

Listening Effort: Order Effects and Core Executive Functions

DOI: 10.3766/jaaa.17024

K. Jonas Brännström*
Elisabeth Karlsson*
Sebastian Waechter*
Tobias Kastberg*

Abstract

Background: Listening effort seems to depend on input-related listening demands and several factors internal to the individual listener. Input-related demands may be listening in noise compared with listening in quiet, and internal factors may be cognitive functions.

Purpose: The purpose was to apply measures of listening effort and perceived listening effort in participants with normal hearing, to determine if there are any presentation order effects, and to explore the relationship between listening effort measured as accuracy, response times, efficiency of information encoding into long-term memory, perceived listening effort, and core executive functions.

Research design: A within-subject design with repeated measures was used and a study of relationships between variables was made.

Study sample: Thirty-two healthy adults with normal hearing.

Data collection and analysis: Participants were tested individually by a listening task using a dual-task paradigm. The listening task was performed in quiet and in multitalker babble noise at 10 dB signal-to-noise ratio (SNR). Perceived listening effort and core executive functions (working memory, inhibitory control, and cognitive flexibility) were also assessed.

Results: The measures of listening effort (correct responses, response times, and immediate and delayed listening comprehension) failed to demonstrate increased listening effort in multitalker babble noise (10 dB SNR) compared with quiet, although a significant test order effect was seen for correct responses indicating that participants who first listened in noise did not improve in quiet. Perceived listening effort increased significantly in noise compared with quiet. No relationship was found between measures of listening effort and ratings of perceived listening effort. Working memory and cognitive flexibility were not related to ratings of perceived listening effort. In contrast, better inhibitory control was related to higher ratings in both quiet and in noise.

Conclusions: It is possible that the SNR and measures used were not as sensitive as required to measure listening effort behaviorally. In the present experimental setup, prior noise exposure impedes the beneficial effects of performing a task in quiet. Self-reports seem to provide a valid measure of perceived listening effort that is related to the individual's inhibitory control. The present findings suggest that participants with better inhibitory control are more susceptible to the task demand level both in quiet and in noise.

Key Words: adverse listening condition, cognitive demands, response time, subjective ratings, working memory.

*Department of Clinical Science Lund, Logopedics, Phoniatics and Audiology, Lund University, Lund, Sweden

Corresponding author: Jonas Brännström, Department of Clinical Science Lund, Logopedics, Phoniatics and Audiology, Lund University, Lund, Sweden; Email: jonas.brannstrom@med.lu.se

The authors alone are responsible for the content and writing of the article. The authors confirm this manuscript is an original contribution, not previously published, and not under consideration for publication elsewhere.

The study received approval from the regional ethics committee. The study adhered to the Declaration of Helsinki ethical principles.

Abbreviations: FUEL = framework for understanding effortful listening; PCA = principal component analysis; RM = repeated measure; SART = sustained attention response test; SD = standard deviation; SNR = signal-to-noise ratio; WMC = working memory capacity

INTRODUCTION

Despite a favorable signal-to-noise ratio (SNR) and a target speech signal that is fully audible, the listening may still be effortful (Larsby et al, 2005). Attentional-related cognitive and behavioral responses such as listening effort seem to be the result of several factors that can be considered internal and external to the individual (Pichora-Fuller et al, 2016). From this point, the terms effortful listening and listening effort are used interchangeably. The framework for understanding effortful listening (FUEL; Pichora-Fuller et al, 2016) states that the amount of listening effort that a person is experiencing in a specific situation is related to input-related demands, arousal, available cognitive capacity, allocation policy of the cognitive capacity (to what extent and to which task should the available capacity be used), evaluation of the demands on cognitive capacity, and the activity that receives allocation of the cognitive capacity. The evaluation of the demands on cognitive capacity depends on arousal and fatigue, which may in turn influence how a person evaluates his/her performance. The allocation policy depends on both automatic attention, for example, detecting your own name at a busy restaurant, and intentional attention, for example, listening with the right ear in a dichotic listening task. Intentional attention is influenced by, for example, motivation and arousal, which influence the willingness to perform a task. FUEL also proposes that all these components influence attention-related responses such as cognitive behavioral responses and self-reports. Within FUEL, listening effort is defined as a “...specific form of mental effort that occurs when a task involves listening,” where mental effort refers to the “...deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task.” (Pichora-Fuller et al, 2016).

The present study concerns listening effort in relation to input-related demands (quiet and noise), available cognitive capacity, allocation policy of cognitive capacity, and the activity that receives allocation of this capacity. This is studied in a dual-task paradigm where the participant has to listen and remember while at the same time monitoring the auditory signal for response cues. This task is completed in quiet and in noise. The underlying assumption in the dual-task paradigm is that the individual's cognitive resources are limited (Kahneman, 1973). When two tasks are simultaneous, the limited pool of resources must be distributed between these tasks. If the primary task requires allocation of more cognitive resources, fewer resources remain available

for the secondary task. Hence, a decrease in the secondary task performance or response times may be considered to constitute an estimation of how much resources or effort the primary task requires (McGarrigle et al, 2014).

Increased input-related demands due to adverse transmission factors increase cognitive demands (Mattys et al, 2012). Adverse transmission factors degrade the speech signal when it is traveling from its origin to the receiver resulting in an imperfect communication channel (Mattys et al, 2012). The relationship between increased input-related demands and increased cognitive demands can be described using the Ease of Language Understanding model to provide a basis for understanding why listening can be effortful. Under ideal listening conditions, listening to speech is a fast and implicit process (Rönnberg, 2003; Rönnberg et al, 2008; 2010). In this implicit process, multimodal input is automatically connected to the individual phonological representations (Rönnberg, 2003; Rönnberg et al, 2008; 2010). After reaching the identification threshold for a speech sound, the input is matched to the phonological representations available in the semantic long-term memory (Rönnberg, 2003; Rönnberg et al, 2008; 2010). Hence, the identification of lexical representations—and in extension their meaning—is achieved through this matching. The opposite of implicit processing is explicit processing. Listening to a speech signal that is degraded because of, for example, adverse transmission factors results in a mismatch between the input and the phonological representations (Rönnberg, 2003; Rönnberg et al, 2008; 2010; Mattys et al, 2012). To come to terms with the imperfect match explicit processing is required which in turn uses cognitive resources (Rönnberg, 2003; Rönnberg et al, 2008; 2010). As a consequence of the increased engagement of cognitive resources during adverse listening conditions, the perceived listening effort increases (Larsby et al, 2005).

Previous studies indicate that explicit cognitive processing may result in decreased performance, increased listening effort and increased perceived listening effort (i.e., subjective ratings of listening effort) in participants with normal hearing and in particular in participants with hearing impairment (Larsby et al, 2005). Houben et al (2013) interpreted increased response times in noise despite almost maximum audibility as an increase in listening effort. By assessing complex cognitive performance related to speech perception such as semantic decision-making (decide if words belong to predefined semantic categories or not), lexical decision-making (decide whether three-letter words are real words or nonwords), name matching (decide if the letters in a pair are the same or

not), and perceived listening effort, Larsby et al (2005) demonstrated that correct scores decreased, response times increased, and perceived listening effort increased for participants with normal hearing when listening to speech in background noise. In addition, the use of more cognitive resources to suppress task-irrelevant input may result in less capacity available to store relevant information (Hygge et al, 2002; 2003). Rönnerberg et al (2014) proposed that listening effort could be assessed as a measure of this less efficient encoding into long-term memory. Because more cognitive resources become engaged when listening and understanding is based on explicit processing, fewer resources will also be available for the process of encoding information into long-term memory. Therefore, an increased listening effort that taxes a sufficient amount of cognitive resources at the time of information encoding will hypothetically result in poorer delayed recall of information. The present study examines this possibility.

