Nonverbal Cues of Meta-Memory Awareness in Older Adults

Mandy Visser (Mandy.Visser@tilburguniversity.edu)

Marie Postma-Nilsenova (M.Nilsenova@tilburguniversity.edu)

Emiel Krahmer (E.J.Krahmer@tilburguniversity.edu)

Marc Swerts (M.G.J.Swerts@tilburguniversity.edu)

Tilburg center for Cognition and Communication (TiCC), Tilburg University, The Netherlands

Abstract

This study aimed to examine nonverbal expressions of older adults performing a Feeling-of-Knowing (FOK) meta-memory task, since nonverbal cues are considered to serve as indicators of memory deficit awareness. In a production experiment, we collected a variety of recalled and unrecalled answers from older adults (mean age = 79.5) and tested their accuracy. Nonverbal behavior was annotated manually and automatically using facial expression detection software. We found an overall effect of FOK ratings on the use of FOK related nonverbal features. For recalled items, the participants used more nonverbal cues with lower FOKs than with higher FOKs. For unrecalled items, the opposite effect was found. A subsequent perceptual study showed that third-party judges were able to estimate older adults' FOK correctly. Overall, this study shows that the elderly can be aware of their memory deficits and display the associated nonverbal cues in a manner comparable to younger age groups.

Keywords: Meta-memory; Feeling-of-Knowing; nonverbal cues; older adults.

Introduction

Oftentimes, when asked a question, we can reliably estimate whether we know the answer or not even before actually retrieving it from our memory. This type of guess is referred to as "Feeling-of-Knowing (FOK)" (Hart, 1965) and is considered to be a form of meta-memory comparable to the well-known Tip-of-the-Tongue phenomenon. Various studies have shown that FOK levels correlate with both verbal and non-verbal expressiveness displayed when responding to questions (e.g., Krahmer & Swerts, 2005; Smith & Clark, 1993). These findings indicate that the presence or absence of particular non-verbal features, employed in the context of monitoring one's memory, signal an individual's awareness of memory deficits.

Research on children and younger adults showed a link between FOK accuracy and nonverbal expressiveness (Krahmer & Swerts, 2005) but there are at least three reasons to assume that the findings cannot be extended to older population. First, in older age groups, FOK accuracy appears to decrease at least for some types of knowledge, like in episodic memory tasks (e.g., Souchay, Moulin, Clarys, Taconnat & Isingrini, 2007). The decrease in accuracy may, in turn, affect the accompanying non-verbal behavior. Second, older adults are arguably less expressive in their non-verbal behavior. Although findings on expressing positive emotions are mixed, negative emotions like fear and anger are found to be less intense in elderly, compared to young adults (Carstensen, Pasupathi, Mayr & Nesselroade, 2000; Levenson, Carstensen, Friesen & Ekman, 1991), possibly thanks to a better-developed emotion regulation mechanism (Charles & Carstensen, 2007). Third, the use of non-verbal expressions by the elderly may be affected by the higher frequency of the experienced retrieval failures and tip-of-the-tongue states (Gollan & Brown, 2006). Older adults are more prone to word retrieval failures than younger adults, for example due to their greater vocabulary knowledge (Burke, MacKay, Worthley & Wade, 1991).

In the research reported here, we set out to obtain a comprehensive overview of the non-verbal cues displayed by older adults, using a variant of the FOK paradigm, as applied in earlier question-answering studies (e.g., Smith & Clark, 1993). In the original paradigm due to Hart (1965), also referred to as the recall-judgment-recognition paradigm, participants are exposed to a three-step procedure starting with a series of general knowledge questions. In response to these questions (i.e., "What is the capital of Switzerland?"), participants are either able to recall the answers ("Bern") or not ("I don't know"). Subsequently, for the unrecalled items, participants are asked to judge whether or not they believed they would be able to recognize the correct answer among several wrong alternatives, e.g., when presented in a multiple-choice test. In the last part of the procedure - the recognition – they are given a multiple-choice test and asked to select the correct answers to the previously queried items. Hart (1965) referred to the participants' judgment elicited during the second step in the procedure as their Feeling-of-Knowing. FOKs serve as assessments that information is available in memory, even when it has not been retrieved (e.g., Eakin & Hertzog, 2012; Hart, 1965; Singer & Tiede, 2008).

