

Review

Not peer-reviewed version

Decline of Affective Prosody Recognition with a Positivity Bias among Older Adults: A Systematic Review and Meta-analysis

Xinran Fan , [Enze Tang](#) , [Minyue Zhang](#) , Yi Lin , [Hongwei Ding](#) * , [Yang Zhang](#) *

Posted Date: 21 December 2023

doi: 10.20944/preprints202312.1643.v1

Keywords: affective prosody recognition, older adults, meta-analysis, systematic review



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Decline of Affective Prosody Recognition with a Positivity Bias among Older Adults: A Systematic Review and Meta-Analysis

Xinran Fan ^a, Enze Tang ^b, Minyue Zhang ^a, Yi Lin ^a, Hongwei Ding ^a and Yang Zhang ^c

^a Speech-Language-Hearing Center, School of Foreign Languages, Shanghai Jiao Tong University, Shanghai, China

^b Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China

^c Department of Speech-Language-Hearing Sciences and Masonic Institute for the Developing Brain, University of Minnesota, Minneapolis, USA

First author: Xinran Fan

e-mail address: xr.fan@sjtu.edu.cn; Tel: +86 17320442650

Postal address: School of Foreign Languages, Shanghai Jiao Tong University, 800 Dong Chuan Rd., Minhang District, Shanghai, 200240, China

Corresponding author 1: Hongwei Ding

e-mail address: hwding@sjtu.edu.cn; Tel: +86-213420-5377

Postal address: School of Foreign Languages, Shanghai Jiao Tong University, 800 Dong Chuan Rd., Minhang District, Shanghai, 200240, China

Corresponding author 2: Yang Zhang

e-mail address: zhanglab@umn.edu; Tel: +1-612-624-7818

Postal address: Department of Speech-Language-Hearing Sciences and Masonic Institute for the Developing Brain, University of Minnesota, Minneapolis, MN 55455, USA

Abstract: Purpose: Understanding how older adults perceive and interpret emotional cues in speech prosody contributes to our knowledge of cognitive aging. This study provides a systematic review with meta-analysis to investigate the extent of the decline in affective prosody recognition (APR) among older adults in terms of overall and emotion-specific performance and explore potential moderators that may cause between-study heterogeneity. **Method:** Incorporated into the literature search were six electronic databases, with a specific emphasis on studies comparing the APR performance of older adults with that of younger adults. This comparison was focused on basic emotions. Meta-regression analyses were executed to pinpoint potential moderators related to demographic and methodological characteristics. **Results:** A total of 19 studies were included in the meta-analysis, involving 560 older adults with a mean age of 69.15 years. The findings indicated a substantial negative effect size ($g = -1.21$). Furthermore, the magnitude of aggregated effect sizes showed a distinct valence-related recognition pattern with positive prosody exhibiting smaller effect sizes. Years of education were found to moderate the overall performance effect estimate while age and gender significantly influenced the effect estimate of happiness. **Conclusions:** The results confirmed a significant decline in APR ability among older adults, but this decline was unbalanced across basic emotions. Educational level emerged as a significant factor influencing APR ability. Moreover, participants with a higher mean age exhibited notably poorer performance in recognizing happy prosody. These findings underscore the need to further investigate the neurobiological mechanisms for APR decline associated with aging.

Keywords: affective prosody recognition; older adults; meta-analysis; systematic review

Introduction

Accurate emotion recognition is essential for effective daily communication (Blair, 2003). Previous research findings show that older adults often experience challenges in comprehending and interpreting emotions in speech (Gurera & Isaacowitz, 2019; Ruffman et al., 2008). This specific difficulty substantially contributes to the social and communication hurdles faced by this demographic.

There are multiple verbal and nonverbal channels to convey and interpret emotions, among which affective prosody, or emotions portrayed via voice intonation, is identified as a crucial element in emotion recognition (Banse & Scherer, 1996; Paulmann et al., 2008). It conveys speakers' emotions through various acoustic parameters such as pitch, intensity, and speech rate (Schmidt et al., 2016). Additionally, it holds pivotal significance in the comprehension and analysis of sentences (Sbattella et al., 2014). Current studies on affective prosody recognition (APR) in older adults consistently indicate that they perform significantly less effectively than their younger counterparts (e.g., Cortes et al., 2021; Dupuis & Pichora-Fuller, 2015; Hunter et al., 2010; Maltezou-Papastylianou et al., 2022; Martzoukou et al., 2022; Mitchell, 2007; Ruffman et al., 2009; Schaffer et al., 2009). These age-related disparities begin to manifest as early as an individual's fourth decade (Mill et al., 2009; Paulmann et al., 2008).

