Research Paper

The role of reward and task demand in value-based strategic allocation of auditory comprehension effort

Min Zhang a, b, *, Greg J. Siegle e, Malcolm R. McNeil b, d, Sheila R. Pratt b, d, Catherine Palmer b, c

a Shanghai Key Laboratory of Clinical Geriatric Medicine, Huadong Hospital, Fudan University, Shanghai, China
b Department of Communication Science and Disorders, University of Pittsburgh, PA, USA
c Department of Otolaryngology, University of Pittsburgh Medical Center, PA, USA
d Geriatric Research, Education, and Clinical Center, VA Pittsburgh Healthcare System, PA, USA
e Department of Psychiatry, University of Pittsburgh, PA, USA

A R T I C L E   I N F O
Article history:
Received 25 February 2019
Received in revised form 30 July 2019
Accepted 31 July 2019
Available online 2 August 2019

Keywords:
Auditory comprehension effort
Strategic allocation
Task demand
Pupillometry

A B S T R A C T

Objectives: Listeners who fail to optimize their allocation of effort during auditory comprehension tasks can experience from compromised performance, fatigue and stress, which might result in reduced engagement in social communication activities. Strategically allocating effort based on costs and perceived benefits are commonly observed in the research of effortful physical and visual behaviors. Whether people manage their effort in a similar manner in audition remains unclear. As the listening performance of people with normal hearing often serves as the goal of auditory rehabilitation for people with hearing loss, this study evaluated how strategy-induced effort allocation, challenged by reward and task demand, interactively impacted auditory comprehension in normal hearing adults.

Design: A value-based strategic effort allocation paradigm was evaluated in 40 normal-hearing young adults. The paradigm included five levels of reward (motivation) and five levels of task demand (speech rate) that were independently manipulated. Effects of reward and task demand on performance accuracy and pupil dilation (a measure of auditory comprehension effort) were examined.

Results: There was a significant interaction effect of reward and task demand on both pupil dilation and comprehension accuracy. At the response stage of speech comprehension processing, pupil dilation significantly decreased as the task demand increased at high reward levels. In contrast, pupil dilation did not vary significantly as a function of task demand at low reward levels. Reward significantly improved performance accuracy at fast and extremely fast rate conditions, but not at the slower rates.

Conclusions: Consistent with previous studies on effort regulation, reward and task demand appear to be associated with auditory comprehension effort allocation in an interactive manner when strategic effort control was required to achieve a reward goal. The young normal-hearing listeners in this study prioritized their effort to relatively easy task items over difficult ones at high levels of reward, suggesting a cost-effective value-based strategic effort allocation. Reward significantly improved task performance in terms of accuracy at difficult listening conditions. These findings support the incorporation of affective factors (e.g., reward) and the utility of the value-based strategic effort allocation paradigm in the experimental setting to understand how clinically relevant factors (such as hearing loss and age) might change strategic auditory comprehension behavior.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Hearing loss reduces available auditory processing accuracy and efficiency (Kricos, 2006; Piquado et al., 2012; Picou and Ricketts, 2014; Pichora-Fuller et al., 2016) and requires compensation by other cognitive resources. While listening effort represents the resources consumed when performing an auditory task to improve
speech comprehension outcomes (Pichora-Fuller et al., 2016), auditory comprehension effort is a broader construct that captures the entire span of cognition in a comprehension task from listening preparation, auditory perception, to response decision and execution, all of which are involved in comprehension. With this conception, auditory comprehension effort represents the result of allocating cognitive resources to processes related to the task (speech processing, responding, etc.) in sequence or parallel over time, with listening effort as a primary component.

The decision to expend or withhold effort plays an important role in determining how efficiently a task goal is achieved (Westbrook and Braver, 2015). Strategically allocating effort based on the value created by its costs (e.g., task demand) and perceived benefits (e.g., reward) has been investigated across disciplines. These include psychophysiological and neuropsychiological research involving physical effort (Pasquareau and Turner, 2013; Mathar et al., 2013; Harris and Lim, 2016; Kleinflügge et al., 2016; Studer and Knecht, 2016), visual processing effort (Croxon et al., 2009; Braun and Arrington, 2018), and cognitive effort (Bronstein and Baruchson-Arib, 2008; Bijleveld et al., 2009; Westbrook and Braver, 2015; Sandra and Otto, 2018). However, the unique or combined contributions of task demand and reward on effort allocation for auditory comprehension remain unclear. Listeners who fail to optimize effort allocation among auditory comprehension tasks can suffer from compromised performance, fatigue and stress, which can result in increased economic, social communication and other burdens (Albanini et al., 2017). These negative consequences are particularly impactful for people with hearing loss (Lambert et al., 2017), who take longer to recover from fatigue compared to people with normal hearing (Nachtegaal et al., 2009, 2012; Hornsby et al., 2016). In order to enhance performance and reduce fatigue in these individuals, better understanding of the mechanisms for the strategic allocation of auditory comprehension effort may be helpful in motivating effort mobilization and prioritizing and balancing mental resources among independent or concurrent tasks. Although currently poorly understood, it is critical to know how people with normal hearing and hearing loss strategically manage auditory comprehension.

The decision to allocate effort can, in part, determine how efficiently a task goal is achieved (Westbrook and Braver, 2015), and includes both cognitive and affective aspects of cognition (Prehn et al., 2011). Cognitive-load/effort relationships have been documented in the auditory domain over the past three decades, demonstrating that listening effort generally increases as the processing requirements of the task increase with or without compromised performance accuracy (Kramer et al., 1997; McCoy et al., 2005; Strauss et al., 2008; Sarampalis et al., 2009; Tun et al., 2009; Zekveld et al., 2010; Mackerse and Cones, 2011; Picou et al., 2011; Desjardins and Doherty, 2014; Picou and Ricketts, 2014; Iverson et al., 2016; Holmes et al., 2018). However, the systematic increase of effort reaches an asymptote when it is at or near resource limits, and declines when processing demands exceed available resources (Granholm et al., 1996; Ohlenforst et al., 2017; Wendt et al., 2018). This finding suggests a cautious interpretation of an observed effort reduction in studying strategic effort allocation behavior, as this sign of disengagement/giving-up might have occurred due to resource limits rather than a choice based on the value judgment of the task.

In contrast, the study of affect/effort relationships in humans’ auditory comprehension activity has a shorter history and relatively sparse database. Motivation, as one aspect of affect, functions as a modulator of the amount of effort applied to task completion (for reviews, see Brehm and Self, 1989; Richter, Gendolla and Wright, 2016). Specifically, when task demand is high, the level of effort expenditure increases in response to the elevation in reward.

In contrast, there is little influence of reward on effort expenditure when task demand is low (Richter, 2016). A recently proposed theoretical Framework for Understanding Effortful Listening (FUEL) (Pichora-Fuller et al., 2016) combines cognitive-behavioral and motivation theories to suggest a dual listening-effort driving mechanism. In this framework, task demand and affective features, such as motivation, are proposed to play essential roles in listening effort control. In support of the FUEL framework, a positive linear relationship between listening effort and various types of motivation, including monetary reward (Richter, 2016; Koelwijn et al., 2018), quiz stress (Picou, 2014), and evaluative threat (Mackersie and Kearney, 2017) have been demonstrated. The common finding from this limited literature is that normal hearing listeners spend more effort in listening when their motivation and auditory task demand are both high, compared to other conditions. However, how task demand and motivation interactively impact effort when flexible and strategic allocation of effort is required in achieving an auditory comprehension goal has not been examined.

Across disciplines, it is generally agreed that effort control involves a cost-benefit trade-off between the evaluation of task demands and the motivation to achieve a goal. When making decisions, the brain appears to weigh costs against benefits by combining neural benefit and cost signals into a single, difference-based neural representation of value (Croxon et al., 2009). Such cost-benefit valuation is thought to contribute to optimizing effort regulation (Kleinflügge et al., 2016; Shenhav et al., 2017). This is consistent with the concept of Hockey’s Compensatory Control Model of effort regulation (1997) where the level of effort mobilization represents the relative commitment of the individual to the task, and is determined by the extent to which the individual regulates or down-regulates the task value. Similarly, motivational intensity theory proposes that the willingness to invest effort into a task is a function of perceived task demand, ability and likelihood that successful performance on the task will achieve a desired outcome (Brehm and Self, 1989; Wright, 1996). This observation would suggest the presence of an interaction model rather than a simple additive effect of task demand and motivation.