When studying measures of listening effort as possible candidates for assessments in clinical practice, it may be worthwhile to examine if the measures of listening effort are influenced by presentation order, that is whether a noise condition is presented first or last. Although to the best of our knowledge this has not been reported earlier in studies of listening effort, a few studies suggest that there might be an interaction between listening condition and presentation order. Using a visual number updating task performed in quiet and during simultaneous auditory presentation of a fictive story, Sörqvist et al (2010) reported an interaction between listening condition and presentation order. They found that participants who started with the task-irrelevant speech signal improved more in the quiet condition than participants who started in quiet who benefitted less from the quiet condition. Testing offline identification of sentences presented together with a task-irrelevant speech signal, Ljung and Kjellberg (2009) reported that listeners increased their error rates less in the second background listening condition when the first background listening condition had shorter reverberation times (better listening condition) compared with when the first listening condition had longer reverberation times (poorer listening condition). Other studies have not shown these presentation order effects (e.g., Ljung and Kjellberg, 2009). The studies by Sörqvist et al (2010) and Ljung and Kjellberg (2009) indicate it is possible that using noise in a first experimental condition may result in different demands involved in a second listening condition potentially causing a presentation order effect. Based on these previous findings, we hypothesize that listening in quiet as the second listening condition will benefit the participants more than listening in quiet as the first listening condition. Also, participants' subjective ratings of listening effort (i.e., perceived listening effort) are important to establish that the experimental setup has indeed generated effortful listening. Therefore, ratings of perceived listening effort are included and related to the measures of listening effort. Measures of core executive functions are

used to study the available cognitive capacity and the allocation policy in relation to the activity that receives this capacity, that is, our listening effort measures.

Working memory, inhibition, and cognitive flexibility can be considered as "core" executive functions (Miyake et al, 2000; Diamond, 2013). These three core executive functions represent theoretical substrates, but in reality, they may overlap and the combination of these core executive functions are required to regulate and control, for example, attentional-related behavior (Miyake et al, 2000). They can therefore together be considered to provide means for higher-level executive functioning, such as planning, reasoning, and problem-solving (Diamond, 2013). Working memory is used for simultaneous processing and storing information. Hence, high processing activity may reduce the capacity for encoding information into long-term memory. Higher working memory capacity (WMC) has been shown to be related to better speech recognition performance in noise (Lunner and Sundewall-Thorén, 2007; Stenbäck et al, 2015). Working memory can be assessed using a visual N-back test which increases the cognitive load as the participant is required to recall items one or several presentations earlier (Braver et al, 1997). This type of visual task has previously been used in audiological research. For example, Sörqvist et al (2012) showed that the auditory brainstem response was reduced in amplitude with increasing cognitive load. Inhibitory control is the ability to control impulses and undesired activity. This ability seems to be important in listening tasks where a target signal needs to be followed while ignoring task-irrelevant input (Engle, 2002; Diamond, 2013; Ellis and Munro, 2013; Ellis and Rönnerberg, 2014). In line with this, inhibitory control has been shown to be related to better speech recognition performance in noise for adults with normal hearing (Stenbäck et al, 2015; 2016). Sustained attention response test (SART) tests one aspect of inhibitory control by assessing the ability to withhold responses to certain stimuli. Correlation coefficients of intermediate effect (according to Cohen, 1988) has been reported between SART and perceived listening effort for easier tasks in quiet and in noise in participants with normal hearing (Hua et al, 2014a). Cognitive flexibility can be defined as the ability to shift between tasks and the update of information as a result of changing conditions. Perceived listening effort is related to cognitive flexibility in participants with normal hearing. Hua et al (2014a) reported that participants with better cognitive flexibility reported less perceived listening effort in noise for easier tasks. In a dual-task listening paradigm, cognitive flexibility may be required to be able to complete the two tasks and to extract new information to comprehend. Based on these previous studies, we hypothesize that lower listening effort for behavioral measures and perceived listening effort should be related to better core executive functioning given that

the task level difficulty is sufficient to influence the behavioral response measures.

To increase the ecological validity when measuring listening effort, it is important to select a task that is similar to listening tasks and listening conditions encountered in everyday life. Narrative tasks such as passage comprehension are frequently occurring in people's lives, where the listener has to understand and remember information. Passage comprehension is a narrative task that relies on perceptual, cognitive, and linguistic skills (Kintsch, 1988; Bishop, 1997; Kim, 2016). In addition, it requires social processing skills. Performance in this type of task reflects the ability to integrate all these skills and also how well the incoming signal is matched to its long-term representations. Furthermore, passage comprehension tasks require the listener to understand the content linguistically and to integrate what is presented with prior experiences and knowledge with the passage content (Bishop, 1997). This makes these types of tasks challenging from a meta-linguistic perspective, but it also suggests that a listener with previous experience and knowledge about the passage content can use top-down processes to improve task performance (Kintsch, 1988; Kim, 2016). Thus, the familiarity with the passage content makes understanding easier, and it becomes easier to store information.

Furthermore, a listener may find the passage content more or less interesting. This may influence task performance. Also, the listener's motivation in conjunction with the task demands should influence listening effort in a given listening situation (Pichora-Fuller et al, 2016). Picou and Ricketts (2014b) showed that the instructions manipulating the listener's motivation influenced the perceived listening effort for short passages. They showed that increased motivation seems to increase perceived listening effort. This suggests that the listener can overcome increasing listening demands by expending more listening effort when the listener's motivation is high to understand the content, but when the listener's motivation is low, the listener may not be interested in spending any more effort even at minor increases in listening demands (Pichora-Fuller et al, 2016).

Based on this, the first aim of the present study is to apply measures of listening effort and perceived listening effort in participants with normal hearing and to determine if there are any presentation order effects. The second aim is to explore the relationship between listening effort measured as accuracy, response times, efficiency of information encoding into long-term memory, perceived listening effort, and core executive functions.

METHODS

Participants

Thirty-seven participants were recruited in the study. Expecting a relevant difference of 10% (of any measure)

and a standard deviation (SD) 1.5 times, this difference yields a power of ~94% for an alpha-level of 0.05 using 30 participants. The participants were university students, colleagues, and friends of the authors. All recruited participants were tested, but only 32 meet the a priori inclusion criterion: pure-tone air conduction hearing thresholds in both ears equal to or better than 20 dB HL for frequencies 125, 250, 500, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz. Hearing thresholds were assessed using a GN Otometrics Astera² with circumaural sound-attenuating Sennheiser HDA200 earphones in accordance with ISO 8253-1 (1998), ISO 389-5 (2006), and ISO 389-8 (2004). These 32 participants were included in the final sample, and the present study is based on their responses. The final sample consisted of 26 women and 6 men, aged 22 to 41 yr (average of 26.8 yr, SD of 3.4 yr). All participants reported normal vision or corrected to normal. The study complied with the Helsinki Declaration and was approved by the Regional Ethics Board in Lund, Sweden (approval number 2014/95).