1.1. Recognition pattern for older adults in APR across emotion types

While there is general consensus on the notable age-related declines in APR among older adults, it is important to recognize that the extent of decline can vary depending on specific emotions (e.g., Cortes et al., 2021; Dupuis & Pichora-Fuller, 2015; Ruffman et al., 2008; Wong et al., 2005). A meta-analysis by Ruffman et al. (2008) revealed that older adults encountered greater difficulty in recognizing happy, sad, and angry voices, while showing no significant difference from younger individuals in recognizing fear, disgust, and surprise voices. However, subsequent empirical studies have presented divergent findings. For example, in a study by Ruffman et al. (2009), no statistically significant difference was found between older and younger groups in recognizing auditory expressions of happiness, surprise, disgust, sadness, and fear. Yet it was observed that older individuals displayed significantly poorer performance in recognizing angry voices. Additionally, Dupuis and Pichora-Fuller (2015) noted that older adults demonstrated inferior performance across nearly all basic emotions in APR, including happy, sad, angry, fear, disgust, and neutral, except for pleasant surprise, when compared to their younger counterparts. In a recent study, Martzoukou et al. (2022) found that the older group exhibited notably lower accuracy in decoding fear, sadness, anger, neutrality, and surprise, in comparison to younger participants. However, no significant difference was observed between the two groups in their ability to recognize happy voices.

Notably, amid the diverse APR findings in among older adults, some researchers have proposed the presence of a positivity bias, suggesting that older adults may identify positive emotions better than negative ones (Hunter et al., 2010; Martzoukou et al., 2022). This bias aligns with observations in other psychological domains, including attention and memory (Reed et al., 2014).

Several explanations based on motivational and biological mechanisms account for this positivity bias. *Socioemotional selectivity theory* (SST), a prominent motivational theory, posits that individuals alter their goals as motivation evolves with age, leading older adults to focus more on intimate interpersonal relationships and reduce exposure to negative emotions (Carstensen et al., 1999; Carstensen & DeLiema, 2018; Cortes et al., 2021). *Dynamic integration theory* (DIT) offers an alternative motivational perspectives, suggesting that negative information possesses greater complexity, making it more challenging to incorporate into cognitive-affective systems than positive information (Labouvie-Vief et al., 2010; Labouvie-Vief, 2003). A more recent model, the *strength and vulnerability integration* (SAVI) model, asserted that experiential knowledge accumulated over a lifetime enhances an individual's ability to choose appropriate emotion regulation techniques, leading to more evident positive effects when employed before encountering negative stimuli (Charles, 2010; Gurera & Isaacowitz, 2019).

From a biological standpoint, some studies have attributed the lower recognition accuracy of negative emotions, including sadness, to age-related changes in the human brain (Martzoukou et al., 2022; Sowell et al., 2003). The identification of vocal emotions is specifically associated with the orbitofrontal cortex (OFC; Wildgruber et al., 2002, 2005). The recognition of angry, sad, and fearful prosody involves the OFC, the inferior frontal gyrus, and the amygdala (Morris et al., 1999; Sander et al., 2005). Early declines in normal aging, notably in regions like the frontal and medial temporal lobes, have been reported in several studies (Sowell et al., 2003). It is anticipated that older individuals

may encounter challenges in accurately recognizing specific emotions, particularly anger, sadness, and fear (Martzoukou et al., 2022).

However, consensus on the existence of a positivity bias in older adults' APR is currently lacking. Wong (2005) reported that older adults exhibited lower performance in identifying sadness and happiness, but not in recognizing anger, fear, disgust, and surprise. Lima et al. (2014), on the other hand, showed decreased accuracy for older adults across all investigated emotions. Recently, Cortes et al. (2021) observed that older adults encountered similar challenges in identifying both positive and negative emotions. They proposed that the previously observed positivity effect might be attributed to a ceiling effect, particularly when happiness was the sole positive emotion considered. This effect diminished when multiple positive emotions and response options were available.