Numerous empirical studies have documented the cost-benefit tradeoff strategy of effort allocation in both humans and animals (Hardy, 1982; Forstmann et al., 2006; Croxon et al., 2009; Hillman and Bilkey, 2012; Pasquareau and Turner, 2013; Chiara et al., 2015; Mathar et al., 2015; Westbrook and Braver, 2015; Cheng et al., 2017; Braun and Arrington, 2018). For example, Croxon et al. (2009) completed an fMRI study with young participants performing a sequence of effortful tasks (repeatedly erasing visual targets on the computer screen) with four difficulty levels to obtain secondary reinforcers associated with either high or low money reward. The results demonstrated stronger blood-oxygen-level dependent (BOLD) responses to the high-reward/low-difficulty trials than the high-reward/high-difficulty and low-reward/low-difficulty trials. A behavioral test administered to the same group confirmed that participants chose the more advantageous task options in terms of high reward when the difficulty levels were equal and low and the reward levels were equal.

Little is known about whether effort for language comprehension is subject to the value-based strategic resource management described above for visual and non-linguistic tasks. That is, research has not explored the phenomenon that listeners fail to work harder than is needed to achieve a reward during a given auditory task. In order to investigate this phenomenon, two considerations need to be accounted for in the study design. One is having sufficient levels of task demand. The intentional control of effort allocation is known to occur at levels where information processing is not easy or automatic (Hockey, 1997; Bijleveld et al., 2008). To describe strategic effort selection among task items, the assessment of more
than one task demand level is required. The other consideration is the task goal setting. The task goal in previous studies of listening effort often was predetermined for the participants to maximize accuracy, reward, and/or minimize response time. This approach has led participants to allocate effort in one direction, which was observed in the studies where maximum effort was spent on high-load/high-reward conditions. A task goal that allows participants to demonstrate various possible control patterns of effort allocation among tasks is needed to observe strategic effort control behavior. To the best of our knowledge, only one listening effort study to date has accounted for the first consideration by including two relevant task demand conditions (Koelewijn et al., 2018). However, their findings of the absent interaction effect of reward and task demand only partially represent the contributions of the two factors on strategic effort allocation, because a traditional task goal was used in the study, and the range of the task demand that requires voluntary effort control was relatively narrow (i.e., only two speech intelligibility levels, 85% and 50%, excluding the speech in quiet condition where the performance reached a ceiling of 99.8%).

To address the above design concerns, we developed an auditory comprehension task to examine how listeners make selections on a trial-by-trial basis within the context of global experiment-wise instructions regarding effort expenditure. In this task, five levels of monetary reward and five levels of task demand (speech rate) were independently manipulated, creating a value continuum from low-reward/high-difficulty to high-reward/low-difficulty for the task items. The overall goal in this experiment was to earn a full payment of $50 by selectively and correctly answering a sufficient number of questions associated with differing points within a limited time. The time limit was imposed in order to optimize strategic effort use to successfully obtain their reward rather than maintaining a high level of effort throughout the task. This goal was different from the goal setting for participants in previous listening effort studies. It allowed participants to decide whether and to what extent to invest effort into valued tasks. Both types of decisions are considered part of strategic effort allocation behaviors (Mowen et al., 1981; Hockey and Earle, 2006; Braun and Arrington, 2018) and necessary to fully evaluate the relationship between task demand and motivation. Participants were explicitly informed that correctly answering more questions than needed would not yield more money. Further, as mentioned above, an individually set time limit was given in order to induce a certain level of strategy use as time pressure has been documented to impact decision making strategies (Verplanken, 1993).

Increased pupil dilation with mental activity has been documented for over a century. The size of the pupils at any given time reflects the balance of the sympathetic and parasympathetic systems (Loewenfeld, 1999) and indicates the extent of cortical activation during cognitive and emotional information processing (Siegle et al., 2003). Task-evoked pupil dilation is a sensitive and reliable psychophysiological measure of mental effort allocation (Beatty, 1982). Larger pupil dilation reflects greater processing demands across various cognitive domains (Beatty, 1982; Schlurff et al., 1986; Hoeks and Levelt, 1993; Just and Carpenter, 1993; Hyönä et al., 1995; Granholm et al., 1996; Gerven et al., 2004; Smith et al., 2006; Verney et al., 2010; Hedderik et al., 2012; Wierda et al., 2012; Koelewijn et al., 2015), including listening (Kramer et al., 1997; Piquado et al., 2010; Zekveld et al., 2010, 2018; Kuchinsky et al., 2011, 2013; Zellin et al., 2011; Koelewijn et al., 2012, 2015; Winn et al., 2015; Megarrigle et al., 2017; Ohlenforst et al., 2017). Pupil dilation also reflects reward information processing (Hong and Hikosaka, 2008; Preuschhoff et al., 2011; Anderson and Yantis, 2012; Chiew and Braver, 2013; Han, 2013; Nejjs, 2013; Ariel and Castel, 2014; Chen et al., 2014; Chiara et al., 2015; Takarada and Nozaki, 2017; Hammerschmidt et al., 2017), with pupil enlargement in response to high reward. Although the neural correlates of pupil dilation during reward information processing are still unclear, researchers have observed an association between pupil dilation and strong BOLD response in the reward neural circuit, including the striatum (Schneider et al., 2018).

The pupillary response reflects an overall aggregate of mental resource allocation that is not limited to a specific part of the information processing system (Just et al., 2003). This interpretation of joint sensitivity to affective and cognitive influences suggests that pupil dilation can be used to index overall brain activity (Prehn et al., 2011). Furthermore, when task demand and reward are incorporated into one task, pupil size can reflect both reward magnitude and the difficulty level involved in effort recruitment decisions (Bijleveld et al., 2009; Viswanathan et al., 2012; Koelewijn et al., 2018). Rewards lead to large pupil dilation (indicating high amount of effort engagement) when the reward requires considerable mental resources, whereas little effect of reward on pupil dilation is present during low demanding tasks. Thus, we posit that pupil dilation can serve as a proxy for auditory comprehension effort, with pupil response reflecting both reward anticipation/reactivity and cognitive load associated with item difficulty.

If the reward effects and cognitive load do not interact, pupil dilation is expected to vary linearly with reward and item difficulty. If there is a value-based effort allocation pattern, pupil dilation is expected to vary as a function of reward × difficulty interaction. Different brain regions are engaged in different processing phases of effort regulation tasks (Croxson et al., 2009; Holroyd and Yeung, 2012). Given the evidence that the dissociable brain areas are selectively responsive to reward and task demand information (Macdonald et al., 2000), we expected the reward and difficulty to display phase-dependent effect patterns on auditory comprehension effort, and that the effort could be indexed by pupil dilation based on its neural correlates (Siegle et al., 2011; Zekveld et al., 2014; Kuchinsky et al., 2016).

The present study uses pupil dilation to investigate how listeners integrate reward and task demand information used in deciding effort allocation when strategic use of effort is required to achieve an auditory comprehension task goal leading to a monetary reward. Based on the literature reviewed above, we hypothesized that costs assessed by manipulating task demand (speech rate), and benefits assessed by varying reward points would be weighed against each other in auditory comprehension effort decision-making. Consistent with optimality principles suggested in the effort literature (Rangel et al., 2008; Croxson et al., 2009; Hillman and Bilkey, 2012), it was expected that people would spend more effort, indexed by larger pupil dilation, on the relatively slow rate items with high rewards (high value) than on relatively fast rate items with the same reward (low value). This interaction effect pattern was hypothesized to occur at the response decision stage of the task. There is a paucity of data in the literature about the direct link between effort (indexed by pupil dilation) and performance accuracy. According to Johnson and Payne (1985), the relationships between task demands, effort and accuracy are complex due to diverse strategies available for selection. They suggest that when effort employment is compensatory, high levels of accuracy would be expected. Thus, we hypothesized that the interactive auditory comprehension effort allocation among task conditions might lead to a similar interaction effect of task demand and motivation on participant accuracy.

2. Materials and methods

2.1. Participants

Forty healthy native English-speaking adults (33 females; 34
Caucasian, 6 African American) aged 19 to 32 (M = 22.43, SD = 2.39 years) were recruited from the campus at the University of Pittsburgh. The sample size was based on Potvin and Schutz (2000), assuming a medium effect size, Cohen’s $\gamma = 0.25$ (or $\eta^2 = 0.059$), and a correlation between the repeated measures ($\rho$) of 0.5. Thirty-five participants were required in order to achieve a power of 0.80. An additional 13% of the calculated sample size (i.e., 5 participants) were recruited a priori to compensate for missing data due to multiple factors, including mechanical failure, and excessively noisy pupil or blink data (Siegle et al., 2008).

Young normal-hearing participants were purposely selected for this study to avoid the potential influences of age and hearing loss on decision-making competence (Bruin et al., 2012; Eckert et al., 2017). All participants were screened for hearing, vision and auditory language processing abilities. All participants met the normal bilateral hearing sensitivity criteria of pure-tone air-conduction thresholds less than 20 dB HL at audiometric test frequencies 250 Hz through 8000 Hz in both ears (ANSI, 2004). In order to avoid potential confounds in pupil dilation measurement, the participants needed normal uncorrected vision. All participants had visual acuity of 20/20 or better when tested under the binocular condition using the standard Snellen chart (Bailey and Lovie, 1980). Participants whose scores fell in the normal range (overall score > 14.17) on the version of the Computerized Revised Token Test (CRTT-L) (McNeil et al., 2015) presented at 65 dB SPL were included in order to rule out auditory language processing and comprehension problems.