Listening Effort

A listening effort task was developed using a dual-task paradigm. The primary task was listening comprehension presented in quiet and in noise. The listening comprehension task consisted of two text passages read by a speech and language pathologist (female, 50-yr old) and five content questions about each passage. The passages came from a standardized test for applications to higher education biannually conducted by the Swedish Council for Higher Education. The two passages were "Litteraturen och biblioteken" ("The literature and the libraries") by Annina Rabe and "Jordgubbar utan allergier" ("Strawberries without allergies") by unknown. The passages contained 446 and 405 words, respectively. The passages were recorded using a Lectret HE-747 microphone connected to a Zoom H2 (Zoom Corporation, Tokyo, Japan) with 44.1 kHz/16-bit sampling frequency. Adobe Audition (version 6; Adobe Systems, San José, CA) was used to normalize the passage recordings to an equal average root-mean-square dB after removing pauses and other silent sections. After the normalization, pauses and other original silent sections were added. At this stage for calibration purposes, a 1000-Hz calibration tone was generated that contained the same average root-mean-square as the speech signals. The durations of the recorded passages used for testing were 2 min and 49 sec and 2 min and 44 sec, respectively. The almost identical durations were caused by a difference in average number of phonemes in a word (5.1 and 5.5, respectively) rather than difference in speaking rates between the recordings.

A multitalker babble noise consisting of two male and two female speakers was recorded. Each speaker was recorded individually while reading different excerpts

from the Swedish version of the “Manual of practical audiometry. Volume 2.” (Arlinger, 1993). The recordings were made using a TSM MT47 condenser microphone connected to an M-audio fasttrack pro soundcard using 44.1 kHz/16-bit sampling frequency. After removing pauses and other silent sections, the four recorded speech signals were normalized to the same average root-mean-square (dB) using the same Adobe Audition software as mentioned above. After adding the original pauses and silent sections to keep prosody intact, the four speech signals were put together into a single signal. Finally, the two passage recordings were combined with the recorded multitalker babble noise to create the listening condition with noise at a 10 dB SNR. The 10 dB SNR was selected to mimic the SNR commonly encountered for speech in noise at the selected speech presentation level (70 dB sound pressure level [SPL]; Lyberg Åhlander et al, 2014) because of the so-called Lombard effect (Lane and Tranel, 1971).

The listening comprehension task was presented through GN Otometrics Astera² with circumaural sound-attenuating Sennheiser HDA200 earphones. After listening to a passage, the participants answered five multiple-choice content questions presented on a paper. Two of these content questions were also taken from the standardized test for applications to higher education biannually conducted by the Swedish Council for Higher Education. Additional three questions were constructed for each passage. An example of the questions is “What of the following reasons does the speaker state as the most important reason to read?” The participant answered the same questions again after approximately 1 hour to assess the ability to encode information into long-term memory. The 1-hour time-frame is similar to previous studies on episodic memory (Enmarker et al, 2006). The same content questions were asked during this second assessment. At this assessment, the questions were presented in the same order, but the response options for each question were presented in reversed order compared with the first assessment. All participants listened to the different passages, one presented in quiet and one in noise. The two background listening conditions and the two passages were balanced across the participants to avoid potential order and fatigue effects.

The secondary task was to correctly identify each occurrence of specific word/word stems while listening to the passage as quickly as possible. The overall instructions to the participants were that they should listen carefully to the passage and that they will be required to answer questions about the passage content afterward. No visual feedback was given. The words/word stems were “litteratur” (“literature”; occurred 12 times in the passage “The literature and the libraries”) and “jordgubbe” (“strawberry”; occurred 13 times in the passage “Strawberries without allergies”), respectively.

Three considerations were made when selecting the keywords: the words should be a content word, should reflect the content of each passage, and have approximately the same frequency in the passages. Responses were made by pressing a designated key on a computer keyboard.

A laptop computer with an internal Conexant Audio Driver sound card connected to Sennheiser HDA 200 earphones was used to present the listening effort task. The software Reaper (version 5; Cockos Incorporated, New York, NY) was used to present and collect the responses. The presentation level was set to 70 dB SPL (± 0.4 dB). This level was verified using a Brüel and Kjaer 2231 sound level meter with a 4134 microphone in a 4153 Artificial Ear by presenting the 1000-Hz calibration tone through the complete setup.

Perceived Listening Effort

After listening to a passage in the listening effort task, the participants rated their perceived listening effort on the Borg CR-10 scale (Borg, 1982; Hua et al, 2014b) using paper and pen. This scale was selected as it has been previously shown to be sensitive to identify differences between different background listening conditions (Larsby et al, 2005). The scale is a combination of ratio and categorical scales. With a combination of numbers and verbal expressions, the participants rate their effort on a scale that ranges from “0”/“none at all” to “10”/“extremely great.” An additional scale step without a number represents “absolute maximum.”

Working Memory

Working memory was assessed using a visual N-back test (Braver et al, 1997). It also provides a measure of the ability to make calculations without writing them down (mental arithmetic; Hubber et al, 2014). In the N-back task, a range of stimuli is presented one at a time. The participant is instructed to determine whether the current stimulus is identical to the stimulus presented N presentations ago. The difficulty level of the task is variable and depends on the N and Braver et al (1997) suggested that the working memory load is proportional to the value of N. In the present study, we used 0-back, where the participant is required to respond if a specific stimulus is presented or not, one-back where the participant is required to decide if the current stimulus is identical to the previous or not, and two-back where the participant is required to decide if the current stimulus is identical to the stimulus before the last. A Swedish version of the N-back test by Ben Robinson and Becky Fuller for E-Studio 2.0 (E-Prime Professional, Sharpsburg, PA) was used (<http://step.psy.cmu.edu/scripts-plus/>). In this version, upper case letters served as stimuli. Using E-Run 2.0 (E-Prime Professional), the letters were presented one after another on a computer

screen and in random order. Each trial consisted of the presentation of one letter during 500 ms followed by 3,000 ms white screen. Responses were allowed during these 3,500 ms. A run consisted of 30 trials. One third of the trials in a run presented target stimuli, that is, a trial where the current stimulus is identical to the stimulus presented *N* presentations ago. The participants were instructed to press the designated keyboard keys for target stimuli and nontarget stimuli. Two consecutive runs were made for each *N*-back condition. One practice run (20 trials, seven targets) was conducted before the actual testing of each *N*-back condition. In the analysis, working memory scores for each *N*-back condition was calculated by dividing the accuracy scores (i.e., proportion correctly identified targets of all possible targets) by the average response times (sec) for target trials (Stenbäck et al, 2016). A higher score indicates better working memory performance.

Inhibitory Control

Inhibitory control, more specifically response inhibition, was assessed using the SART (Manly et al, 1999). The present version of the SART was written in E-Studio 2.0 (E-Prime Professional) and tested using E-Run 2.0 (E-Prime Professional). A stimuli consisted of single numbers (range one to nine) presented one at a time and in random order on a computer screen. In each trial, the task was to press spacebar on a computer keyboard as fast as possible after the number had been presented but to withhold response when the number shown was 3. The numbers were presented during 1,000 ms or until a response had been provided. A 500-ms interstimulus interval was used, and in total, 120 trials were made in a single run. Twenty-four trials contained the number “3.” In the analysis, inhibitory control was calculated as the proportion correctly withheld responses in these trials.