1.2. Why an updated meta-analytic review is needed

There has been a notable surge in research on auditory emotion recognition in aging subsequent to Ruffman et al.'s (2008) meta-analysis. However, many studies have been unable to replicate the age-related patterns observed by Ruffman et al. (2008). These divergent findings and the lack of consensus regarding the positivity bias highlight the complexity and intricacy of the research topic and the need of quantitative evidence with a meta-analysis approach to guide further research to advance our understanding of age-related decline of APR. The previous meta-analysis was conducted 15 years ago to examine older adults' emotion recognition in three modalities (Ruffman et al., 2008). There is currently a gap in the literature focusing on the auditory domain. A recent systematic review has summarized various studies on this topic (Baglione et al., 2023), but it fell short of providing a meta-analysis to statistically aggregate the results to obtain a more precise estimate and in-depth analysis of the moderating factors (X. Zhang et al., 2022). To our best knowledge, this meta-analysis represents the first with a focus on emotion perception in the auditory domain specifically in older adults, aiming to provide updated quantitative evidence for APR research in this population.

What sets our study apart from the previous meta-analysis (Ruffman et al., 2008) lies in a detailed examination of the moderating effects of demographic characteristics and methodological aspects to help explain the convergent or divergent findings. By considering these moderators, our new meta-analysis could contribute to a more nuanced understanding of the research findings and illuminate the influence of these potential factors to guide future studies.

1.3. Objectives and hypotheses of the current study

The primary objective of this study is to evaluate the extent of the decline in APR among older adults and to ascertain if the age-related APR pattern reported by Ruffman et al. (2008) persists considering recent research. To concentrate on assessing older adults' APR ability in the auditory domain, we exclusively focused on research explicitly investigating age-related effects on APR. Studies involving multimodal tasks or rating tasks assessing valence or arousal were excluded. We hypothesize that older adults would show lower APR than younger adults, consistent with most studies. The focus is on whether the age effect across emotion types aligns with the previous meta-analysis that reported no evidence of a positivity bias (Ruffman et al., 2008). If age-related challenges in APR are equally observed across basic emotion types, it would imply a general age-related cognitive decline. If older individuals are better at identifying some emotions than others, motivational models such as the SST, DIT, or SAVI, or biological mechanisms could provide plausible accounts.

The second objective is to systematically investigate the potential factors or moderators impacting study outcomes, including demographic characteristics and methodological factors. Although it is difficult to predict the exact effects, it is conceivable that demographic characteristics, including mean age (Amorim et al., 2021), gender distribution (Lambrecht et al., 2014), and years of education (Demenescu et al., 2014), as well as methodological aspects in task design, material type and task difficulties (M. Zhang et al., 2022), may contribute to result heterogeneity.

2. Methods

2.1. Eligibility criteria

The studies were considered eligible if they were published in English and adopted an experimental method comparing APR between older adults (mean age above 60) and younger adults (mean age under 35). Participants with psychiatric disorders (e.g., schizophrenia, bipolar disorder) or physical disabilities (e.g., blindness, hearing loss, brain damage) were excluded. Studies had to involve at least one of the six basic emotions defined by Ekman (1992) (*happiness, sadness, anger, disgust, fear, and surprise*) in the auditory modality. Studies that focused on complex or social emotions (e.g., sarcasm, amusement, embarrassment, and contempt) were excluded. Considering the limited number of existing studies, stimuli could be pure prosody without semantic content or prosody with verbal semantics. Moderator tests indicated that the semantic information did not affect the meta-analysis results.

Eligible studies should be primary empirical research with original data published in peer-reviewed journals. To be included, studies needed to contain sufficient statistical information to calculate an effect size of the difference in APR performance between older and younger adults. Studies lacking sufficient data for effect size calculation were excluded unless additional data could be obtained from the author(s). Review articles, editorials, and meta-analyses were excluded. Neuroimaging and electroencephalographic studies were also eligible as long as they reported sufficient behavioral data. Studies that exclusively focused on emotional memory, emotional production, emotion processing, or emotion perception in other modalities (e.g., visual modality) were excluded.

2.2. Search strategy

Six electronic databases were searched for comprehensive research: Web of Science Core Collection, PubMed, APA PsycInfo, APA PsycArticles, and Academic Search Complete. The advanced search keywords were “emotional prosody OR affective prosody” AND “older adults OR older people OR aging OR aged OR elderly OR seniors OR geriatrics”. There are no restrictions on publication dates, and the search was completed in August 2023. Furthermore, to identify additional potential studies, articles derived from the references of the selected ones and other review articles were also included.

2.3. Study selection

Two researchers (XF, ET) independently examined titles and abstracts to identify articles that needed to be reviewed in full text. Articles meeting the specified criteria and reporting sufficient data for meta-analysis were included in the quantitative analysis after full-text examination. Studies meeting other criteria but lacking adequate statistical information for effect size calculation were subjected to qualitative analysis. In cases of disagreement between the two researchers regarding article eligibility, consensus was sought through re-evaluation and discussion.