Participants were monetarily compensated for their participation (i.e., certain reward points redeemed a certain amount of money). Although participants were told that they would be paid at the conclusion of the study according to how many points they earned as described below, in reality, all participants received the maximum possible payment of $50.00. A debriefing letter was provided at the end of the study. This deception was used to motivate the participants to earn the necessary points within a time-limited activity. All of the participants signed a statement of informed consent that had been approved by the University of Pittsburgh’s Institutional Review Board.

Data collection errors and excessive blinking resulted in the elimination of five participants yielding a total of 35 participants on whom data were analyzed.

2.2. Speech materials

Speech rate has not been frequently used as a task demand factor in the listening effort literature. To the best of our knowledge, there are only two published studies to date examining listening effort under speeded conditions. Koch and Janse (2016) used a word recognition paradigm and showed that pupil peaks were non-significantly higher and delayed at faster speech rates (>5.71 syllables/second) than at slower speech rates (<5.71 syllable/second). Müller et al. (2019) adopted a picture-matching paradigm and found that the fast speech rate of 304 syllables/minute evoked significantly higher peak-pupil dilation than their slow speech rate of 182 syllables/minute. The mixed results might be due to the different depth of cognitive processing required in the auditory tasks, in which picture-matching tasks involve deeper comprehensive processing than word recognition tasks. We chose to use speeded speech because, not only can it provide a wide range of task demand within an individual’s capacity (Wingfield, 1996; Wingfield et al., 1995; Niimi and Nishio, 2001; Barac-Gikoja, 2004; Shaw, 2006; Jenstad and Souza, 2007; Adams and Moore, 2009; Nourski et al., 2009; Adams et al., 2012), but it also allows room for individuals to change performance in accordance with goals. Listeners might adjust their resource use through strategic mechanisms as attending to sound sources selectively, accelerating storing and retrieving information from memory, seeking context information to improve understanding, and generating appropriate responses quickly. In order to optimize the probability that the experimental task demand range would overlap with participants’ processing capacity range, 5 speech rates conditions were selected from slow to extremely fast based on a preliminary tests (see supplement material, Fig. 1). The intermediate 3 speech rates were manipulated to help in capturing participants’ processing limit if the extremely fast rate exceeded their capacity.

2.2.1. Speech stimuli construction

A total of 432 structurally equated spoken sentences were developed for the study. They were variants of a reasoning question about the spatial relationship between two of three objects (see examples in Table 1). Prior to each sentence presentation, there was an auditory cue for the listener about the sentence speech rate and the reward value in points.

A male native English speaker read the 432 sentences without exaggerated emphasis on $f_0$ contour and intensity. The speech stimuli were recorded at 44.1 kHz sampling rate through stereo channels in a sound-treated chamber on a Marantz solid state digital recorder (PMD 670) with a Shure SM48 dynamic microphone. The five speech rates (i.e., 130 wpm_slow, 230 wpm_normal, 330 wpm_slightly fast, 380 wpm_fast, and 430 wpm_extremely fast) were chosen based on our preliminary tests (see supplementary materials, Fig. 1). To minimize the adverse impact of time compression on speech intelligibility, the experimental speech rates were generated based on 3 original recordings with slow, conversational and fast speeds from the male speaker.

2.2.2. Speech rate manipulation

According to the literature, when speaking fast, talkers unintentionally change relative attributes of their speech such as pause durations (Jansse, 2004). In order to create speeded conditions similar to naturally produced fast speech, the speech-to-pause ratio was calculated at the 3 speed levels of the recorded stimuli. The speech-to-pause ratio was defined as the duration of speech segments divided by that of pause segments. We linearly regressed the speech-to-pause ratio against the speech rate and applied the obtained linear equation (i.e., $y = 0.0403x - 2.4758$) to generate the 5 target speech rates.

The pause boundaries in the recorded speech were identified and marked automatically by the audio editing program (i.e., Adobe Audition 3.0). Using the heuristics derived from Eisler (1968) and Niimi and Nishio (2001), pause (or silence) was defined as a segment in the waveform that had an intensity lower than 10 dB for
at least 250 ms. The pauses in the sentences were adjusted in duration to achieve the target speech-to-pause ratios before Adobe Audition uniformly expanded or compressed the whole sentence. Based on the measured speech duration and pause duration within a sentence, precisely calculated silent segments were either added to or removed from the original pause segments so that the speech-to-pause ratio reached the target value. Once the speech-to-pause ratio was set, the processed sentence was then time-compressed or time-expanded by Adobe Audition without changing the pitch. The processing algorithm was based on the pitch synchronous overlap and add method (PSOLA) (Moulines and Laroche, 1995; Kawahara et al., 1999).

None of the time-compression ratios of the speech rate manipulations exceeded 30%. The full-version of a speech stimulus token comprised a 2 s auditory cue phrase, a 2.5 s silent presentation period, a 0.5 s 1000 Hz pure tone signaling the onset of the sentence to help direct attention, and a 3.7–12.5 s speech presentation depending on the speech rate condition. Following each response, there was a 1 s interval between the button push and the onset of the next trial. An example of a speech stimulus token is shown in Fig. 1.

### Table 1

Examples of auditory stimulus materials and expected correct response. The durations of sentences for each speech rate condition are 12.5 s (slow), 7.1 s (normal), 4.9 s (slightly fast), 4.3 s (fast) and 3.7 s (extremely fast).

| Cue (2.5s) | Sentence | Correct response |
|-----------|----------|------------------|
| Slow 5 points | The stone is on the left of the ball, the kite is on the right of the ball, where is the stone in relation to the kite? | Left button push |
| Fast 9 points | The kite is in front of the stone, the stone is in front of the ball, where is the kite in relation to the stone? | Front button push |
| Extremely fast 1 point | The kite is on the right of the ball, the stone is on the right of the kite, where is the kite in relation to the ball? | Right button push |
| Normal 3 points | The ball is in back of the kite, the kite is in back of the stone, where is the ball in relation to the stone? | Back button push |
| Slightly fast 7 points | The stone is in front of the ball, the kite is in back of the ball, Where is the kite in relation to the Stone? | Back button push |

## 2.4. Procedures

All testing was carried out in a 1.5 × 2.2 m double-walled sound-treated booth that meets specifications for maximum permissible ambient noise levels (ANSI, 2003). The participants were seated at a table in front of a computer monitor with a 5-key response keypad in front of them. The experiment instructions and a fixation image were displayed on the monitor and the stimuli were presented from a desktop computer under experiment software control (SuperLab 5, Cedrus, Phoenix, Arizona). The speech signals were fed through a diagnostic audiometer (Beltone, 2000) to ER-3A insert earphones. Presentation level was calibrated with a speech-shaped noise matched to the Long-Term Average Speech Spectrum (LTASS) of the speech produced by the male speaker used in this experiment. The sound intensity level calibration was performed using a Larson-Davis 824 sound level meter with a KEMAR 2 cc coupler to ensure that the output from each insert earphone was 65 dBSPL.

Prior to the experiment, participants performed a sequence of four blocks of 25 practice trials (supplement materials Table 1) to ensure familiarity with the response keypad and the speech stimuli, and to ensure that they understood the task and directions. Based on the continuous monitoring of eye and pupil images in the eyetracking system, all participants were able to fixate on the cross image throughout trials including during button press window.

In the experimental procedure, 250 sentences (5 speech rate levels x 5 reward levels x 10) were presented in 5 blocks of 50 sentences with breaks between blocks. In order to control for both order and carryover effects, a digram balanced Latin square table was used to create the test conditions of the speech rate and reward point combination. As the level of independent variables was an odd number, two Latin square tables were generated so that each reward point level was followed by a different point level with the same probability. Participants were randomly assigned to receive one of the testing condition sequences shown in Table 2.

The participants were informed both verbally and in writing (shown on the computer screen) that: “There are 5 blocks in total. Please remember, each correctly answered question will earn you certain points (1, 3, 5, 7 or 9), and each point is worth 5 cents. The maximum points you can possibly get is 3700, but you only need 1000 points to get the full payment ($50). Your goal is to EARN the FULL PAYMENT within a limited time. A clock ticking sound will appear when you have spent half of your time in each block. Your response won’t be registered until the sentence presentation is complete. Don’t forget to look at the cross image on the screen all the time.”

The participants were allowed to press the response button at any time after the speech sentence offset. They were required to fixate on a cross-hair presented in the middle of the screen while listening to the speech stimuli and while responding. The participants could either answer the question by pressing one of the four buttons (i.e., left, right, front, and back) on the keypad, or skip to the
next stimulus by pressing the 5th button in the middle of the keypad, depending on their effort control strategies. The time limit was adjusted individually based on the amount of time they spent on the 25-trial simulation practice block. In order to avoid missing data, there was no actual time limit for the experiment. The fixation image remained on the screen and pupil dilation was tracked throughout the block.