Cognitive Flexibility

Two measures of cognitive flexibility were used. The Number–Letter task was used to measure shifting ability (Miyake et al, 2000; Hua et al, 2014a,b). The present version was written in E-Studio 2.0 (E-Prime Professional) and tested using E-Run 2.0 (E-Prime Professional). In this task, the stimuli consisted of pairs consisting of one number and one letter. The number in a pair could be odd or even and the letter could be an upper or lower case letter. These pairs are presented one after the other in a clockwise manner: the first pair is presented in the upper left corner of the computer screen, the second pair in the upper right corner, the third in the lower right corner, the fourth in the lower left corner, and so on. When stimulus pairs were presented in the upper half of the computer screen, the task

was to decide whether the number was “odd” or “even” by pressing the designated computer keyboard keys. When stimulus pairs were presented in the lower half of the computer screen, the task was to decide whether the letter was in “upper case” or in “lower case.” In each trial, a pair was presented during 10,000 ms or until a response had been provided. After 12 practice trials, 38 trials were made. In the analysis, a measure of shifting ability was calculated: the average response time difference between shift trials (i.e., the preceding trial was a no-shift trial) and no-shift trials (i.e., the preceding trial was a shift trial).

The Keep Track Task assessed the ability to update information (Miyake et al, 2000; Hua et al, 2014a,b). The present version was generated and presented using Microsoft PowerPoint 2010 (version 14.0). In each trial, four of six possible categories of words were presented at the top of a computer screen. These categories were “Countries,” “Animals,” “Colors,” “Metals,” “Fruits,” and “Relatives.” Twelve content words from all six categories (e.g., “Sweden,” “Dog,” “Blue,” “Iron,” “Banana,” and “Cousin”) were presented one after another and in random order. Each word was presented during 3,000 ms with 500-ms interstimulus interval. In each trial, one of the four presented categories was represented by one content word, one by two content words, one by three content words, and one by four content words. The remaining two content words in a trial were selected from the two categories not presented at the top of the screen. At the end of each trial, the task was to report the last content word in each of the four categories using paper and pen. For example, this means that the category represented by four presentations of content words in a trial needed to be updated three times to be reported correctly. After a demonstration and two practice trials, each participant completed six trials. In the analysis, information updating was calculated as the proportion correctly recalled content words in these trials.

Overall Procedures

The participants were tested individually. All tests were made during one session where pauses were allowed when required by the participant. Pure-tone audiometry was conducted in a sound proof booth complying with the maximum permissible background noise levels in ISO 8253-1 (1998). The listening effort task, the ratings of perceived listening effort, *N*-back, SART, the Number–Letter Task, and the Keep Track Task were performed in a quiet room, and the participants wore sound-attenuating Sennheiser HDA200 earphones to block out any unwanted external noise in all these tasks. All participants performed all tasks in the same predefined order: otoscopy, pure-tone audiometry, listening effort task, *N*-back, SART, the Number–Letter Task, and the Keep Track Task. As the study is a part

of a larger project exploring listening effort in both participants with normal hearing and with hearing loss, we choose to use visual cognitive tasks to avoid audibility issues when testing participants with hearing loss and comparing their responses with those from participants with normal hearing.

RESULTS

The results are reported as listening effort (proportion correct responses in the secondary task, average response times in the secondary task, and immediate and delayed listening comprehension scores in the primary task), subjective ratings of perceived listening effort, working memory scores (proportion correctly identified targets of all possible targets divided by the average response times in seconds), inhibitory control (the proportion correctly withheld responses), and cognitive flexibility (shifting ability, i.e., the average response time difference between shift and no-shift trials, and information updating, i.e., the proportion correctly recalled content words). Parametric statistics was used, and an alpha level of 0.05 was considered statistically significant. As age may be a confounding factor, an initial correlation analysis using Pearson’s correlation coefficients was conducted to test the relationship between age and all other variables, that is listening effort correct responses, average response times, and immediate and delayed listening comprehension responses, subjective ratings of perceived listening effort, working memory, inhibitory control, shifting ability, and information updating. Age was not correlated with any of these variables ($r \leq 0.243, p \geq 0.180$) and was therefore not included in the following analyses.

Construct Validity for the Core Executive Function Tests

Table 1 shows the average responses for working memory scores, inhibitory control, shifting ability, and information updating along with SDs, minimum, and maximum. To evaluate underlying factors and estimate interactions between the measures of the core executive functions, a confirmatory principal component analysis (PCA) with varimax rotation and Keiser normalization was conducted. In this analysis, it was assumed that the PCA should identify three factors. The unrotated PCA showed that the three factors together accounted for 77.6% of the total variance encountered in the PCA. After rotation and suppression of loadings less than 0.60, factor 1 accounted for 38.9% of the variance encountered, factor 2 accounted for 21.2%, and factor 3 accounted for 17.5%. Table 1 shows the factor loadings. Factor 1 consisted of the three N-back measures and was therefore labeled working memory. Factor 2 consisted of the shifting ability and information updating measures and was labeled cognitive flexibility. Factor 3 consisted of the measure of inhibitory control and was labeled inhibitory control. Pearson’s correlation coefficients between working memory scores, inhibitory control, shifting ability, and information updating showed that only the three N-back measures were significantly correlated ($r \geq 0.518, p \leq 0.002$) whereas all other correlations were not significant ($r \leq \pm 0.303, p \geq 0.092$). These findings together suggest that N-back, inhibitory control, shifting ability, and information updating seem to measure three different constructs: working memory, inhibitory control, and cognitive flexibility.

Table 1. Average Working Memory Scores (Proportion Correctly Identified Targets of all Possible Targets Divided by the Average Response Times in sec), Inhibitory Control (The Proportion Correctly Withheld Responses), Shifting Ability (The Average Response Time Difference Between Shift Trials), and Information Updating (The Proportion Correctly Recalled Content Words) Along with Standard Deviation SDs, Minimum, and Maximum Values

	Mean	SD	Min	Max	PCA		
					Factor 1 (Working Memory)	Factor 2 (Cognitive Flexibility)	Factor 3 (Inhibitory Control)
Working memory scores 0-back	2.07	0.32	1.37	2.7	0.826		
Working memory scores 1-back	1.81	0.32	1.21	2.49	0.927		
Working memory scores 2-back	1.33	0.39	0.74	2.2	0.815		
Inhibitory control	0.78	0.13	0.42	1.00			0.985
Shifting ability	684	306	212	1,358		-0.702	
Information updating	0.87	0.11	0.54	1.00		0.846	
Variance explained (%)					38.9	21.2	17.5

Note: Factor scores (scores lower than 0.6 are suppressed) are shown along with amount of total variance explained by each factor and tentative factor labels. N = 32.

Listening Effort in Quiet and in Noise

Three separate repeated measure (RM) ANOVAs were conducted to test the effect of background noise on correct responses in the secondary task, response times in the secondary task, and immediate and delayed responses in the primary task and to determine if there are any presentation order effects. Average results are presented in Table 2 along with SDs, minima and maxima. In the first RM ANOVA, the effect of background noise on correct responses in the secondary task was tested. The within-subject factor was correct responses and the between-subject factor was presentation order (quiet first and noise last or noise first and quiet last). No covariates were added. No main effect was seen for correct responses [Wilks' lambda = 0.973, $F_{(1,30)} = 0.820$, $p = 0.373$, $\eta^2 = 0.027$]. A significant between-subject effect was seen [$F_{(1,30)} = 4.345$, $p = 0.046$, $\eta^2 = 0.127$] indicating that participants beginning in quiet performed better than participants beginning in noise. A significant interaction effect was seen [Wilks' lambda = 0.874, $F_{(1,30)} = 4.345$, $p = 0.046$, $\eta^2 = 0.126$]. Figure 1 shows the interaction effect. This interaction indicates a presentation order effect: participants beginning in quiet performed worse in noise than in quiet and the participants beginning in noise did not improve their performance in quiet.