2.4 Data extraction

Two authors (XF, ET) extracted the following key elements from each included study: (1) study characteristics (e.g., publication year, author information, and region), (2) participant-related information (e.g., sample size, mean age, gender ratio, and years of education), (3) methodology-related information (e.g., materials, number of speakers, and number of trials), (4) major results that were necessary for meta-analysis of overall performance and emotion-specific meta-analysis (e.g., means and standard deviations, *t*-test values, and *F* values).

2.5. Outcome measures

Accuracy data of APR in performance-based measures such as identification and discrimination were extracted from studies to reveal differences between older adults and younger adults. If the

results were reported as the number of trials corrected or errored, or as percentages of errors, the corresponding percent correct data were manually calculated. Studies that did not summarize overall accuracy but only reported data on each emotion were excluded from overall performance estimates but synthesized in emotion-specific meta-analysis.

2.6. Moderator variables

Moderators were chosen based on their potential to impact the observed effect size estimate. The following factors were considered: (1) demographic characteristics: mean age, gender distribution, and years of education; (2) methodology-related variables: material type (stimuli with pure prosody (SPP), stimuli with neutral semantic content (SNSC), stimuli with congruent meaning and prosody (SCMP), or stimuli with incongruent meaning and prosody (SIMP)), task difficulty (including the number of speakers, number of trials, and number of answer options), emotional valence of positivity (only for the emotion of surprise, since some studies employed pleasant surprise while some used negative surprise). However, material type was not included in the moderator analysis of emotion-specific performance due to an insufficient number of studies that specified material type.

2.7. Statistical analysis

Meta-analyses were conducted using software R (version 4.2.1). This study adopted Hedges' g for effect size estimates, which uses a correction factor for small sample bias (X. Zhang et al., 2022). The magnitude of Hedges' g was interpreted using benchmarks proposed by Cohen (1988): 0.2 for small, 0.5 for medium, and 0.8 for large. However, these benchmarks have been challenged by researchers who argued that the magnitude of effect sizes should be compared against the related prior studies (Hemphill, 2003; Thompson, 2007). Therefore, we also interpreted effect sizes with benchmarks proposed by Gaeta and Brydges (2020) for audiology and speech-language pathology research: 0.25 for small, 0.55 for medium, and 0.95 for large. In addition to employing statistical benchmarks to determine the magnitude of effect sizes, we also utilize the effect size of neutral as a baseline for comparing the age-related decline across different emotions.

Considering that demographic characteristics and experimental design varied across the included studies, we completed meta-analyses with the random-effects model (DerSimonian-Laird estimate) to estimate effect sizes. A prior power analysis indicated that the number of included studies was sufficient to detect moderate to large effects in the meta-analysis, with the statistical power above 95%.

The Cochran's Q statistic (Borenstein, 2009), I^2 statistic (Higgins & Thompson, 2002), and prediction intervals (IntHout et al., 2016) were adopted to assess between-study heterogeneity and the impact of heterogeneity. If the result of Cochran's Q -test is significant, it indicates that the estimated effect sizes are widely dispersed. I^2 represents the ratio of true heterogeneity in the effect sizes that is not caused by sampling error, with 25%, 50%, and 75% being considered as low, moderate, and high degrees of heterogeneity, respectively (Higgins et al., 2003). If the value of I^2 is 0, it suggests that these studies are completely homogeneous, and the random-effects model is simplified to a fixed-effect model. The prediction intervals provide an expected effect range that future similar studies may fall into based on the observed evidence.

Outliers were defined as those with effect sizes so significantly different from the overall effect that their confidence intervals did not overlap with the confidence interval of the pooled effect (Harrer et al., 2021). We used the Leave-One-Out method to make influential analyses to ensure that the pooled effect size estimate was not influenced by outlier cases. This method systematically assesses the impact of individual studies by iteratively removing one study at a time.

Publication bias was evaluated by funnel plots and by Egger's test to detect asymmetry patterns (Egger et al., 1997). A trim-and-fill method was employed to calculate a bias-corrected estimate (Duval & Tweedie, 2000) if Egger's test indicated publication bias ($p < 0.05$). The bias-corrected effect size estimate and its 95% CI were reported after detecting outliers and calculating the adjusted effect.

To identify sources of between-study heterogeneity, we employed a mixed-effects model for meta-regression analyses to examine the contributions of various potential moderators. Moderators

were considered if information was available from a sufficient number of studies ($N \geq 4$; Tang et al., 2022; Velikonja et al., 2019). The regression coefficients, the Q_{model} (Q_M) statistics, and the p -values were reported respectively. To quantify the amount of heterogeneity explained by each moderator, R^2 was calculated to indicate the strength of their relationship with the estimated effect (Borenstein, 2009).