2.5. Pupillary data analysis strategy

2.5.1. Data selection and cleaning

Pupillary data were cleaned using procedures described in Siegle et al. (2008). The data cleaning process included blink identification and interpolation, smoothing, and artifact trial removal. As noted in Siegle et al. (2008), blinks were identified as large changes in pupil dilation occurring too rapidly to signify actual dilation or contraction. Specifically, blinks were coded as samples with estimated pupil diameter meeting any of the following criteria: 1) below 1 mm; 2) below the minimum diameter in a subject’s waveform +0.1 mm; 3) below the median diameter minus 4 mm; 4) below two times the interquartile range below the 25th percentile (i.e., the Tukey extreme outlier hinge). Consistent with Granholm et al. (1996), trials that required interpolation on 50% or more of all samples within a trial were completely removed from further analysis. This liberal removal criteria was applied because fairly long trials were used in the auditory comprehension task and the blinks were not expected to occur at the same place on every trial. The analysis on the blinks showed a mean of 20% blinks at any given sample, and the blinks were not systematically related to reward, task demand or reward x task demand interaction when using this criteria (see supplement Table 2 for detailed analysis on the blinks).

Linear interpolations beginning 4 samples before and ending 9 samples after a blink replaced blinks throughout the data set. This technique prevented interpolation to poor pre- and post-blink pupil estimates due to partial lid closure. Afterwards, a three-point moving average filter was applied twice to smooth the de-blinded trials and to remove any high frequency artifacts. Finally, the average pupil diameter in the 10 samples (0.17 s) preceding the onset of the tone was subtracted from the pupil diameter to produce pupil dilation difference score indices. Determining the appropriate baseline period from previous studies is difficult because it varies greatly from study to study. Some researchers have used long baseline periods of up to 1000 ms, while other authors (e.g., Kret and Sjak-Shie, 2018; Granholm et al., 2016; Mathar et al., 2015; Steinhauer et al., 2015; Siegle et al., 2011) have used shorter ones. Long baselines have the disadvantage of being susceptible to pupil-size fluctuations during the baseline period from such artifacts as blink, and short baselines have the disadvantage of being susceptible to recording noise (Mathot et al., 2018). Given that there is no established norm in psychophysiological research using event-related designs (Kelbsch et al., 2019), the short baseline window was used in the current study to minimize the likelihood of blink artifact in the baseline computation (Siegle et al., 2008; Mathot et al., 2018). Post hoc analysis using a 1s long baseline showed equivalent results (see supplement Table 2) that selected in this study.

The pre-stimulus baseline served as a uniform baseline across the processing time windows within trials, as auditory comprehension effort was conceptualized as an integrative construct that comprises diverse effort components necessary for task completion. Additional analyses using a separate baseline for each processing window of interest (window-specific baseline) were performed to separate effort components at each processing stage.

Intentionally incorrect trials where participants pressed the skip button were maintained in the analysis. Participants were instructed to use the skip button when they decided not to spend effort on answering the question. Therefore, those skipped trials indicated that the participants focused attention on the task and followed the instructions. Pupil dilations in those trials were valid to represent their effort exertion decision. Other incorrect and missed trials were excluded. Data selection and cleaning procedures resulted in the elimination of 38 trials (out of 250 trials) on average per participant (SD = 12.5, Median = 24 trials) due to contaminated pupillary data. Five participants were dropped due to trial removal rates greater than 30%.

2.5.2. Determination of windows of significant differences

The effect of reward and speech rate on pupil dilation across the entire trial was of interest in this study as previously mentioned that distinct information processing might engage in different phases of a cognitive task. The pupillary data at each sample point were included in the analysis. We segmented each trial into 3 temporal phases, including the cue window defined as the time period between the onset of the auditory cue and the onset of the speech sentence, the sentence presentation window defined as the time period between the onset and offset of the sentence stimulus, the response window defined as the time period between the offset of the sentence presentation and the onset of the next trial.

Contrasts on pupil dilation were examined via a two-level analysis model. The first level included a linear multiple regression with predictors of reward, speech rate and the interaction of reward and speech rate at each time point along pupil dilation for each individual. The second level included an aggregation across participants and the subsequent t-tests on the regression coefficients at each time point. To control type I error for this large number of tests, a technique originally described by Guthrie and Buchwald (1991) and regularly employed in analyzing pupil dilation waveforms (e.g., Siegle et al., 2008) was used. This technique involves using Monte Carlo simulations to estimate the number of consecutive significant differences that were long enough to be judged as not occurring by chance with p < .05 given the temporal auto-correlation of the data. Thus, contiguous sample-by-sample tests were considered replications.

Single lag autocorrelations were calculated as recommended by Guthrie & Buchwald’s original manuscript. 1000 Monte Carlo
simulations of data with the same level of autocorrelation and number of participants as in the empirical dataset were performed, storing the number of significant tests in a row in the simulated data for each. After removing four principal components which could reflect task-related responses (Guthrie and Buchwald, 1991), the autocorrelation of the residual condition-related pupillary reactivity waveforms (i.e., mean baseline-corrected pupillary responses for each condition) was 0.98. Monte Carlo simulations suggested that a window of 85 ms (i.e., 1.41 s given the 60 Hz sampling rate) consecutive regression tests on pupillary reactivity along the time course (significant at \( p < .1 \)) would yield a window of differences significant at \( p < .05 \). When the \( p \)-values were reported for an entire time window, they represented significance tests of the mean regression coefficients in a window of consecutive significant differences.

Additionally, in order to examine whether using pupillometry could evaluate effort at individual level, the Chi square analysis was performed on the proportions of participants for whom the individual results (i.e., main effect of speech rate, main effect of reward and the interaction effect) were significant.

### 2.5.3. Analysis within temporal windows of interest

Because sentences with different speech rates had different durations, pupil waveforms by each valid trial in stimulus presentation window and the end of the sentence presentation were unaligned in time, with slow rate conditions having much longer durations than fast rate conditions. This unequal temporal window may confound the comparison of pupil dilation across conditions if the conventional period amplitude variables are used, such as dilation peak, mean dilation or area under the curve. In order to compare pupillary reactivity across conditions without this confound, a different analysis approach was used for the sentence presentation window. Specifically, trials segments of the pupillary waveforms were aligned to the beginning of the sentence for analyses of the cue window. The segments were then aligned to the end of the sentence for analyses of the response window. Two baseline corrections were used in analyzing pupillary response in these two windows. One was a uniform baseline correction, which used the baseline preceding the sentence presentation throughout the trial. This baseline was used to examine the effect of task demand and reward on auditory comprehension effort as an integrative construct. The other approach was a window-specific baseline correction by using the baseline preceding each specific window. This approach helped to parse the integrative effort construct indexed by pupil dilation. For example, it was used to separate the effort component sensitive to perceptual demand of listening (i.e., carried-over from the sentence presentation window) from the potential task-unrelated component (e.g., pupil size variation susceptible to time) from the effort component sensitive to reaction to the task.

As Jones et al. (2010) has demonstrated that the pupillary motility occurring at the frequency of task stimuli reflects task-related cognitive resource allocation, their statistical technique of using wavelet decomposition to quantify pupillary motility at the speech presentation frequency were adopted for the sentence presentation window.

Wavelet analysis was performed for the entire sentence presentation window to estimate variation in pupillary motility at the stimulus frequency in this window for all conditions. Continuous wavelet power using a Morlet mother wavelet at the stimulus frequency was estimated for each of the valid pupillary trials (as in Jones et al., 2010) for each participant. Pupillary log wavelet power values, scaled by frequency, were computed for each of the 25 conditions and served as the dependent variable in analyses for the sentence presentation window.

### 3. Results

#### 3.1. Auditory comprehension effort indicated by pupillary data

Pupillary waveforms displayed an elevated dilation during the auditory cue presentation in the cue window, a trough at the 1000 Hz tone signal before the sentence presentation, and a peak dilation in the response window. The 2.5-s silence between the cue offset and the sentence onset was the longest interval within a trial that allowed the pupil to return to resting size. As expected, the minimum pupil size was reached at the end of the cue window (preceding sentence onset) for each condition (see supplement Figure 4 left panels). This resting pupil size served as the uniform baseline for dilation calculations. Fig. 2 shows the grand mean pupillary waveforms and the pupillary wavelet powers by speech rate (upper panel) and by reward (lower panel) along the time course of task trials across all 35 participants. The pupillary waveforms during the sentence presentation window were not shown in Fig. 2 because of the alignment issue (see the actual pupil response plot in supplement material Fig. 4).

