampling rates. *ITL-International Journal of Applied Linguistics, 166*(2), 278-306.

Ibabe, I., & Sporer, S. L. (2004). How you ask is what you get: On the influence of question form on accuracy and confidence. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 18*(6), 711-726.

Kapiley, K., & Mishra, R. K. (2024). Language contexts induced by the interlocutors’ proficiencies modulate bilingual language monitoring. *Bilingualism: Language and Cognition*, 1-14.

Kavé, G., & Halamish, V. (2015). Doubly blessed: older adults know more vocabulary and know better what they know. *Psychology and Aging, 30*(1), 68-73.

Kavé, G., & Yafé, R. (2014). Performance of younger and older adults on tests of word knowledge and word retrieval: Independence or interdependence of skills? *American Journal of Speech-Language Pathology, 23*(1), 36–45.

Keysar, B. (1994). The illusory transparency of intention: Linguistic perspective taking in text. *Cognitive Psychology, 26*(2), 165-208.

Keysar, B., & Bly, B. (1995). Intuitions of the transparency of idioms: Can one keep a secret by spilling the beans?. *Journal of Memory and Language, 34*(1), 89-109.

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349–370.

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review, 119*(1), 80-113.

Kumar, A. P., Nayak, A., Shenoy, M., & Goyal, S. (2023). A novel approach to generate distractors for multiple choice questions. *Expert Systems with Applications, 225*, 120022.

Lau, B. K. Y., Geipel, J., Wu, Y., & Keysar, B. (2022). The extreme illusion of understanding. *Journal of Experimental Psychology: General, 151*(11), 2957–2962.

Liu, M. (2006). Anxiety in Chinese EFL students at different proficiency levels. System, 34(3), 301-316.

MacIntyre, P. D., Noels, K. A. & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning, 47*(2), 265-287.

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research, 50*(4), 940-967.

Markovits, H., Thompson, V. A., & Brisson, J. (2015). Metacognition and abstract reasoning. *Memory and Cognition, 43*, 681-693.

Masson, M. E. (2013). How L1 loanwords can create a false sense of familiarity with L2 vocabulary meaning and usage. *Vocabulary Learning and Instruction, 2*(1), 8-14.

Monaghan, P., Chang, Y. N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language, 93*, 1-21.

Mor, B., & Prior, A. (2020). Individual differences in L2 frequency effects in different script bilinguals. *International Journal of Bilingualism, 24*(4), 672-690.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), The psychology of learning and motivation: Advances in research and theory (pp. 125–173). Academic Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(3), 435–451.

Segal, D. (2023). Sustained attention plays a critical role in reading comprehension of adults with and without ADHD. *Learning and Individual Differences, 105*, 102300.

Segal, D., & Gollan, T. H. (2018). What’s left for balanced bilinguals? Language proficiency and item familiarity affect left-hemisphere specialization in metaphor processing. *Neuropsychology, 32*(7), 866-879.

Segal, D., & Kavé, G. (2024). Perspective judgment across adulthood: Evidence from bilinguals. *Psychology and Aging, 39*(3), 324-336.

Sharon, A. J., & Baram-Tsabari, A. (2014). Measuring mumbo jumbo: A preliminary quantification of the use of jargon in science communication. *Public Understanding of Science, 23*(5), 528-546.

Shulman, H. C., Dixon, G. N., Bullock, O. M., & Colón Amill, D. (2020). The effects of jargon on processing fluency, self-perceptions, and scientific engagement. *Journal of Language and* Social *Psychology, 39*(5-6), 579-597.

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. *Learning and Instruction, 51*, 61-73.

Stell, G., & Dragojevic, M. (2017). Multilingual accommodation in Namibia: An examination of six ethnolinguistic groups’ language use in intra-and intergroup interactions. *Journal of Language and Social Psychology, 36*(2), 167-187.

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2019). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition, 22*(3), 516-536.

Tullis, J. G. (2018). Predicting others’ knowledge: Knowledge estimation as cue utilization. *Memory and Cognition, 46*(8), 1360-1375.

Tullis, J. G., & Feder, B. (2023). The “curse of knowledge” when predicting others’ knowledge. *Memory and Cognition, 51*(5), 1214-1234.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. Science, 185(4157), 1124-1131.

Wilson, E., Chen, A. H., Grumbach, K., Wang, F., & Fernandez, A. (2005). Effects of limited English proficiency and physician language on health care comprehension. *Journal of General Internal Medicine, 20*(9), 800-806.

Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science, 9*(2), 111-125.