During the course of answering questions, people undergo several alternating processes. While being questioned, people actively search in their mind for the correct answer. Simultaneously, this retrieval is monitored on a meta-cognitive level. This means that while formulating or searching for the correct answer, people are continuously evaluating whether they are capable to answer the question

correctly or not (e.g., Nelson & Narens, 1990). FOK can be used as a guidance for monitoring the search for a correct answer and can help in deciding to continue the search or resign oneself to an unsuccessful retrieval.

Traditionally, in psychological studies, FOK is used to describe solely *prospective* memory tasks, i.e., the feeling of being able to recognize a correct answer for unrecalled items (Hart, 1965). In psycholinguistic literature, on the other hand, FOK is typically examined on a par with indications of confidence (e.g., Brennan & Williams, 1995). In this case, FOK is used to describe both *prospective and retrospective* memory tasks, and refers to the participant's estimate of being able to recognize the correct answer both for recalled and unrecalled items. Since the outcomes of psycholinguistic studies on (non-)verbal expressions of FOK show similarities between signals of high FOK for unrecalled items and low FOK for recalled items, this study investigates FOK in both contexts.

Past studies of non-verbal cues associated with FOK identified specific visual and auditory cues displayed during the task. For instance, a high FOK experienced for previously unrecalled items is typically signaled by auditory cues including *linguistic hedges* and *fillers* such as 'perhaps' (Smith & Clarke, 1993) and 'um' (Corley & Stewart, 2008), as well as by visual cues, like *averted gaze* or *brow movements* (Swerts & Krahmer, 2005). Similar cues appear to be used for low FOK in the case of recalled items, together with *high-rise terminals* (Scherer, London & Wolf, 1973), or visual cues like *smiles* and *"thinking faces"* (Swerts & Krahmer, 2005).

In the study reported here, we set out three objectives. First, we explore older adults' metamnemonic awareness in relation to their use of nonverbal cues for varied degrees of FOK experiences. To our knowledge, the nonverbal behavior accompanying FOK in an older age group has not been studied before, despite the fact that there are reasons to believe it may differ from younger age groups, as outlined above. In this study, we will code nonverbal features manually, based on a coding scheme used in earlier FOK studies (e.g., Swerts & Krahmer, 2005). Given that this existing set of cues is identified for younger age groups, we will also analyse the visual features with the help of a comprehensive automatic procedure. Second, we compare older adults' FOK accuracy and nonverbal cues of *recalled* items (answers) and *unrecalled* items (non-answers), where earlier FOK studies seem to focus merely on either recalled or unrecalled items, or merge the two categories. As a third and final objective, we explore the decoding of FOK cues displayed by the older participants by third-party judges (in the literature referred to as "Feeling-of-Another's-Knowing", FOAK for short; see Brennan & Williams, 1993). Assuming that the non-verbal cues associated with FOK fulfil, at least partly, social functions, we expect them to be recognized by independent observers.

Production Experiment

Method

Participants In total, 24 participants (12 female) took part in the production experiment. Prior to the analysis, the data of one of the female participants had to be discarded due to a recording error, resulting in a sample of 23 participants. They were recruited in a nursing home in Rotterdam (N = 16) and an activity centre in Tilburg (N = 8) in The Netherlands. Participants' age ranged from 70 to 95 years (M = 79.5, SD = 6.3) and according to their caretakers, they did not suffer from any major cognitive impairment. Beforehand, participants signed a consent form by which they gave permission to be filmed during the experiment and for the recordings to be used for scientific purposes.