2.8. Quality assessment

Two reviewers (XF, ET) independently assessed the quality of the included quantitative studies using the 14-item standard quality assessment (SQA) criteria (Kmet et al., 2004). Given that all the studies were quantitative behavioral experiments, the SQA criteria were deemed appropriate for quality evaluation. However, as the relevant studies did not involve interventions, ratings for item 5 (If interventional and random allocation was possible, was it described?), item 6 (If interventional and blinding of investigators to intervention was possible, was it reported?), and item 7 (If interventional and blinding of subjects to intervention was possible, was it reported?) were excluded. The two reviewers conducted their assessments individually and then reached a consensus through discussion.

3. Results

3.1. Literature search results

The electronic search yielded a total of 802 articles. Titles and abstracts of these studies were checked according to the inclusion and exclusion criteria, resulting in 54 articles remained for further scrutiny. Two reviewers (XF, ET) independently examined the full texts of these studies and identified 21 eligible articles for qualitative analysis. An additional five studies were discovered from the reference lists, bringing the total eligible articles to 26. Among these, ten articles lacked sufficient data for effect size estimates and were excluded from quantitative synthesis. The remaining 16 articles, constituting 19 studies/experiments, were included in the meta-analysis (13 studies for the overall APR effect size calculation, and 14 studies for the emotion-specific APR meta-analysis). Details of the literature screening are presented in Figure 1. Articles included by one review but excluded by the other were re-evaluated to reach a consensus after thorough discussion.

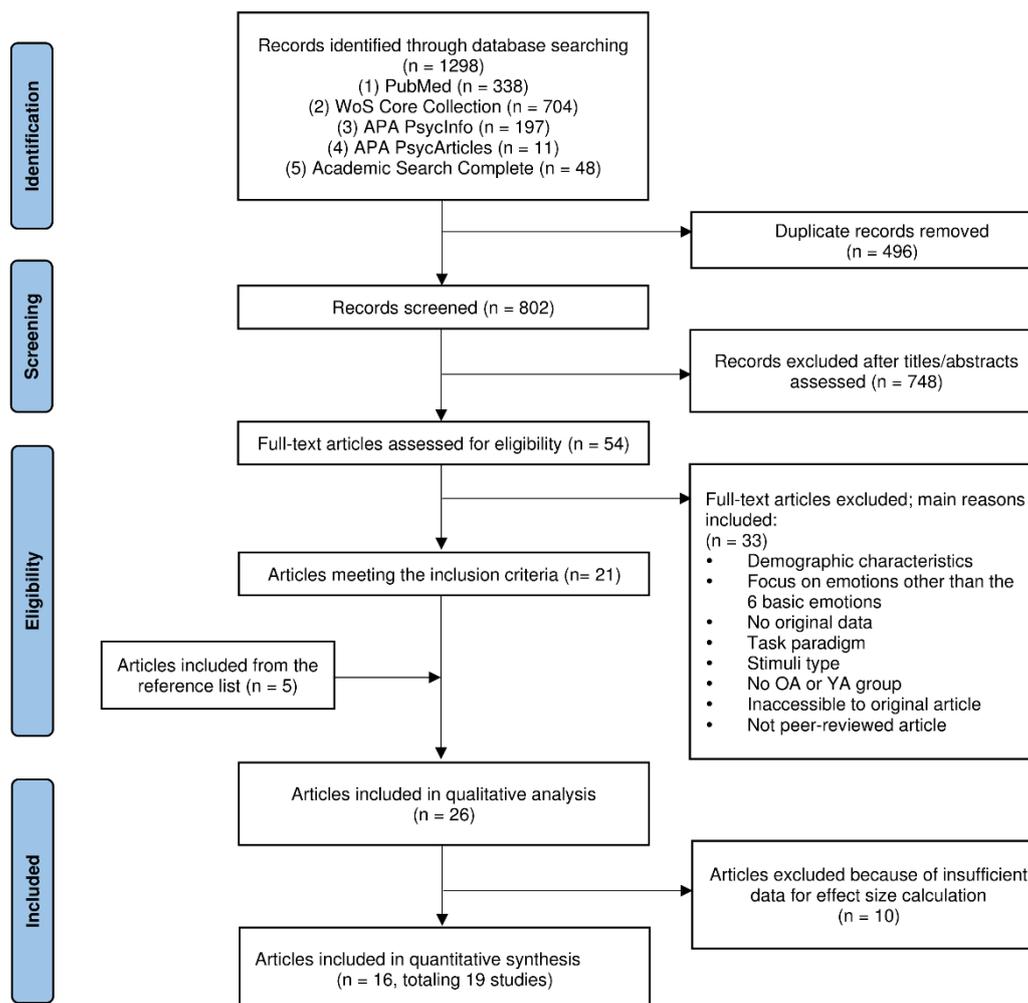


Figure 1. Flow chart for the process of the systematic review and meta-analysis.