**Cue window effects.** During the cue period (panels on the left in Fig. 2), there was a significantly long window of consecutive tests for the main effect of speech rate extending from 0.02 to 4.02 s after the beginning of a trial. Speech rate was positively associated with pupil dilation after adjusting for reward in this window, \( t \) \( (34) = 6.59, p < .01, d = 1.11 \), with 0.02 mm increase in dilation for every unit increase in speech rate. There also was a significant window of positive association of pupil dilation with reward after adjusting for speech rate, from 0.02 to 1.70s, \( t \) \( (34) = 3.15, p < .01, d = 0.53 \), with 0.01 mm increase in dilation for every unit increase in reward. The same effects remained when the mean pupil diameter in the initial second of the trial was used as the baseline for the cue window response. The significant main effect of speech rate extended from 0.23 to 5.22 s after the beginning of a trial, \( t \) \( (34) = 6.69, p < .01, d = 1.13 \). The main effect of reward was also significant in the 0.52–4.75 s time period after the trial onset, \( t \) \( (34) = 4.63, p < .01, d = 0.78 \). These was no interaction effect of speech rate and reward during this time window.

**Sentence presentation window effects.** For the sentence presentation window (panels in the middle in Fig. 2), speech rate was significantly positively associated with wavelet power of pupillary dilation at the stimulus frequency with 0.04 mm increase in dilation for every unit increase in speech rate, \( t \) \( (34) = -4.48, p < .01, d = 0.76 \), as was reward with 0.05 mm increase in dilation for every unit increase in reward, \( t \) \( (34) = 3.10, p < .01, d = 0.52 \). No significant interaction effect of speech rate and reward was observed in this window, \( B < 0.001, t \) \( (34) = 0.47, p = 0.639, d = 0.08 \).

**Response window effects.** As pupillary response is a delayed response, in order to quantify the effort response of all stages (e.g., listening, response selection and execution) involved in auditory comprehension, the pupillary response analysis window was expanded from the sentence offset to 4s after (panels on the right in Fig. 2) to be conservative. In this time period, there was a significant interaction effect of speech rate and reward from window 0–2.97s relative to the offset of the sentence presentation, \( B = 0.004, t \) \( (34) = -2.84, p = .01, d = 0.48 \), suggesting that pupil dilation change as a function of speech rate was modulated by reward levels. The 2.97 s time window was sufficiently long to capture the effort response that did not overlap with the pupil activity from the preceding listening phase of the trial because the longest condition-related mean response time (from sentence offset to button press) was 1.58 s. When the mean pupil diameter in the 1 s preceding the response was used as the baseline for the response window, the interaction effect of speech rate and reward remained significant, from window 0–2.13s relative to the sentences offset,
To illustrate the results, the averaged predicted pupil dilation based on the mean regression coefficient in each condition is plotted in Fig. 3 (left panels) and surface plots with all of the participants’ data are displayed in the right panels. With the uniform baseline (Fig. 3 upper panels), at the low reward levels of 1, 3 and 5 points, pupil dilation did not significantly vary with the speech rate, $B = .003$, $t(34) = .867$, $p = .392$, $\eta^2_p = .022$, $B = -.004$, $t(34) = -.895$, $p = .377$, $\eta^2_p = .023$, and $B = -.011$, $t(34) = -1.593$, $p = .120$, $\eta^2_p = .069$ respectively. However, when the reward was high, the linear association between pupil dilation and speech rate became significant. Specifically, pupil dilation significantly decreased as the speech became fast at the reward of 7 points, $B = -.020$, $t(34) = -2.146$, $p = .039$, $\eta^2_p = .119$, and 9 points, $B = -.029$, $t(34) = -2.519$, $p = .017$, $\eta^2_p = .157$. The differences among the regression slopes were significant, $F(4, 136) = 9.653$, $p <.001$, $\eta^2_G = .221$. Pairwise comparisons with Bonferroni adjustment revealed that pupil dilation was significantly more negatively associated with speech rate with the reward of 7 points than with low reward levels (i.e., 1, 3 and 5 point), $F(1, 34) = 9.503$, $p = .041$, $\eta^2_G = .218$, $F(1, 34) = 9.920$, $p = .034$, $\eta^2_G = .226$, and $F(1, 34) = 9.650$, $p = .038$, $\eta^2_G = .221$. Similar results were found when comparing the relationship of pupil dilation and speech rate between the 9 points reward condition and the low reward conditions, with $F(1, 34) = 10.708$, $p = .025$, $\eta^2_G = .240$, $F(1, 34) = 10.942$, $p = .022$, $\eta^2_G = .243$, $F(1, 34) = 10.448$, $p = .027$, $\eta^2_G = .235$, respectively. No significant difference on the regression slope was found between rewards of 7 and 9 points, $F(1, 34) = 3.757$, $p = .009$, $\eta^2_G = .100$.

With the 1 s baseline prior to the response, the interaction effect of speech rate and reward remained significant with differences, mainly in the low reward conditions (Fig. 3 lower panels). Specifically, at the low reward levels of 1, 3 and 5 points, pupil dilation significantly decreased with the speech rate, $B = -.038$, $t(34) = -3.126$, $p <.01$, $\eta^2_p = .237$, $B = -.032$, $t(34) = -2.855$, $p = .010$, $\eta^2_p = .194$, and $B = -.028$, $t(34) = -2.563$, $p = .022$, $\eta^2_p = .160$, respectively. At high reward levels, pupil dilation did not significantly vary across the speech rate, with $B = -.016$, $t(34) = -.963$, $p = .092$, $\eta^2_p = .103$ for the 7 points condition and $B = -.004$, $t(34) = -.591$, $p = .134$, $\eta^2_p = .023$ for the 9 points condition. The differences among the regression slopes were significant, $F(4, 136) = 10.272$, $p <.001$, $\eta^2_G = .246$. Pairwise comparisons with Bonferroni adjustment showed that pupil dilation was significantly more negatively associated with speech rate with low reward condition 1,3 and 5 points than with the 9 points high reward, $F(1, 34) = 9.872$, $p = .032$, $\eta^2_G = .272$, $F(1, 34) = 9.910$, $p = .034$, $\eta^2_G = .226$, and $F(1, 34) = 9.650$, $p = .038$, $\eta^2_G = .221$. The negative relationship in the 1 and 3 points
conditions was also significantly stronger than in the 7 points condition, $F(1, 34) = 10.179, p = .026, \eta^2_p = .258, F(1, 34) = 9.347, p = .039, \eta^2_p = .213$. No other significant regression slope differences were found.

These results suggested that regardless of baseline correction approaches, effort indexed by the pupil dilation spent on completing the auditory comprehension task after the speech presentation did not simply vary linearly with reward or speech rate, instead, it varied as a function of reward by speech rate interaction. More effort was invested on relatively easy items than on difficult ones, indicating an allocation strategy based on the cost-benefit valuation. This phenomenon was independent of the carry-over pupillary response from the preceding listening phase and possible pupil change over time that was unrelated to the task.

In addition to the significant interaction effect, there was a significant main effect of speech rate extending from 0 to 6 s with 0.04 mm increase in dilation for every unit increase in speech rate, $t(34) = 8.35, p < .01, d = 1.41$, as well as a significant main effect of reward extending from 0 to 2.77 s with 0.01 mm increase in dilation for every unit increase in reward, $t(34) = 2.79, p = .01, d = 0.47$. These two main effects on pupil dilation relative to the uniform baseline were similar to the cue and sentence presentation windows. However, when the window-specific baseline was used, main effects were eliminated.

Pearson correlation analysis was performed on the effect indices (i.e., regression coefficients) derived for 3 window periods. The results showed that the influence of task demand and reward on pupil activity at different auditory comprehension processing stages were highly correlated (see Table 3), which suggested a low separability of cognitive components attributing auditory comprehension effort.

Further, the number of individual’s $p$ values less than 0.05 for each effect (i.e., main effect of speech rate, main effect of reward and the interaction effect) in this window were computed. The
result of Chi square analysis on the proportion of $p < .05$ indicates that the individual level of significance did not occur by chance, with main effect of speech rate (26 out of 35), $\chi^2 = 35.110, p < .005$, main effect of reward (22 out of 35), $p < .005$, and interaction effect (20 out of 35), $\chi^2 = 22.216, p < .005$ respectively.