In the second RM ANOVA, the effect of background noise on response times in the secondary task was tested. The within-subject factor was response times and the between-subject factor was presentation order (quiet first or noise first). No covariates were used. No main effect was seen for response times [Wilks' lambda = 0.999, $F_{(1,30)} = 0.021$, $p = 0.887$, $\eta^2 = 0.001$]. No between-subject effect [$F_{(1,30)} = 1.691$, $p = 0.203$, $\eta^2 = 0.053$] and no interaction effect was seen [Wilks' lambda = 0.953, $F_{(1,30)} = 1.491$, $p = 0.232$, $\eta^2 = 0.047$].

In the third RM ANOVA, the effect of background noise on immediate and delayed listening comprehension in the primary task was tested. The within-subject factors were listening comprehension (immediate and delayed) and background noise (quiet or noise), and

the between-subject factor was presentation order (quiet first or noise first). No covariates were used. No main effect was seen for listening comprehension [Wilks' lambda = 0.993, $F_{(1,30)} = 0.516$, $p = 0.478$, $\eta^2 = 0.017$]. No main effect was seen for background noise [Wilks' lambda = 1.000, $F_{(1,30)} = 0.005$, $p = 0.944$, $\eta^2 < 0.001$]. No between-subject effect [$F_{(1,30)} = 1.691$, $p = 0.203$, $\eta^2 = 0.053$] and no interaction effects were seen [Wilks' lambda ≥ 0.983 , $F_{(1,30)} \leq 2.670$, $p \geq 0.232$, $\eta^2 \leq 0.082$].

Perceived Listening Effort in Quiet and in Noise

Using a RM ANOVA, the effect of background noise on perceived listening effort was tested. The within-subject factor was perceived listening effort and the between-subject factor was presentation order (quiet first or noise first). No covariates were added. A significant main effect was seen for perceived listening effort [Wilks' lambda = 0.436, $F_{(1,30)} = 38.845$, $p \leq 0.001$, $\eta^2 = 0.564$]. The corresponding data are shown in Figure 2. No between-subject effect [$F_{(1,30)} = 1.085$, $p = 0.306$, $\eta^2 = 0.035$] or interaction effect [Wilks' lambda = 0.993, $F_{(1,30)} = 0.213$, $p = 0.648$, $\eta^2 = 0.007$] was seen. The subjective ratings of perceived listening effort increased significantly in noise compared with quiet.

Correlations Between Listening Effort, Perceived Listening Effort, and Core Executive Functions

In a first analysis, Pearson's correlation coefficients were calculated to assess the relationship between listening effort (measured as correct responses in the secondary task, average response times in the secondary task, and immediate and delayed listening comprehension responses in the primary task) in quiet and in noise, working memory, inhibitory control, shifting ability, and information updating. In this analysis, to adjust for the large number of correlations calculated, the false discovery rate was used to calculate a corrected alpha level (Benjamini and Hochberg, 1995).

Table 2. Average Proportion of Correct Responses in the Secondary Task, Response Times in the Secondary Task, Immediate and Delayed Passage Comprehension Responses in Quiet and in Noise, and Perceived Listening Effort in Quiet and in Noise Along with Standard Deviation SDs, Minimum, and Maximum Values

	Quiet				Noise			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Correct responses	0.95	0.08	0.69	1.00	0.94	0.08	0.75	1
Response times	673	174	425	1,245	668	163	431	1,068
Immediate listening comprehension	2.94	1.05	1	5	2.81	1.06	1	5
Delayed listening comprehension responses	2.72	1.17	0	5	2.88	1.10	1	5
Perceived listening effort	2.59	1.76	0	7	4.28	2.25	2	9

Note: N = 32.

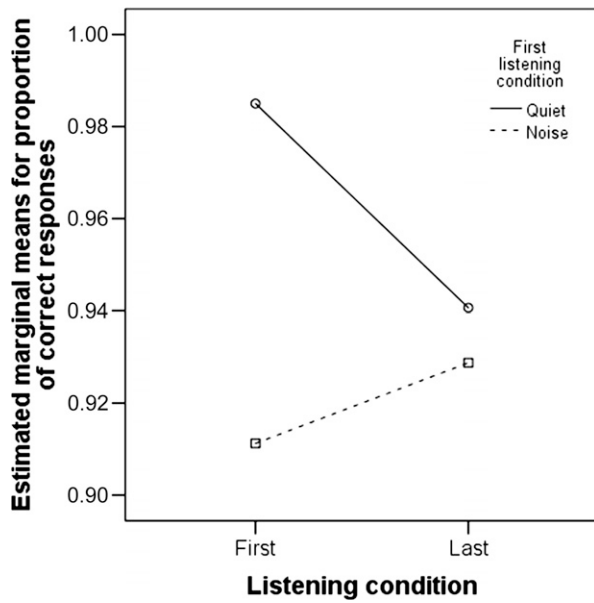


Figure 1. Estimated marginal means for proportions of correct responses in quiet and in noise for listening comprehension as a function of first and last listening condition (quiet or noise). Drawn line indicates that the first listening condition was in quiet and the last in noise. Dashed line indicates that the first listening condition was in noise and the last in quiet. Error bars show SDs. N = 32.

Listening effort assessed as correct responses, response times, immediate listening comprehension responses, and delayed listening comprehension responses in quiet and in noise were not correlated with any of the measures of core executive functions ($r \leq \pm 0.439, p \geq 0.012$). A significant negative correlation was seen between listening

effort correct responses in quiet and listening effort response times in noise indicating that increasing effort correct responses in quiet were related to decreasing response times in noise ($r = -0.528, p = 0.002$). Significant positive correlations were found between immediate and delayed listening comprehension responses both in quiet ($r = 0.802, p < 0.001$) and in noise ($r = 0.698, p < 0.001$). These findings indicate that higher immediate responses in quiet and in noise were related to higher delayed responses for both listening conditions.

In a second analysis, Pearson’s correlation coefficients were calculated to assess the relationship between perceived listening effort in quiet and in noise, working memory, inhibitory control, shifting ability, and information updating. Here also, the false discovery rate was used to calculate a corrected alpha level. A significant positive correlation was found between perceived listening effort in quiet and perceived listening effort in noise ($r = 0.741, p < 0.001$). Significant positive correlations were found between inhibitory control and perceived listening effort in quiet ($r = 0.460, p = 0.008$) and in noise ($r = 0.454, p = 0.009$) (c.f., Figure 3) indicating that better inhibitory control was related to higher ratings of perceived listening effort in both quiet and in noise.