Stimuli Similar to earlier FOK studies (e.g., Smith & Clarke, 1993), participants were asked knowledge questions in a quiz-like setup. In order to collect a substantial amount of lexically distinct answers, while keeping the duration of the experiment within reasonable limits, participants were exposed to one of two question sets, each of which contained twenty knowledge questions, selected from Trivial Pursuit board games. Both clusters of questions resulted in answers that were likely to be either easy or hard to retrieve. To prevent feelings of frustration, participants were assured beforehand that the range of question difficulty varied and they were not expected to be able to answer all questions correctly. The question sets were pretested with four older adults (65 to 92 years old) who were not a part of the experimental group. They provided both answers and non-answers to the questions in the sets.

Experimental Procedure The production experiment took place in environments familiar to the participants. Following the FOK paradigm (Hart, 1965), participants underwent a three-step procedure, without any time restrictions.

First, participants were asked to answer a series of twenty questions, presented orally to them by the experimenter. Participants sat in a chair in front of a video camera that recorded them during the experiment. The experimenter was positioned behind the camera and aimed to respond to comments of the participants as less as possible, except repeating a question if needed. In this way, participants were unable to pick up any feedback cues about the (in)correctness of their answers. In the second part of the experiment, the participants were given a paper form, which listed the exact same sequence of questions. For each question, participants were asked to indicate (on a seven-point Likert scale) how sure they were that they would recognize the correct answer if it was presented in a multiple-choice test (the FOK score). In the third and last part of the experiment, the same set of questions was presented again, as a multiple-choice test in which the correct answer was mixed with three plausible alternatives. Participants were urged to respond to every question, even if this meant they had to guess.

Results

Types of Responses All twenty-three participants answered twenty questions; in total, 460 utterances were collected in the production experiment. Recordings contained correctly recalled items, incorrectly recalled items and unrecalled items, see Table 1. The majority of the FOK ratings were of level 7, in line with the attempt to make most of the questions easy to answer to prevent participants' discomfort, leaving a sufficient number of lower FOK ratings to be used in follow-up measurements.

FOK and Recall Analysis of variance showed that mean FOK ratings were higher for recalled items than for unrecalled items (with participants as random factor, $F1(1,22) = 2.36$, $p < .001$, $\eta^2_p = .11$; with items as random factor: $F2(1,19) = 2.33$, $p < .001$, $\eta^2_p = .10$). Moreover, participants indicated higher FOK ratings for correctly recalled items than for incorrectly recalled items, ($F1(1,21) = 2.18$ $p < .01$, $\eta^2_p = .17$; $F2(1,17) = 2.20$, $p < .05$, $\eta^2_p = .14$). Average FOK ratings as a function of different answer categories are presented in Table 1.

Table 1: Average FOK ratings for different response categories.

Response category		N	*M*	*SD*
Open questions	All answers	267	6.35	1.25
	Corrects answers	192	6.79	.65
	Incorrect answers	75	5.23	1.64
	All non-answers	193	2.88	2.16
Multiple choice questions	Correct answers	311	5.65	2.13
	Incorrect answers	149	3.33	2.20

FOKs for Unrecalled Items Only In order to establish the accuracy of the FOK judgments, we compared the FOK ratings of unrecalled items that were correctly recognized, to the FOK ratings of incorrectly recognized unrecalled items. A T-test for independent samples revealed a significant difference between the two groups, $t(183.95) = 2.88$, $p = .004$ (equal variances not assumed). The FOK ratings were higher for correctly recognized unrecalled items ($M = 3.32$, $SD = .51$) than for the incorrectly recognized ones ($M = 2.44$, $SD = .39$), indicating that the elderly participants were accurate at predicting the recognition outcome.