### 3.2. Task performance

A 5*5 within-subjects analysis of variance was performed on accuracy as a function of speech rate (slow, normal, slightly fast, fast and extremely fast) and reward (1 point, 3 points, 5 points, 7 points and 9 points) for the same 35 participants. As the assumption of sphericity was violated, statistics are reported with Huynh-Feldt correction. There was a significant interaction effect of reward and speech rate on accuracy, $F(12.982, 493.301) = 2.634, p = .001, \eta_p^2 = .065$, indicating that the pattern of the performance accuracy among the reward levels was significantly dependent on speech rate (see Fig. 4). At slow and normal speech rates, accuracy did not change significantly as a function of reward; however, reward was significantly associated with increased accuracy in the fast and extremely fast conditions, $F(4, 152) = 4.469, p = .002, \eta_p^2 = .105, F(4, 152) = 2.669, p = .034, \eta_p^2 = .066$, respectively. In both rapid speech conditions, the performance was significantly better when high rewards (5, 7 and 9 points) were received than when low reward (1 and 3 points) were received, $F(1, 38) = 48.880, p < .001, \eta_p^2 = .563$, and $F(1, 38) = 52.214, p < .001, \eta_p^2 = .579$. There was no significant difference in the accuracy among 5, 7 and 9 points reward conditions at fast speech rate, $F(2, 76) = 1.607, p = .207, \eta_p^2 = .041$, or at extremely fast speech rate, $F(2, 76) = 1.750, p = .181, \eta_p^2 = .044$.

There was a significant main effect of speech rate on accuracy averaged across reward, $F(2.801, 106.439) = 50.006, p < .001, \eta_p^2 = .568$. Performance was significantly lower at faster speech rates than at slower rates in all pairwise comparisons except for the slow/normal pair; and slightly fast/fast pair comparisons. There was no significant main effect of reward on accuracy averaged across speech rate, $F(2.961, 112.535) = 2.690, p = .050, \eta_p^2 = .066$.

In order to examine the relationship between auditory comprehension effort (indexed by pupil dilation) and the resulting performance accuracy, a mixed effects analysis with two levels was conducted on the mean accuracy by the mean pupil dilation (0–2.97s post stimulus time window) with participants as a random factor on the second level. There was a significant negative association between the pupil dilation and the accuracy after accounting for the individual random effect, $B = -.176, t (735.749) = 2.511, p = .012$. However, we observed a ceiling effect on the performance at slow rate conditions. To avoid a biased negative relationship, the accuracy at those conditions was excluded from the analysis. Results from the mixed effect analysis showed no significant association between the pupil dilation and performance accuracy, $B = .039, t (613.200) = .668, p = .504$. These results suggested that high effort did not necessarily lead to high performance accuracy in the auditory comprehension task under the context of global goal setting.

4. Discussion

The goal of the current study was to understand 1) how reward and task demand interact on strategic effort allocation on an auditory comprehension task in normal hearing listeners, 2) whether auditory comprehension effort allocation is value-based and abides by the optimal cost-benefit policy, and 3) whether auditory comprehension effort relates to ultimate performance accuracy. The three research questions were addressed by manipulating both reward value and speech rate across a relatively wide range, and by measuring pupil dilation as the index of effort and percent correct as the index of performance accuracy. We hypothesized that, under a global goal setting where selectively responding to the experimental speech stimuli was required to achieve the full payment, the impact of task demand (speech rate) on auditory comprehension effort allocation would depend on the level of reward (point value). People would spend more effort, indexed by larger pupil dilation, on the relatively easy items with high rewards (high value) than on relatively difficult items with the same reward (low value). This interaction effect pattern was hypothesized to occur at the response stage of the task, and leading to a similar performance accuracy display. Our hypotheses were partially supported.

To the extent that pupillary data are, as we assumed, proxies for effort in the present study, they suggested that effort allocation during auditory sentence comprehension is driven by both reward and task demand. Their effects depended on the information processing stage during the task. Additive effects of reward and task demand were present in the cue and sentence presentation windows, whereas an interaction effect was present in the response window, which is consistent with the notion that effort allocation was taxed within the decision-making/response selection process in this study.

During the cue window, participants received auditory information about speech rate and the reward associated with the

| Table 3 |
| --- |
| The correlation between indices derived for each period. |
| --- |
| Coefficients of speech rate effect | .654** | .816** | .649** |
| Coefficients of reward effect | .328 | .372* | .615** |
| Coefficients of interaction effect | .425* | .702** | .493** |

**. Correlation is significant at the 0.01 level (2-tailed); *. Correlation is significant at the 0.05 level (2-tailed).
upcoming sentence. Among other cognitive processes, effort at this stage included auditory attention, perception, comprehension and valuation. As Macdonald et al. (2000) demonstrated, strategic control at the preparation stage is an instruction-related process that focuses on representing and maintaining task demands. Such preparatory set construction could yield the observed main effects of speech rate and reward cue on pupil dilation within this time frame (see Fig. 2, cue window). These main effects were unlikely a result of carry-over effects from the response phase of the preceding trial, because when the pupillary responses were adjusted relative to the baseline prior to the cue window instead of the uniform baseline, the effects remained, and even sustained longer. This result implies that the effort allocated on task preparation is robustly impacted by the auditory cue of task demand and reward information.

During the sentence presentation period, participants used whatever strategy (e.g., finger alignment, visualizing) they chose to solve the spatial questions while listening to the spoken sentences. Effort at this stage likely included auditory perception and comprehension, storage and encoding in working memory, response planning and physical effort if finger strategies were used. During this listening period, effort significantly increased as speech rate and the designated reward point increased (see Fig. 2, sentence presentation window), suggesting that the effort spent on listening to the sentence stimuli for the purpose of comprehension, solving the problem and obtaining the reward points was linearly driven by both speech rate and reward. These results are in line with Koelewijn et al.’s (2018) finding that the pupil activity increased with increased reward independent of task demand in the sentence recognition. The results also were consistent with Richter’s (2016) finding that the effort indexed by cardiovascular pre-ejection period reactivity in an auditory discrimination task was highest when task demand and success importance were both high. At this stage, the pattern of effort allocation among the speech rate levels might not vary as a function of reward levels because cognitive processing focused on perceiving the auditory signal and solving the problem (i.e., implementing effort) rather than weighting cost against benefit. Therefore, it was not surprising that no significant interaction effect was observed.

In the post stimulus presentation period, participants made decisions and answered spatial questions quickly given the time pressure and the overall goal of gaining the full payment. Effort in this period included the decision-making, attention to answering the question, physical response, performance monitoring, as well as attention to maintaining the task goal, all of which might have contributed to the formation of the pupil dilation peak. This implies that auditory comprehension effort has distinct contents in this phase from other phases of a trial. Pupil activity in this time window might have indexed a combination of diverse sources of effort that were necessary for completing the auditory comprehension task. When the pupil reactivity waveforms were corrected based on the mean pupil diameter prior to the sentence onset (uniform baseline), the main effects and the interaction effect were preserved. When attempting isolate the mixed effort components by using the baseline prior to the sentence offset, the main effects consistent with the well-established listening effort effects were eliminated, while the interaction effect was preserved. This result supports that the processes reflected by the pupil dilation occurring after the presentation of the auditory stimulus are not attributed solely to listening (Zekveld et al., 2018). Given that participants were required to push keys to respond as quickly as possible during this time interval, the pupil dilation likely comprised the accumulative pupillary reactivity carried over from the listening period (listening effort) and the reactivity associated with response selection and execution. The latter component was sensitive to the interaction between task demand and reward.

As the nature of the task goal in the current study required a strategic use of effort, the interaction effect of reward and speech rate on effort in this period supported our hypothesis of cognitively and affectively optimized effort allocation (see Fig. 3). Effort was no longer prioritized to high demanding and high rewarding tasks as an additive model would suggest, rather, it was allocated to relatively easy tasks (i.e., high values), indicating a type of strategic effort control. This behavior is consistent with previous research on strategic recruitment of resources in other fields and the Motivational Intensity Theory (Brehm and Self, 1989) that proposes that the effect of task difficulty on effort exertion depends on the magnitude of potential motivation. Less effort mobilization is expected when the task difficulty outweighs the value of the potential gain than when effort is considered worthwhile. The lack of impact of task demand on effort at low reward levels was consistent with the prediction of the Framework of Understanding Effortful Listening (Pichora-Fuller et al., 2016) and previous empirical listening effort studies (Picou, 2014; Koelewijn et al., 2018).

Both Koelewijn et al.’s (2018) study and the current study investigated the effect of reward and task demand on effort (indexed by pupil dilation) within individuals’ cognitive capacities. A significant interaction effect was observed in the current study but not by Koelewijn et al. (2018). One possible reason is that the task goal was set differently. In Koelewijn et al.’s (2018) study, participants were encouraged to earn as much reward as possible by repeating 70% or more of the sentences. In contrast, participants in our study were encouraged to intentionally “give up” items that were perceived to be not worth the effort, given that gaining the full payment within a time limit only required a comprehension accuracy of 50%–80% depending on which items were selected. These two goals might have changed how people processed the reward and task demand information in deciding effort allocation.