In a third final analysis, Pearson’s correlation coefficients were calculated to assess the relationship between listening effort (i.e., correct responses, average response times, and immediate and delayed listening comprehension responses) in quiet and in noise and perceived listening effort in quiet and in noise. A corrected alpha level was calculated using the false discovery rate. No significant correlations were found between

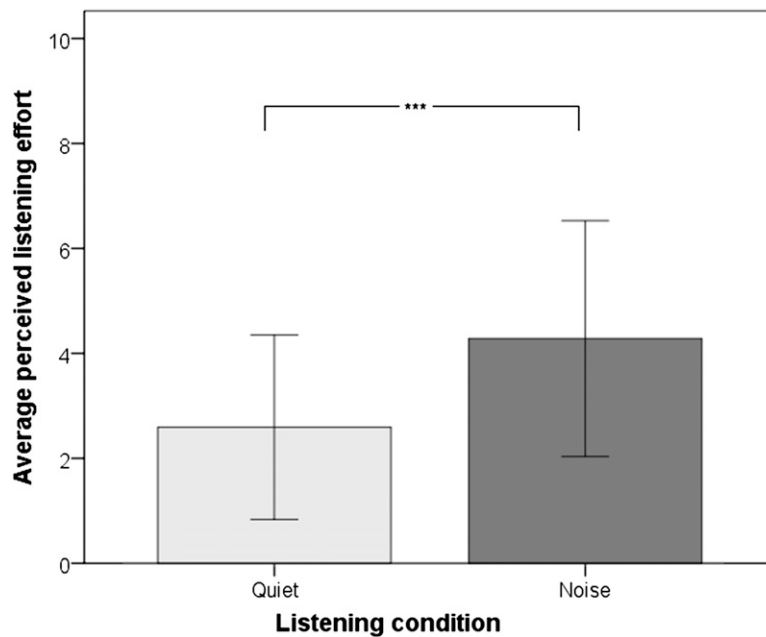


Figure 2. Average perceived listening effort in quiet (light gray) and in noise (dark gray). Error bars show SDs, and the asterisks indicate the significant effect ($p < 0.001$). N = 32.

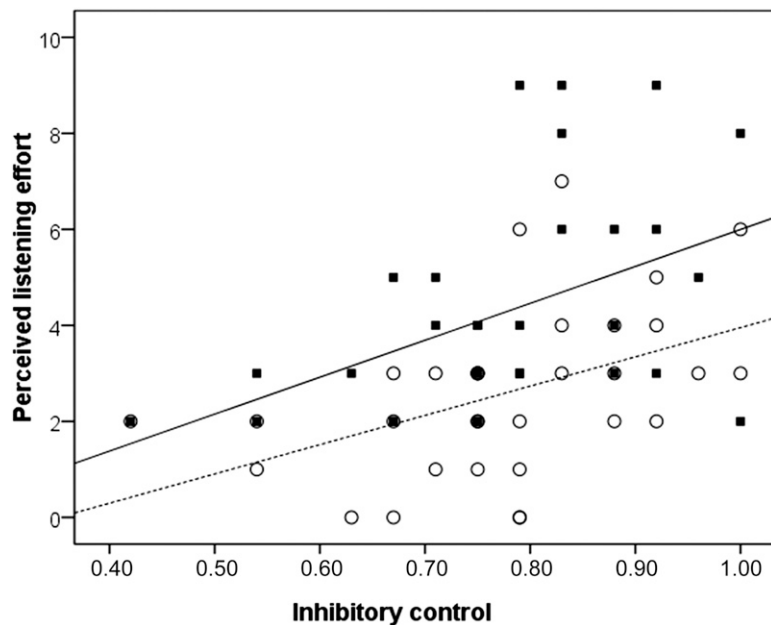


Figure 3. Scatter plot showing the relationship between perceived listening effort and inhibitory control for individual participants. Open circles denote perceived listening effort in quiet and filled squares perceived listening effort in noise. The lines represent the linear interpolation in quiet (dashed line) and in noise (drawn line). $N = 32$.

the listening effort measures in quiet and in noise and perceived listening effort in quiet and in noise ($r \leq 0.272$, $p \geq 0.132$).

DISCUSSION

The present findings suggest that our measures of listening effort, that is correct responses in the secondary task, response times in the secondary task, and immediate and delayed listening comprehension in the primary task, failed to demonstrate increased listening effort in background noise (10 dB SNR) compared with quiet. In addition, a significant test order effect was seen for correct responses indicating that participants who first listened in noise did not improve in quiet. However, the subjective ratings of perceived listening effort increased significantly in noise compared with quiet. No relationship was found between measures of listening effort and ratings of perceived listening effort. Working memory, inhibitory control, and cognitive flexibility were not related to performance in the different measures of listening effort. Working memory and cognitive flexibility were not related to ratings of perceived listening effort, but better inhibitory control was related to higher ratings in both quiet and in noise.

The Impact of Noise and Test Order on Listening Effort and Perceived Listening Effort

The first aim of the present study was to apply measures of listening effort and perceived listening effort in

participants with normal hearing and to determine if there were any presentation order effects. Despite a significant main effect of multitalker babble noise on perceived listening effort, the present findings suggest that the measures of listening effort (correct responses, response times, immediate and delayed listening comprehension) failed to detect any detrimental effects of noise presented at 10 dB SNR. This suggests that, although the participants experienced that the listening was more effortful in noise than in quiet, it did not affect their correct responses, response times, and immediate and delayed listening comprehension performance. It is likely that the SNR was too positive in relation to the tasks applied: Ohlenforst et al (2016) reported that listening effort estimated as pupil dilatation in participants with normal hearing increased with increasingly more challenging SNR in speech recognition to a certain point where the dilatation started to decrease because of a too difficult task. According to FUEL, pupil dilatation represents a measure of automatic arousal responses but is also related to cognitive capacity. Thus, in the present study, the relationship between the target speech signal and the background noise in combination with the level of the language processing required (Kintsch, 1988; Kim, 2016) may have provided a dual task that was too easy to result in noise effects at this SNR in these measures of listening effort. In the listening comprehension task used here, the listener is required to both understand the content linguistically and to integrate any earlier knowledge about the world and experiences with the passage content (Bishop, 1997). Thus, the listening comprehension task is challenging from a meta-linguistic perspective

(Kintsch, 1988; Kim, 2016). However, it also means that the language and information redundancy in the signal for individuals with previous experiences and knowledge about the passage content will facilitate their top-down processes which in turn will improve the efficiency of the performance in the task and at the end, the ability to store information. This may have reduced the impact of noise on our measures of listening effort in these individuals as more previous experiences and knowledge about the passage content may have allowed for more allocation of cognitive resources to the secondary task that are not necessarily related to the cognitive capacity of the individual. This explanation would also explain that the measures of listening effort in the present study were not at all related to performance in the cognitive tests.

In addition, some participants may have found the passage contents more or less interesting to listen to which may have influenced their performance. FUEL suggests that listener motivation in relation to the listening demands influences listening effort in a given listening situation (Pichora-Fuller et al, 2016). When motivation is high, the listener can overcome increasing listening demands by expending more listening effort. However, when the motivation is low because of, for example, an uninteresting story, the listener may not be interested in spending any more effort even at minor increases in listening demands (Picou and Ricketts, 2014a). Future studies should incorporate a measure of listener motivation. Furthermore, Pichora-Fuller et al (2016) suggest that listening effort experienced in an experimental setup at the laboratory may be different from real-life experience of listening effort because of, for example, much longer task durations in real-life listening. This could suggest that the short passages used in the present setup were too short (less than three minutes long) to assess listening effort.

The present findings indicate a significant test order effect for correct responses. Participants listening first in quiet performed more poorly in noise whereas participants listening first in noise did not improve in quiet. This finding is in contrast to our initial hypothesis that listening in quiet as the second listening condition would benefit the participants more than listening in quiet as the first listening condition. There are several reasons why the present findings differ from those previously reported by Sörqvist et al (2010) and Ljung and Kjellberg (2009). Sörqvist et al (2010) that participants who started with a task-irrelevant speech signal improved more in a quiet condition than participants who started in quiet who benefitted less from the quiet condition. However, they used a visual number updating task in quiet and during simultaneous auditory presentation of a fictive story and it is possible that the use of a dual-task paradigm in only the auditory modality may result in different order effects.