FOK and Nonverbal Cues All 460 utterances were manually transcribed and categorically coded for the presence or absence of the auditory and visual features as described in table 2, based on earlier work of Smith and Clark (1993) and Krahmer and Swerts (2005). With respect to vocal features, Brennan and Williams (1993) found correlations between FOK and the use of *delays, fillers* and *high intonation*, when answering a question. Similar to the study of Krahmer and Swerts (2005), we based the three visual features on the Facial Action Coding System by Ekman and colleagues (e.g., Ekman & Rosenberg, 1997). In this system, facial expressions are described by means of Action Units (AUs), i.e., numbered muscular actions: *smiling* is related to AU 12, 13 and/or 14; *eye brow movement* is related to AU 1 and/or 2; and a *puzzled face* is related to AU 14, 15, 18, 20 and/or 24, which describe *lip movements*, like *lip pucker* and *dimpler* in combination with AU 1, 2, and/or 5, which describe *eyebrow movements*, and AU 9 for a *nose wrinkle*. For representative examples of visual features used by participants, see figure 1.

Following an explicit labeling protocol, two independent coders labeled part of the data (a standard 15%) with fairly acceptable inter-coder agreements (Cohen's Kappa's were .86 for fillers, .72 for *high intonation*, .69 for *delays*, .69 for *eyebrow movements*, .78 for *smiling* and .65 for *puzzled faces*); the remaining utterances were labeled by one coder. Both coders were blind to FOK ratings and the questions preceding the utterances.

Table 2: Description of coded features.

Feature:	Description
Filler	The use of fillers (like "um", or "I'm not sure, but I think this is…").
Intonation	Ending an answer with a high boundary tone.
Delay	A silence > 1 sec, preceding an answer.
Eye brow movement	Moving (one of) the eyebrows from neutral position.
Smiling	Moving the corners of the lips upwards.
Puzzled face	Combining brow movements, possible nose wrinkle and lip pucker/dimpler.

FOK and Expressivity An analysis of variance showed an overall effect of FOK ratings on the number of nonverbal features used in recalled items, $F(1, 266) = 11.75$, $p < .001$, $\eta^2_p = .21$; with lower FOKs, participants used a larger amount of nonverbal features than with higher FOKs. For unrecalled items, there was an opposite effect, $F(1, 266) = 5.99$, $p < .001$, $\eta^2_p = .16$; participants used a larger number of features for high FOK non-answers than for low FOK non-answers.

Figure 1: Stills illustrating the coded features (from left to right: eyebrow movement, smiling and puzzled face).

Specifying Nonverbal Cues The labelled features were analysed individually by means of paired sample T-tests for all items, comparing the FOKs in the presence and absence of a feature. Table 3 shows that the presence of the nonverbal features in recalled items corresponds with a lower FOK rating, with the exception of *eyebrow movement* and *smiling*. Contrasting results are shown in table 4, which displays the presence and absence of nonverbal features in unrecalled items. For delay and high intonation, the FOKs were higher when the nonverbal feature was present compared to when it was absent.

Table 3: Mean individual FOK ratings for recalled items as a function of presence and absence of FOK nonverbal features (N representing the number of participants that could be used to calculate individual means).

	N	Present	Absent	Difference
Filler	22	6.04 (0.66)	6.54 (0.81)	-0.50 (1.00)*
Delay	17	5.37 (1.50)	6.45 (0.47)	-1.08 (1.39)**
Intonation	20	5.35 (1.62)	6.52 (0.52)	-1.70 (1.52)**
Eyebrow	22	6.07 (0.84)	6.43 (0.64)	-0.36 (0.99)
Smile	15	6.21 (0.71)	6.45 (0.41)	-0.24 (0.84)
Puzzled face	8	4.13 (1.89)	6.13 (0.74)	-2.01 (2.15)*

*$p < .05$, **$p < .01$

Table 4: Mean individual FOK ratings for unrecalled items as a function of presence and absence of the FOK nonverbal features (N representing the number of participants that could be used to calculate individual means).