Another possible reason might be the measurement of different effort constructs. The pupillary reactivity in the post-stimulus interval is the most commonly analyzed period in effort research. Pupil dilation is a delayed response (Beatty, 1982). During its time approaching the peak after the stimulus offset, it likely overlaps with pupillary responses that associate with reactions required by the task (e.g., repetition, recall, answer selection or cognitive preparation for those actions). This has made the effort indexed by a single pupil value (e.g., peak, mean) difficult to decipher. In Koelewijn et al.’s (2018) study, the peak and mean pupil dilation were the main dependent variables. Although there was a 0.5s pause before participants were prompted to repeat sentences, the average pupil dilation peaks occurred after the 0.5s pause, indicating some overlap of the pupil reactivity for listening and other processes were reflected in the reaction. Moreover, their pupillary traces were corrected using a baseline prior to the sentence onset, similar to the uniform baseline used in the current study. Therefore, the construct measured in the two studies were similar and may both involve more than just listening effort and reflect the integrative process of sentence comprehension.

By using the window-specific baseline, we were able to identify that listening effort was one primary component of auditory comprehension effort in the context of the current study. Other factors involved in effort allocation that are also essential for auditory comprehension and might be more sensitive to the item value judgment. Pearson correlations between the regression coefficients derived for each of the 3 window periods suggested that the different auditory comprehension processing stages were moderately to highly correlated (see Table 3). Thus, we conclude there were some common and some independent influences of reward and task demand at each stage. Effects appear to have been the most common for task demand, and most separable for reward.
Further research is needed to evaluate the contribution of different effort components in similar tasks.

It might be argued that the counterintuitive decrease of effort at the fast and extremely fast speech conditions with high rewards was due to the task demand exceeding participants’ cognitive capacity. However, based on an inspection on each individual’s condition-dependent pupil dilation, performance accuracy and response time, none of the participants was found to display a sign of disengagement or giving-up, which would meet the following criteria: 1) an inverted U shaped pupil dilation across task demand levels; 2) a steep downward slope of accuracy as a function of task demand, with the lowest accuracy at or less than chance level; and 3) an inverted U shaped response time across task demand levels, with the shortest response time occurring at the highest demanding task condition. Even though the performance decreased with task demand in general, the response time reliably increased when speech became fast (supplement material Figure 6), suggesting an active task engagement.

Taken together, the pupillary findings of the current study provide evidence of an interaction of reward and task demand on auditory comprehension effort. The findings are in line with the effort associated cost-benefit decision-making literature (Hardy, 1982; Rangel et al., 2008; Croxson et al., 2009; Hillman and Bilkey, 2012; Shenav et al., 2017), and high level voluntary effort regulation theories (Brehm and Self, 1989; Hockey, 1997). Comparable to other cognitive activities, the effort allocation indexed by pupil dilation during an auditory comprehension task appears to be a value-based decision-making process and highly dependent on motivation. Our results support the FUEL framework’s (Pichora-Fuller et al., 2016) notion that motivation and task demand, together, drive effort for listening. However, our results demonstrate a different interaction pattern from what the FUEL proposes. In the FUEL framework, effort is hypothesized to monotonically increase as the task demand increases, except for the very low motivation where effort does not change across task demand levels. This positive relationship gets stronger as motivation becomes higher. In contrast, our pupillary data suggest that at high motivation levels (i.e., 7 and 9 reward points), larger effort was spent on relatively easy items instead of difficult ones, consistent with the idea that effort is expended on rewards that people view as clearly attainable. This explanation is supported as performance did drop precipitously for the highest difficulty levels even in the high reward conditions. The finding is similar to the drop in pupillary reactivity in very difficult conditions in a digit span, and earlier drop in pupil dilation with lower cognitive resources (Granholm et al., 1997).

The present study provided foundational evidence of strategic auditory comprehension effort allocation behavior in young normal-hearing individuals and requires replication in other populations. It supplements the FUEL framework with an alternative relationship between motivation and task demand factors when driving effort allocation in auditory tasks. Specifically, effort for auditory comprehension was flexibly regulated based on reward/task demand value judgment (interaction effect) when the task goal allowed listeners to achieve the maximum reward without 100% accuracy. This alternative effort pattern suggests that the association between the nature of task goals and how listeners process the auditory task demand and motivational information should be accounted for in future research.

The strategic value-based effort allocation phenomenon observed in the present study could serve as a reference or a goal of audiological rehabilitation for people with hearing loss in experimental and clinical settings. The findings support the incorporation of affective factors (e.g., reward) and the utility of the value-based strategic effort allocation paradigm in investigating how clinically important factors (such as hearing loss and age) might influence strategic effort allocation behavior. However, special attention should be paid on the choice of effective reward type when replicating the experiment in other populations, because evidence has shown that age impacts the sensitivity to various rewards. For example, older adults are more motivated to social and health-related rewards than monetary rewards, and young people tend to be motivated by monetary rewards (Rademacher et al., 2014; Seaman et al., 2016).

The results of this study support manipulating speech rate as task demand variable in studying effort. The significant effect of speech rate across trials was consistent with Müller et al. (2019) that faster rate induced more effort mobilization. The current study also support the use of pupillometry as an appropriate method to examine auditory comprehension effort at both group and individual levels, as pupil diameter did not monotonically vary as a function of reward or task demand. Instead, it varied as a function of combinations consistent with human effort investment behavior and the implications of underlying neural activities. Reporting on the effects of reward and task demand on pupil dilation, including the entire temporal window of the auditory comprehension task, could be helpful for further understanding the components of auditory comprehension effort at various cognitive processing stages. Future studies might find it beneficial to document the cognitive neural correlates of auditory comprehension effort and whether effort allocation is differentially influenced by determinant factors with time as demonstrated in the current study.

Regarding the influence of reward and task demand on ultimate performance accuracy, our results showed that under relatively easy conditions in which task demand was well within the individuals’ capability, reward did not significantly influence performance. However, under relatively difficult conditions where there was more room to increase performance, the reward significantly improved performance. These data suggested that when reward motivated the listeners, they utilized their capacity and found strategies to maximize performance.

The hypothesis that the reward and speech rate would show a similar interaction effect pattern on pupil dilation and performance accuracy was not supported. We failed to observe a monotonically linear relationship between auditory comprehension effort (indexed by pupil dilation) and performance accuracy. On one hand, this result might reflect the complexity of the relationship between effort and accuracy in strategy selection. Central to a cost-benefit analysis of strategic selection is the existence of an accuracy-effort tradeoff, a continuum of rules in which increases in effort result in increases in accuracy (Johnson and Payne, 1985). However, this rule varies with context such as the number of dominated alternatives (Klein and Yadav, 1989), anticipation and experience (Fennema and Kleinmuntz, 1993). Recent evidence on listening effort (indexed by pupil dilation) has shown that factors that change task demand also can alter the effort-performance relationship. Ohlenforst et al. (2018) and Wendt et al. (2017) reported that noise reduction algorithms in hearing aids significantly reduced the effort required for successful speech communication while improving the performance accuracy. The negative relationship between effort and accuracy at slow speech rate condition in the present study is consistent with their findings. In contrast, the lack of significant effort-performance association in the range from “normal rate” to “extremely fast” is consistent with Koelwijn et al.’s (2018) study in which speech recognition performance was not elevated in accordance with an increase in effort due to the task demand and reward change. In the current experiment, the absence of the effort-performance relationship might be due to the fact that participants had numerous alternative selections to achieve the
goal of earning the $50 reward such as investing high effort on difficult items with high reward, investing low effort on easy items regardless of reward, or a mixed strategy. This result might also be attributable to the near perfect performance on the task, which prevented the examination of the relationship of pupil dilation to performance accuracy.

The range of speech rates was not large enough to surpass the capability of some participants. There were a total of 29 participants (74.4%) who gained enough reward points to receive the full $50 payment with accuracy ranging from 78% to 98%. Thirteen participants had an overall accuracy above 90%. Having some items at extremely difficult levels is preferred to demonstrate the subjective control on selecting items to which to respond. Because all levels of speech rate were manageable for those participants, the interaction effect of speech rate and reward might have been restricted. Future studies may benefit from using an individualized task demand range. Our results suggested that accuracy and effort allocation may be viewed as two separate dimensions of task performance, but are profitably examined simultaneously when judging how well an individual is likely to perform a task.

Several limitations should be considered when evaluating the current results. First, due to convenience of the available sample, 80% of the participants were females, which may limit the generalizability of the results to the male population. Second, only 10 trials were presented per condition due to the long duration of each trial and many conditions. After rejecting noisy pupillary data and inaccurate trials, a minimum of 6 trials remained in each condition, which was low for psychophysiological measures. More trials should be included in future replications of the study. Third, even though using a common pupillary baseline in all examined window was reasonable because speech comprehension-associated processes may not be completely independent, and our primary interest was auditory comprehension effort rather than listening effort only, the 1 s interval between the response endpoint (button push) and the onset of the next trial was short for slow physiological responses such as pupil dilation. As a result, using the same baseline for the cue window as for the other windows might be problematic. The pupillary responses to the auditory cue seemed to have overlapped with the pupillary recovery phase from the previous trial, therefore, the results of the cue window should be interpreted with caution. Fourth, although we had participants undertake a sequence of practice trials to ensure that they were able to quickly and accurately locate and press the desired button while fixating on the computer screen, the response action can still consume some effort and impact pupil dilation. A non-physical response form might be preferable to indicate the endpoint of the comprehension task in future studies. Last, we did not include the response time in the analysis because our research questions mainly focused on the decisions of effort allocation (indexed by pupil dilation) driven by reward and motivation. Time pressure impacts individuals’ speed-accuracy tradeoff decisions (Swensson, 1972; Liesefeld et al., 2014; Drugowitsch et al., 2015), so the interpretation of response time in the context of our experimental design is complex and deserves a separate and thorough discussion. We have included response time data in the supplementary materials (see supplementary materials Figure 5).