Ljung and Kjellberg (2009) tested visual identification of sentences presented to the listener in quiet and during task-irrelevant speech. They found that listeners increased their error rates less in the second background listening condition when the first background listening condition had shorter reverberation times (better listening condition) compared with when the first listening condition had longer reverberation times (poorer listening condition). The sentence identification task offline is different from the online measures acquired in a passage comprehension task in the present study and may therefore yield different presentation order effects. The present findings could suggest that it is possible that the potentially detrimental impact of noise on our measures of listening effort was masked by a test order effect that suggests prior noise exposure influences subsequent performance in quiet. These findings seem also to demonstrate the benefit of asking the participants about their perceived listening effort which was clearly negatively influenced by the presence of noise and without any order effects. In relation to FUEL (Pichora-Fuller et al, 2016), it may be possible to interpret the present findings that the input-related demands may be influenced by the experimental design. Although the present study examined participants with normal hearing, it may be the most viable way in the clinical setting to ask about perceived listening effort rather than using these types of measures. However, studies on participants with hearing impairment are required.

We found no relationship between the different measures of listening effort and perceived listening effort. This lack of relationship may be the result of that the task level demands were too low to show any effect on the measures of listening effort. We used several operational definitions of listening effort in the present study. All these measures have been proposed to measure listening effort (McGarrigle et al, 2014). Therefore, from a theoretical point of view, it is problematic that we cannot establish a consistent effect of noise on our behavioral measures of listening effort or an association between behavioral measures and subjective reports. It is possible that our behavioral measures of listening effort are only accurate or sensitive at more challenging SNRs. However, the present finding is similar to those in previous studies that suggest that different measures of listening effort and perceived listening effort are not necessarily related (Hick and Tharpe, 2002; Sarampalis et al, 2009; Fraser et al, 2010; Gosselin and Gagné, 2011; Mackersie and Cones, 2011).

The Relationship with Core Executive Functions

In the present study, four tests (working memory, inhibitory control, shifting ability, and information updating) were used to assess core executive functions,

that is, working memory, inhibitory control, and cognitive flexibility (Diamond, 2013; Miyake et al, 2000). Initially, a factor analysis was conducted to evaluate underlying factors and estimate interactions between the measures of the core executive functions. In addition, a correlation analysis was conducted to explore the relationships between these measures. These analyses suggested that the four tests seem to assess three different underlying constructs, which we interpreted as these three core executive functions. Thus, the cognitive tests seem to provide information on separate aspects of executive functioning.

The second study aim was to explore the relationship between listening effort measured as accuracy, response times, encoding information into long-term memory (i.e., delayed listening comprehension), perceived listening effort, and core executive functioning. The present findings suggest no significant relationships between our measures of listening effort and core executive functioning. Previous studies suggest that higher WMC is related to better sentence-based speech recognition performance in noise (Lunner and Sundewall-Thorén, 2007; Stenbäck et al, 2015). However, the relationship depends on the SNR. Stenbäck et al (2015) found that WMC as assessed by a reading span task was related to sentence-based speech recognition but only at an SNR targeting 80% word recognition. In a similar fashion, Stenbäck et al (2015; 2016) reported that better inhibitory control was related to better speech recognition performance in noise for adults with normal hearing. Cognitive flexibility has been shown to be related to perceived listening effort (Hua et al, 2014a) but not to performance in noise. The test order of listening effort and the core executive functions was not balanced across participants. This is a similar design as in previous studies (Nilsson et al, 1997; Enmarker et al, 2006). Future studies should balance or randomize the test order to avoid any potential order and fatigue effects.

As more cognitive resources are believed to be used in the suppression of task-irrelevant sounds, we expected that fewer resources would be available to encode information into long-term memory which would thus reflect listening effort (Hygge et al, 2002; 2003; Rönnerberg et al, 2014). The present findings are in contrast to these previous findings and assumptions because we found no effect of background noise in delayed responses in the listening comprehension task. It is possible that the 10 dB SNR used in the present listening comprehension task was too favorable and did not provide sufficient adverse listening condition that required explicit processing (Rönnerberg, 2003; Rönnerberg et al, 2008; 2010; Mattys et al, 2012). This in combination with the higher level of the language processing required (Kintsch, 1988; Kim, 2016) may, as already been discussed above, have provided a dual-task that was too easy to result in noise effects in these measures of listening effort.

The present findings showed that the perceived listening effort increased in noise compared with in quiet. Furthermore, the present findings suggest a significant relationship between perceived listening effort and inhibitory control where participants with better inhibitory control reported higher perceived listening effort in quiet and in noise. This may seem contrainuitive, but it is possible that only those with sufficient amount of cognitive resources not engaged in the task were able to perceive the task as more demanding both in quiet and in noise (Lavie, 2005).

CONCLUSIONS

Four different behavioral measures of listening effort, that is, correct responses, response times, immediate, and delayed listening comprehension, were unaffected by the presence of multitalker babble noise presented at 10 dB SNR. These measures were also unrelated to assessments of three core executive functions. One of these measures suggested a clear order effect where prior noise exposure impedes the beneficial effects of performing the task in quiet. It is possible that the SNR and measures used were not as sensitive as required to measure listening effort behaviorally. Self-reports seem to provide a valid measure of perceived listening effort that is related to the core executive function inhibitory control. It seems as if participants with better inhibitory control are more susceptible to the task demand level both in quiet and in noise.

REFERENCES

- Arlinger S. (1993) *Manual of Practical Audiometry*. Vol. 2. London: Whurr Publishers Ltd.
- Benjamini Y, Hochberg Y. (1995) Controlling for the false discovery rate: a practical and powerful approach to multiple testing. *JR Soc* 57:289–300.
- Bishop D. (1997) *Uncommon Understanding: Development and Disorders of Language Comprehension in Children*. Hove, United Kingdom: Psychology Press.
- Borg GA. (1982) Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 14(5):377–381.
- Braver TS, Cohen JD, Nystrom LE, Jonides J, Smith EE, Noll DC. (1997) A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 5(1):49–62.
- Cohen J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.
- Diamond A. (2013) Executive functions. *Annu Rev Psychol* 64: 135–168.
- Ellis RJ, Munro KJ. (2013) Does cognitive function predict frequency compressed speech recognition in listeners with normal hearing and normal cognition? *Int J Audiol* 52(1):14–22.