	N	Present	Absent	Difference
Filler	19	3.56 (1.57)	2.25 (1.40)	1.31 (1.89)*
Delay	19	4.32 (1.72)	2.66 (1.36)	1.67 (2.09)*
Intonation	4	4.50 (2.38)	3.21 (1.90)	1.29 (1.67)
Eyebrow	20	3.14 (1.43)	2.95 (1.42)	0.19 (1.36)
Smile	13	3.70 (1.60)	3.19 (1.36)	0.50 (1.93)
Puzzled face	17	3.43 (2.29)	2.92 (1.19)	0.51 (1.90)

*$p < .05$, **$p < .01$

Automatic Analysis

For the automatic analysis, we used the software tool for frame-based automatic facial expression recognition CERT (Computer Expression Recognition Toolbox; Littlewort et al., 2011). Based on a machine-learning algorithm, the tool identifies the face region in a video and detects with a reasonably high accuracy (comparable to human annotators) the 44 Facial Action Units in the Facial Action Coding System (Ekman & Rosenberg, 1997). In total, 440 video fragments were analysed for the averaged probability of a particular facial action unit being present in the fragment. Twenty fragments were discarded because the software was unable to detect the facial region reliably. The set of fragments contained both recalled (57%) and unrecalled (43%) items.

In order to explore a possible link between FOK and different Action Units, we performed a multiple regression analysis for the two conditions (recalled, unrecalled) separately, see Table 5. As suggested by Berry (1993), we excluded all weakly correlated variables prior to the analyses, ignoring Action Units with correlations < .3 between a given unit and FOK, leaving *dimpler* (-.31), *chin raise* (-.29), *lip pucker* (-.40), *lips part* (.28) and *fear brow* (.41) for recalled items and *nose wrinkle* (-.27), *dimpler* (-.29) and *lids tight* (-.26) for unrecalled items. The correlation analyses showed no correlation coefficients exceeding > .9 between the Action Units selected as predictors in the regression models, thus satisfying the assumption regarding multi-collinearity.

The regression analysis for recalled items revealed a significant effect of the *dimpler, lip pucker, lip parting*, and the *fear brow*. For unrecalled items, the *dimpler* and *nose wrinkle* were significantly related to the FOK score. In addition to the AUs associated with the puzzled face used in the manual labelling (*nose wrinkle, dimpler* and *eye brow movement*), the comprehensive automatic analysis thus helped to identify other cues associated with FOK for recalled items, especially in the lip area. Moreover, different cues appear to be predictive of the FOK ratings for recalled items compared to unrecalled items.

Table 5: Linear Regression Models Predicting FOK by Facial Action Units

	Recalled Items		Unrecalled Items	
	B	*SE*	*B*	*SE*
(Constant)	11.038	0.859	7.928	0.739
Fear Brow	1.739***	0.286	-	-
Lip Pucker	-38.991***	5.993	-	-
Lips Part	0.961***	0.254	-	-
Dimpler	-0.849***	0.208	-1.932***	0.351
Chin Raise	0.117	0.220	-	-
Nose Wrinkle	-	-	-9.660**	1.833
Lids Tight	-	-	-0.385	1.155
R^2	.40	**	.20	**
F Change	31.29	**	15.630	**

*$p < .05$, **$p < .01$

Perception Experiment

In the final part of the research reported here, we examined to what extent the nonverbal cues displayed by the older participants in the production experiment could be interpreted as cues to FOK by third-party judges (Feeling of Another's Knowing, FOAK, Brennan & Williams, 1993).

Method

Participants Forty-two younger adults participated as third-party judges in this perception study (24 women, age *M* =

22.8, *SD* = 6.0). All participants were students of Tilburg University and received course credits for their input.

Stimuli In total, sixty-four utterances were selected from the corpus collected in the production study, with an equal distribution of recalled and unrecalled items, and low and high FOK ratings. Given the individual differences in the use of the FOK scale, the lowest or second lowest (or highest and second highest) score for all answers of that particular participant were used as instances of a low (or high) FOK score. Note that the selection of low (and high) FOK utterances could differ between participants, given that individuals differed in what they experienced as difficult or easy questions. This gave a 2 × 2 design (high/low FOK × recalled/unrecalled item). The stimuli for the perception test were randomly selected, but utterances were iteratively replaced until the following criteria were met: the answers given in the selected clip had to be lexically different from each other to avoid that participants in the perception test would have to judge clips with similar content and the speaker should appear in clips representing all four conditions (recalled/unrecalled answer x high/low FOK). To assure judgments were only based on the (non-)verbal expression of the speaker, and not by the participants' own estimation of the correctness of answers, stimuli were presented without the questions that preceded answers.