3.2. Study characteristics and quality assessment

The 19 included studies spanned from the years 1995 to 2022 and comprised a combined sample of 751 younger adults (Mean age 23.02 across studies) and 560 older adults (Mean age 69.15 across studies). Six studies (Brosigle & Weisman, 1995; Chaby et al., 2015; Cortes et al., 2021; Lima & Castro, 2011; Seddoh et al., 2020; Wong et al., 2005) were excluded from the meta-analysis of the overall performance of APR due to insufficient data for overall accuracy calculation.

Sample sizes varied across the studies, with the younger group ranging from 10 to 155 participants and the older group ranging from 10 to 61 participants, with an average of 39.53 younger adults and 31.11 older adults in one study. The percentage of male participants ranged from 20.00% to 50.00%. Eleven studies reported the education years of the participants, and the mean education years were 13.85 for the younger group and 14.82 for the elderly group. There was no significant difference in education level between the younger and older adults ($t = -0.64, p = 0.53$).

General characteristics, statistical measures, and quality assessment of included studies for quantitative analysis were summarized in Supplemental Sheet 1. The inter-rater correlation coefficient between the two reviewers, calculated using Spearman's correlation (Gwet, 2014), was 0.86. Any disagreements were resolved through follow-up discussions and reached consensus. Overall, the articles scored medium to high in the SQA (Todorova et al., 2019; M. Zhang et al., 2022; Mean = 0.87; SD = 0.09, min-max = [0.68–1.00]).

The majority of the studies included in the analysis were conducted in Europe, accounting for ten studies (52.63%). Following Europe, seven studies (36.84%) were conducted in North America, and two studies (10.53%) were in Oceania. English was the predominant language in 14 studies, while

Swedish was used in two studies, and Greek, French, and Portuguese were each used in one study. In terms of publication timeline, seven studies (36.84%) were published before 2010, and 12 studies (63.16%) were published after 2010, indicating a more recent focus on the topic.

All included studies employed at least one identification task, with two studies incorporating both identification and discrimination tasks (Schaffer et al., 2009). Among studies that reported material type, the majority utilized one type ($N = 14$), while two employed a combination of two types, and two used three types. In the included studies, SNSC was the most utilized material type ($N = 8$), following SPP ($N = 4$). Four studies employed a combination of materials: two used SPP and SNSC, and the other two used a combination of SNSC, SIMP, and SCMP. Stimulus forms included sentences ($N = 12$), nonverbal vocalizations ($N = 3$), monophthongs ($N = 2$), and a combination of nonverbal vocalizations and sentences ($N = 1$). All studies provided participants with answer choices rather than requiring labeling.

3.3. The overall accuracy of older adults on APR

Supplemental Sheet 2 presents a summary of results for the overall APR accuracy of older adults compared with that of younger participants in each included study. Hedges' g ranged from -2.68 to -0.51 . According to Cohen's benchmarks, three studies achieved the medium to large magnitude effect size, and ten studies reached the large benchmark. Applying more stringent benchmarks in the clinical audiology field, one study reported a small to medium magnitude, three studies demonstrated medium to large benchmark, and nine reached the large magnitude benchmark.

The standardized mean effect size was large and significant (Hedges' $g = -1.21$, 95% CI -1.50 to -0.92 , $p < 0.01$; Figure 2.). There was significant heterogeneity between studies ($Q(12) = 35.68$, $p < 0.01$), indicating variability in effect sizes among the included studies. This was further supported by the moderate to high degrees of true heterogeneity between the included studies ($I^2 = 66.4\%$), independent of heterogeneity caused by sampling error. The prediction interval $[-2.21, -0.21]$ predicted that future studies would all report APR impairments in older adults. Out of these studies, one outlier (Martzoukou et al., 2022) was identified through examination of the confidence intervals (CIs). After removing the outlier, the standardized mean effect size remained large and significant (Hedges' $g = -1.09$, 95% CI -1.31 to -0.86 , $p < 0.01$). Influential analyses showed no significant change in results, with effect sizes ranging from -1.09 (95% CI -1.31 to -0.86) to -1.27 (95% CI -1.55 to -0.98), suggesting that no study unduly influenced the overall findings (Supplemental Figure 1.).