- Ellis RJ, Rönnerberg J. (2014) Cognition and speech-in-noise recognition: the role of proactive interference. *J Am Acad Audiol* 25(10): 975–982.
- Engle RW. (2002) Working memory capacity as executive function. *Curr Dir Psychol Sci* 11:19–23.
- Enmarker I, Boman E, Hygge S. (2006) Structural equation models of memory performance across noise conditions and age groups. *Scand J Psychol* 47(6):449–460.
- Fraser S, Gagné JP, Alepins M, Dubois P. (2010) Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *J Speech Lang Hear Res* 53(1):18–33.
- Gosselin PA, Gagné JP. (2011) Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. *Int J Audiol* 50(11):786–792.
- Hick CB, Tharpe AM. (2002) Listening effort and fatigue in school-age children with and without hearing loss. *J Speech Lang Hear Res* 45(3):573–584.
- Houben R, van Doorn-Bierman M, Dreschler WA. (2013) Using response time to speech as a measure for listening effort. *Int J Audiol* 52(11):753–761.
- Hua H, Emilsson M, Ellis R, Widén S, Möller C, Lyxell B. (2014a) Cognitive skills and the effect of noise on perceived effort in employees with aided hearing impairment and normal hearing. *Noise Health* 16(69):79–88.
- Hua H, Emilsson M, Kähäri K, Widén S, Möller C, Lyxell B. (2014b) The impact of different background noises: effects on cognitive performance and perceived disturbance in employees with aided hearing impairment and normal hearing. *J Am Acad Audiol* 25(9):859–868.
- Hubber PJ, Gilmore C, Cragg L. (2014) The roles of the central executive and visuospatial storage in mental arithmetic: a comparison across strategies. *Q J Exp Psychol (Hove)* 67(5):936–954.
- Hygge S, Boman E, Enmarker I. (2003) The effects of road traffic noise and meaningful irrelevant speech on different memory systems. *Scand J Psychol* 44(1):13–21.
- Hygge S, Evans GW, Bullinger M. (2002) A prospective study of some effects of aircraft noise on cognitive performance in school-children. *Psychol Sci* 13(5):469–474.
- ISO 8253-1. (1998) ISO 8253-1. Acoustics: Audiometric Test Methods Part 1: Basic Pure-Tone Air and Bone Conduction Threshold Audiometry. Geneva: International Organization for Standardization.
- ISO 389-8. (2004) ISO 389-8. Acoustics: Reference Zero for the Calibration of Audiometric Equipment. Part 8: Reference Equivalent Threshold Sound Pressure Levels for Pure Tones and Circumaural Earphones. Geneva: International Organization for Standardization.
- ISO 389-5. (2006) ISO 389-5. Acoustics: Reference Zero for the Calibration of Audiometric Equipment. Part 5: Reference Equivalent Threshold Sound Pressure Levels for Pure Tones in the Frequency Range 8 kHz to 16 kHz. Geneva: International Organization for Standardization.
- Kahneman D. (1973) *Attention and Effort*. Englewood Cliffs, NJ: Prentice Hall, Inc.
- Kim YS. (2016) Direct and mediated effects of language and cognitive skills on comprehension of oral narrative texts (listening comprehension) for children. *J Exp Child Psychol* 141:101–120.
- Kintsch W. (1988) The role of knowledge in discourse comprehension: a construction-integration model. *Psychol Rev* 95(2):163–182.
- Lane H, Tranel B. (1971) The Lombard sign and the role of hearing in speech. *J Speech Hear Res* 14:677–709.
- Larsby B, Hällgren M, Lyxell B, Arlinger S. (2005) Cognitive performance and perceived effort in speech processing tasks: effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *Int J Audiol* 44(3):131–143.
- Lavie N. (2005) Distracted and confused?: selective attention under load. *Trends Cogn Sci* 9(2):75–82.
- Ljung R, Kjellberg A. (2009) Long reverberation times decreases recall of spoken information. *Build Acoust* 16(4):301–311.
- Lunner T, Sundewall-Thorén E. (2007) Interactions between cognition, compression, and listening conditions: effects on speech-in-noise performance in a two-channel hearing aid. *J Am Acad Audiol* 18(7):604–617.
- Lyberg Åhlander V, Pelegrín García D, Whitting S, Rydell R, Löfqvist A. (2014) Teachers' voice use in teaching environments: a field study using ambulatory phonation monitor. *J Voice* 28(6): 841.e5–841.e15.
- Mackersie CL, Cones H. (2011) Subjective and psychophysiological indexes of listening effort in a competing-talker task. *J Am Acad Audiol* 22(2):113–122.
- Manly T, Robertson IH, Galloway M, Hawkins K. (1999) The absent mind: further investigations of sustained attention to response. *Neuropsychologia* 37(6):661–670.
- Mattys SL, Davis MH, Bradlow AR, Scott SK. (2012) Speech recognition in adverse conditions: a review. *Lang Cogn Process* 27: 953–978.
- McGarrigle R, Munro KJ, Dawes P, et al. (2014) Listening effort and fatigue: what exactly are we measuring? A British Society of Audiology cognition in hearing special interest group 'white paper'. *Int J Audiol* 53(7):433–440.
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. (2000) The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognit Psychol* 41(1):49–100.
- Nilsson L-G, Bäckman L, Erngrund K, Nyberg L, Adolfsson R, Bucht G, Karlsson S, Widing M, Winblad B. (1997) The betula prospective cohort study: memory, health, and aging. *Ageing Neuropsychol Cogn* 4:1–32.
- Ohlenforst B, Zekveld A, Lunner T, Wendt D, Naylor G, Wang Y, Versfeld NJ, Kramer SE. (2016) Impact of stimulus-related and listener-related factors on cognitive processing load as indicated by pupil dilatation. Conference proceedings of Hearing Across the Lifespan. Como, Italy, June 2–4, 2016.
- Pichora-Fuller MK, Kramer SE, Eckert MA, Edwards B, Hornsby BWY, Humes LE, Lemke U, Lunner T, Matthen M, Mackersie CL, Naylor G, Phillips NA, Richter M, Rudner M, Sommers MS, Tremblay KL, Wingfield A. (2016) Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear* 37 (Suppl 1):S5–S27S.

- Picou EM, Ricketts TA. (2014a) The effect of changing the secondary task in dual-task paradigms for measuring listening effort. *Ear Hear* 35(6):611–622.
- Picou EM, Ricketts TA. (2014b) Increasing motivation changes subjective reports of listening effort and choice of coping strategy. *Int J Audiol* 53(6):418–426.
- Rönnerberg J. (2003) Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: a framework and a model. *Int J Audiol* 42(Suppl 1):S68–S76.
- Rönnerberg J, Rudner M, Foo C, Lunner T. (2008) Cognition counts: a working memory system for ease of language understanding (ELU). *Int J Audiol* 47(Suppl 2):S99–S105.
- Rönnerberg J, Rudner M, Lunner T. (2014) Comments from Dr. Jerker Rönnerberg, Mary Rudner, Thomas Lunner. *Int J Audiol* 53(7): 441–442. discussion 444–445. [Peer commentary on “Listening effort and fatigue: What exactly are we measuring? A British Society of Audiology Cognition in Hearing Special Interest Group ‘white paper’,” by Ronan McGarrigle, Kevin J. Munro, Piers Dawes, Andrew J. Stewart, David R. Moore, Johanna G. Barry, & Sygal Amitay].
- Rönnerberg J, Rudner M, Lunner T, Zekveld AA. (2010) When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12(49):263–269.
- Sarantis A, Kalluri S, Edwards B, Hafer E. (2009) Objective measures of listening effort: effects of background noise and noise reduction. *J Speech Lang Hear Res* 52(5):1230–1240.
- Stenback V, Hällgren M, Larsby B. (2016) Executive functions and working memory capacity in speech communication under adverse conditions. *Speech Lang Hear* 19:218–226.
- Stenback V, Hällgren M, Lyxell B, Larsby B. (2015) The Swedish Hayling task, and its relation to working memory, verbal ability, and speech-recognition-in-noise. *Scand J Psychol* 56(3):264–272.
- Sörqvist P, Halin N, Hygge S. (2010) Individual differences in susceptibility to the effects of speech on reading comprehension. *Appl Cogn Psychol* 24:67–76.
- Sörqvist P, Stenfelt S, Rönnerberg J. (2012) Working memory capacity and visual-verbal cognitive load modulate auditory-sensory gating in the brainstem: toward a unified view of attention. *J Cogn Neurosci* 24(11):2147–2154.