Experimental Procedure Participants were placed in front of a computer screen in an isolated booth. On the screen, two sets of thirty-two stimuli (recalled and unrecalled items) were presented one by one. First, the set containing only recalled items was shown in one of two random orders. Participants saw the stimulus ID (1 to 32) and then the actual stimulus. During a stimulus-interval of three seconds, participants were instructed to estimate to what extent speakers were certain about their answer, on a seven-point Likert scale (the FOAK score). When participants finished this first set, a second set of thirty-two stimuli was presented in one of two random orders, containing only unrecalled items. A stimulus ID was presented (33-64) before the actual stimulus and the three seconds stimulus interval. Participants were asked to estimate the chance that the speaker would recognize the correct answer when the question would have been presented as a multiple-choice question instead. Participants were to judge this on a seven-point Likert scale (again the FOAK score). To get familiar with both tasks, participants practised with example stimuli beforehand.

Results

We conducted two repeated measures analyses with the participants' judgment scores (FOAK) as dependent variable and FOK (high or low FOK) as factor, for both recalled and unrecalled items. Participants were able to distinguish between speakers' high and low FOK for recalled items, $F(1, 41) = 976.28$, $p < .001$, $\eta^2_p = .96$. Speakers' high FOK recalled items were judged as more certain than speakers' low FOK recalled items (high FOK: $M = 5.39$, $SD = .43$; low FOK: $M = 2.83$, $SD = .47$). A comparable effect was found with respect to FOK for unrecalled items, $F(1, 41) = 403.93$, $p < .001$, $\eta^2_p = .91$. This means that speakers were judged as more capable of recalling a correct answer when presented in a multiple choice test, when they responded with a high FOK unrecalled item than when they responded with a low FOK unrecalled item (high FOK: $M = 4.33$, $SD = .56$; low FOK: $M = 2.40$, $SD = .54$).

Discussion and Conclusion

This study aimed to examine older adults' accuracy of FOK experiences and their use of nonverbal cues for both recalled and unrecalled items. Therefore, we conducted a production experiment with older adults, in which we collected recordings of a variety of answer utterances and accompanying FOK ratings. We coded all utterances for the presence of various cues (manually and automatically) and presented a selection of utterances to third-party judges in a perception test.

The results of our study support the view that the elderly are able to produce an accurate FOK for both recalled and unrecalled items. With regards to recalled items, participants indicated higher FOK ratings for correctly recalled items than for incorrectly recalled items. Additionally, FOK ratings were higher for correctly recognized unrecalled items than for the incorrectly recognized ones, which is in line with studies by Hertzog and colleagues (e.g., 2012), but contradicts earlier results found by Souchay et al. (2007). Overall, these results indicate that the elderly may be as accurate at assessing their performance in a metamnemonic task as younger age groups (e.g., Hart, 1965; Krahmer & Swerts, 2005).

With respect to nonverbal cues associated with FOK, we expected the older participants to signal their FOK differently than younger age groups. In particular, earlier research has shown that emotional expressiveness appears to decrease with age (Carstensen et al., 2000; Gross et al., 1997; Levenson et al., 1991). Therefore, one might expect older adults to suppress their FOK expressions. In general, our manual coding study showed an overall effect of FOK ratings on the use of FOK related nonverbal features in recalled items. More specifically, with lower FOKs, older participants used more FOK cues (which were identified with earlier studies of younger adults and children), than with higher FOKs. For unrecalled items, we found the opposite effect; participants were more expressive for high FOK non-answers than for low FOK non-answers. We can conclude that similarly to younger age groups, the elderly tend to display cues to low FOK, despite their lower emotional expressiveness. According to Charles and Carstensen (2007), the decline in expressiveness is caused by a better-developed emotion regulation system by older adults, which would explain why they express their FOK similarly to younger adults. FOK expressions are argued to have a self-presentational, face-saving nature: expressing a

low FOK might lower recipients' expectations regarding the speaker's mnemonic performance (Smith & Clark, 1993).