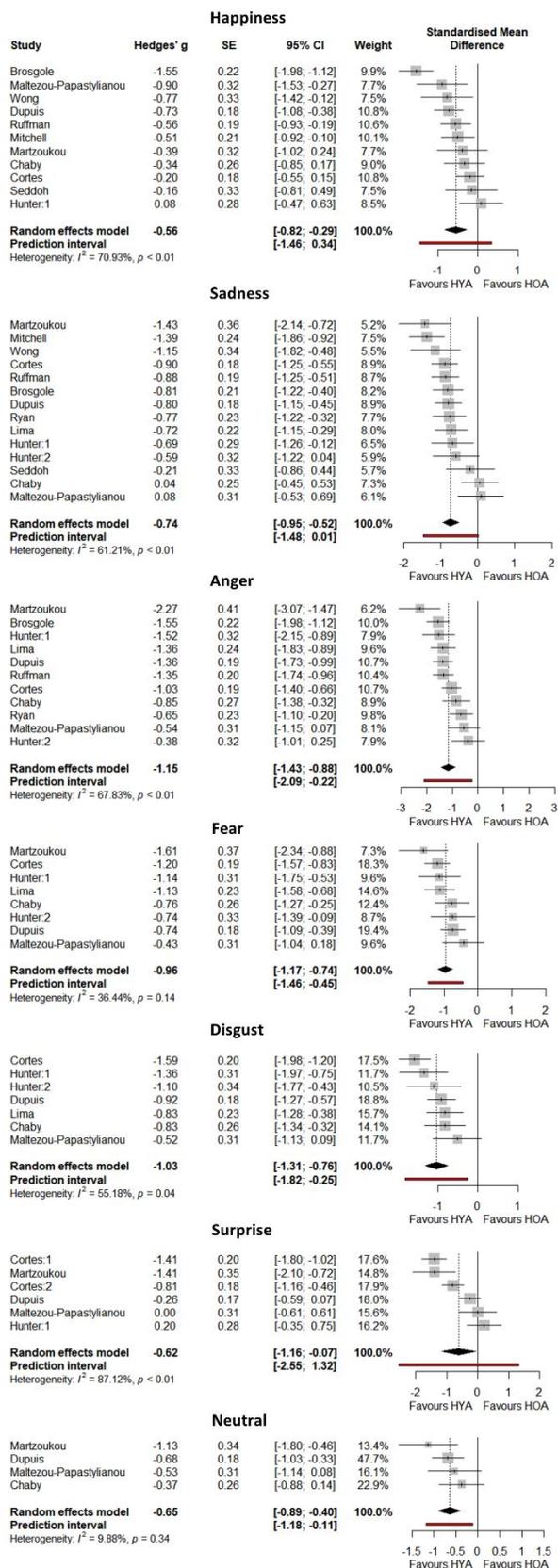


Figure 2. Forest plot for overall performance, with individual and pooled effect sizes (Hedges' g), standard errors, 95% confidence intervals and weights for the included studies.

For publication bias assessment, a contour-enhanced funnel plot (Peters et al., 2008) was generated to evaluate the relationship between effect dispersion and statistical significance, which detected asymmetry (Supplemental Figure 2.). Egger's test confirmed significant publication bias ($p < 0.01$). To adjust the pooled effect size by imputing potentially missing publications, the trim-and-fill analysis was conducted twice, with all studies included or the outlier removed (Supplemental Figure 3.). When all studies were considered, five studies were added, leading to an adjusted smaller but still significant estimate (Hedges' $g = -0.93$, 95% CI -1.28 to -0.59 , $p < 0.01$, prediction interval -2.37 to 0.50). In the other calculation with the outlier removed, four studies were added, producing a similar effect estimate (Hedges' $g = -0.99$, 95% CI -1.36 to -0.61 , $p < 0.01$, prediction interval -2.50 to 0.53).