5. Conclusions

Reward and task demand interactively drive auditory comprehension effort allocation and impact auditory comprehension performance accuracy in young normal hearing adults. Given a global task goal that requires strategic effort control, listeners prioritized their effort to relatively easy task items over difficult ones at high levels of reward, suggesting a value-based strategic effort allocation to some extent. Reward significantly improved task performance in terms of accuracy at difficult listening conditions. These findings support the need for additional experimental and clinical study in incorporating reward and value-based strategic effort allocation for evaluating the effects of hearing loss on auditory comprehension. The lack of a linear relationship between effort and performance accuracy revealed complex connection between action and outcome in auditory comprehension. More research on auditory comprehension effort is needed to explore the underlying mechanisms of strategic effort allocation behavior. A thorough understanding of these mechanisms will facilitate the integration of listening and auditory comprehension effort concepts into clinical practice, including auditory capability assessment with effort dynamic range, auditory training centered on efficient effort management, and effort-oriented signal processing amplification technology development.

Financial disclosures

This research was funded by University of Pittsburgh SHRS Research Development Fund.

Declarations of interest

The authors have no conflicts of interest to disclose.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

The authors thank all the volunteers for their participation in this study. The authors gratefully acknowledge Associate Professor Michael Dickey from the Department of Communication Science and Disorders, University of Pittsburgh, for the technical and material support, and his advice on the experiment design. The authors also thank particularly the Distinguished Professor Charles Perfetti from the Department of Psychology, University of Pittsburgh, for his assistance in the design and valued comments on this manuscript. All authors contributed equally to this work. Study concept and design: M.Z, C.P., S.R.P, M.R.M. Study supervision: C.P. Statistical analysis: G.J.S., M.Z. Performing of the experiment and drafting of the manuscript: M.Z. All authors assisted in the interpretation of data and critical revision of the manuscript for important intellectual content. All authors discussed the results and implications, commented on the manuscript at all stages and approved the final manuscript. Approval for this study was obtained through the institutional review board (IRB) of University of Pittsburgh. The study was financially supported by the The School of Health and Rehabilitation Science Research Development Fund of University of Pittsburgh.

Malcolm R. McNeil and Sheila R. Pratt received salary support from research grants funded by the Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Rehabilitation Research and Development Service and the Geriatric Research Education and Clinical Center in the VA Pittsburgh Healthcare System. However, the contents do not represent the views of the Department of Veterans Affairs or the United States Government.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heares.2019.107775.
Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. J. Exp. Psychol. Learn. Mem. Cogn. 41, 1140.

Loewenfeld, L. 1959. The Pupil: Anatomy, Physiology and Clinical Applications. Butterworth Heinemann, Boston.

Macdonald, A.W., Cohen, J.D., Stenger, V.A., Carter, C.S., 2000. Dissociating the role of the dorsolateral prefrontal cortex in cognitive control. Science 288, 1835–1838.

Macdonald, A.W., Cohen, J.D., 2011. Subjective and physiological indices of listening effort in a competing-talker task. J. Am. Acad. Audiol. 22, 113–122.

Mackersie, C.L., Kearney, L. 2017. Autonomic nervous system responses to hearing-related demand and evaluative threat. Am. J. Audiol. 26, 373–377.

Mahr, D., Hortsmann, R., Böllinghoff, K., Wohlgemuth, J., 2015. WM and the effort? Novel insights into obesity-associated alterations in cost-benefit decision-making. Front. Behav. Neurosci. 9.

Mathôt, S., Fabius, J., Van Heusden, E., Van der Stigchel, S., 2018. Simple and sensible preprocessing and baseline correction of pupil-size data. Behav. Res. Methods 50, 94–106.

McCoy, S.L., Tun, P.A., Cox, L.C., Colangelo, M., Stewart, R.A., Wingfield, A., 2005. Hearing loss and perceptual effort: downstream effects on older adults’ memory for speech. Q. J. Exp. Psychol. A 58, 22–33.

Mcgarrigle, R., Daren, S., Stewart, A.J., Kuchinsky, S.E., Munro, K.J., 2017. Measuring listening-related effort and fatigue in school-aged children using pupillometry. J. Exp. Child Psychol. 161, 95.

McNeil, M.R., Pratt, S.R., Szuminsky, N., Sung, J.E., Fossett, T.R.D., Fassbinder, W., Lim, K.Y. 2015. Reliability and validity of the computerized Revised Token Test: comparison of reading and listening versions in persons with and without hearing loss. Ear Hear. 38, 211–219.

Moulens, E., Laroche, J. 1995. Non-parametric techniques for pitch-scale and timescale modification of speech. Speech Commun. 16, 175–205.

Mowen, J.C., Middendorf, R.D., Luther, D. 1981. Joint effects of assigned goal level and incentive structure on task performance: a laboratory study. J. Appl. Psychol. 66, 598–603.

Müller, J.A., Wendt, D., Kollmeier, B., Debener, S., Brand, T. 2019. Effect of speech rate on neural tracking of speech. Front. Psychol. 10, 449.

Nachtgea1, J., Festen, J.M., Kramer, S.E. 2012. Hearing ability in working life and its relationship with sickness leave and self-reported work productivity. Ear Hear. 33, 94–103.

Nachtgea1, J., Kuik, D.J., Anema, J.R., Goverts, S.T., Festen, J.M., Kramer, S.E., 2009. Hearing status, need for recovery after work, and psychosocial work characteris- tics: results from an internet-based national survey on hearing. Int. Audiol. 48, 684–691.

Neijs, K.L.A. 2013. Money Makes the World Go Round: an Integrative Approach to Understanding Proximate and Evolutionary Mechanisms of Money’s Effects. Thesis. University of Amsterdam.

Niimi, M., Nishio, S. 2001. Speaking rate and its components in dysarthric speakers. Neurosci. Lett. 299, 91–94.

Nachtegaal, J., Kuik, D.J., Van der Meer, E., Festen, J.M., Kramer, S.E., 2009. Comparison of reading and listening versions in persons with and without hearing loss. Ear Hear. 30, 311–318.

Nakamura, A., Nakamura, T., 2005. A framework for studying the neurobiology of value-based decision making. Nat. Rev. Neurosci. 9, 545–551.

Neff, G., & Whiteside, M. 2005. Objective measurement of listening effort: effects of background noise and noise reduction. J. Speech Lang. Hear. Res. 52, 1230–1240.

Schluroff, M., Zimmerman, J. E., Jr., Hofmeister, K., Lorscheid, T., Weber, A. 2016. Pupillometry responses to syntactic ambiguity of sentences. Brain Lang. 27, 322–344.

Schneider, M., Leuchs, L., Czisch, M., Samann, P.G., Spooksmear, V.1, 2018. Disen- gagement: Reward anticipation with simultaneous pupillometry/MRI. Neuroimage 178, 11.

Seaman, K.L., Gorkin, M.A., Vekaria, K.M., Hsu, M., Zald, D.H., Samazeur-Larkin, G.R., 2016. Adult age differences in decision making across domains: increased dis- couragement of social and rewards-related psychosocial models. Psychophysiology 31, 737–733.

Shaw, B.J., 2008. Intelligibility of Clear Speech at Normal Rates for Older Adults with Hearing Loss. Annu. Rev. Neurosci. 40.

Siegle, G.J., Ichikawa, N., Steinhaeuser, S., 2008. Blink before and after you think: blinks occur prior to and following cognitive load induced by pupillary re- sponses. Psychophysiology 45, 679–687.

Siegle, G.J., Steinhaeuser, S.R., Friedman, E.S., Thompson, W.S., Thase, M.E., 2011. Remission prognosis for cognitive therapy for recurrent depression using the pupil: utility and neural correlates. Biol. Psychiatry 69, 730–733.

Siegle, G.J., Steinhaeuser, S.R., Mcgarrigle, R., Dawes, P., Stewart, A.J., Kuchinsky, S.E., Munro, K.J., 2017. Toward a rational and mechanistic account of mental effort. Annu. Rev. Neurosci. 40.

Spencer, T.J., Whitwell, M. & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.

Spencer, T.J., Whitwell, M., & Lee, J. 2000 Eye movements and pupil dilation during event perception. Symposium on Eye Tracking Research & Applications. City. pp. 48–48.