The specific cues older adults use to signal their FOK for recalled items include the presence of *fillers*, *delay, high intonation* and *puzzled face*, according to the manual coding. In addition, the automatic analysis helped to detect movements involving the lip area, i.e*., lip pucker* and *lip parting*. In the case of unrecalled items, for delay and high intonation, the FOKs were higher when the nonverbal feature was present compared to when it was absent. Even though the manual labelling did not identify any relevant visual features, with the help of the automatic analysis we found the effect of a *nose wrinkle* and *dimpler* for unrecalled items as well, thus adding to the list of FOK cues.

Finally, we examined how older adults' FOK is perceived by third-party judges, by using the Feeling-of-Another's-Knowing (FOAK) paradigm. Ours study showed that signals of FOK seem to be perceived as such, as the judges were able to estimate speakers' FOK correctly. These results are similar to FOAK studies with younger age groups (Brennan and Williams, 1993; Krahmer & Swerts, 2005).

To conclude, this study shows that the elderly can be aware of their memory deficits and display the associated nonverbal cues in a manner comparable to younger age groups. Future FOK studies can distinguish between different functions of the nonverbal cues and their effect on third-party judgments. In particular, it could be the case that some expressions are automatic and primarily associated with the affective (e.g., movements of the lips) and cognitive states experienced by the participant (e.g., eye brow movement), while others serve a more communicative function (e.g., hedges, filled pauses).

Acknowledgments

We would like to thank all participants, especially the older adults in the production experiment. Thanks to Eric Postma for technical support, and to Tineke van den Hoek, Elise Peters and Manuela Luitjes for their help with the data collection.

References

Berry, W. D. B. (1993). *Understanding Regression Assumptions*. Sage Publications.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language, 34,* 383–398.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language, 30,* 542–579.

Carstensen, L. L., Pasupathi, M., Mayr, U., & Nesselroade, J. R. (2000). Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology*, *79*(4), 644.

Charles, S. T., & Carstensen, L. L. (2009). Emotion Regulation and Aging. In J. J. Gross, (Ed.), *Handbook of Emotion Regulation*. Guilford Press.

Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass, 2,* 589–602.

Eakin, D., & Hertzog, C. (2012). Age invariance in feeling of knowing during implicit interference effects. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences,* 67(5), 555-562.

Ekman, P., & Rosenberg, E. L., (1997). *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press.

Gollan, T. H., & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology,135*(3), 462-483.

Gross, J. J., Carstensen, L. L., Pasupathi, M., Tsai, J., GotesUm-Skorpen, C, & Hsu, A. (1997). Emotion and aging: Experience, expression, and control. *Psychology and Aging, 12,* 590-599.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56,* 208–216.

Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and Speech, 48,* 29–53.

Levenson, R. W., Carstensen, L. L., Friesen, W. V., & Ekman, P. (1991). Emotion, physiology, and expression in old age. *Psychology and Aging, 6,* 28-35.

Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011). The computer expression recognition toolbox (CERT). In *Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on* (pp. 298-305). IEEE.

Nelson, T., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26,* 125–141.

Scherer, K.R., London, H., & Wolf, J. J. (1973). The voice of confidence: Paralinguistic cues and audience evaluation. *Journal of Research in Personality, 7,* 34-44.

Singer, M., & Tiede, H. (2008). Feeling of knowing and duration of unsuccessful memory search. *Memory & Cognition, 36,* 588–597.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language, 32,* 25–38.

Souchay, C., Moulin, C. J. A., Clarys, D., Taconnat, L., & Isingrini, M. (2007). Diminished episodic memory awareness in older adults: Evidence from feeling-of-knowing and recollection. *Consciousness and Cognition, 16,* 769–784.

Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language, 53,* 81–94.