3.4. Emotion-specific accuracy of older adults on APR

In separate random-effects meta-analyses examining performance for each emotion, older adults, compared to younger participants, demonstrated significantly lower accuracy in recognizing happiness (Hedges' $g = -0.56$, 95% CI -0.82 to -0.29 , $p < 0.01$), sadness (Hedges' $g = -0.74$, 95% CI -0.95 to -0.52 , $p < 0.001$), anger (Hedges' $g = -1.15$, 95% CI -1.43 to -0.88 , $p < 0.001$), fear (Hedges' $g = -0.96$, 95% CI -1.17 to -0.74 , $p < 0.001$), disgust (Hedges' $g = -1.03$, 95% CI -1.31 to -0.76 , $p < 0.001$), surprise (Hedges' $g = -0.62$, 95% CI -1.16 to -0.07 , $p = 0.03$), and neutral (Hedges' $g = -0.65$, 95% CI -0.89 to -0.40 , $p < 0.001$). Based on Cohen's benchmark, effect sizes of happiness, sadness, surprise, and neutral demonstrated the medium to large magnitude, and effect sizes of anger, fear, and disgust reached the large benchmark. Similarly, according to the more rigorous benchmarks within the field of clinical audiology, effect sizes of happiness, sadness, surprise, and neutral fell within the medium to large range, while effect sizes of anger, fear, and disgust met the benchmark of large magnitude.

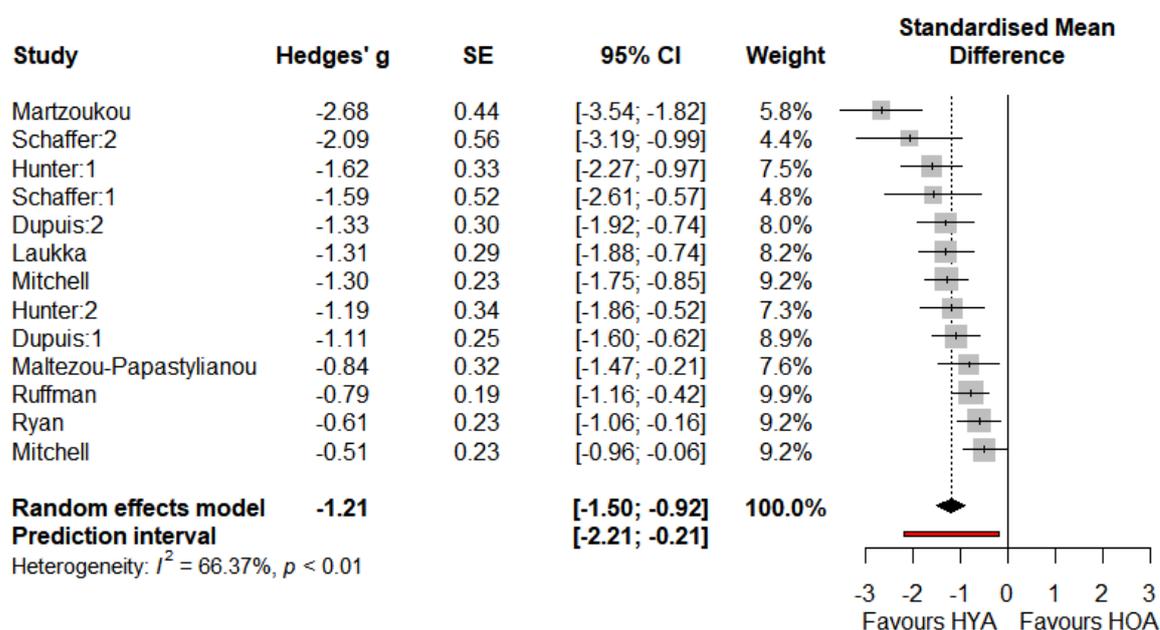


Figure 3. Forest plot for emotion-specific performance, with individual and pooled effect sizes (Hedges' g), standard errors, 95% confidence intervals and weights for the included studies.

3.5. Moderators for overall and emotion-specific APR performance of older adults

Meta-regression analyses (refer to Supplemental Sheet 3 for details) were conducted to examine moderators influencing the heterogeneity of overall and emotion-specific APR effect sizes. Across eight analyses encompassing biological and study-related variables, only years of education ($Q_M(1) = 8.34$, $R^2 = 100.00\%$, $p = 0.004$) significantly accounted for the heterogeneity of true effect sizes in overall

APR performance. Effect size differences were not significantly associated with older adults' mean age ($Q_M(1) = 1.69, p = 0.19$), male percentage ($Q_M(1) = 3.74, p = 0.053$), material ($Q_M(3) = 1.86, p = 0.60$), number of speakers ($Q_M(1) = 0.18, p = 0.67$), number of trials ($Q_M(1) = 0.001, p = 0.97$), and number of answer options ($Q_M(1) = 0.40, p = 0.53$).

Regarding specific emotions, different moderators were associated with the effect size estimates. For happiness, older adults' mean age ($Q_M(1) = 13.97, R^2 = 85.91\%, p < 0.01$) and male percentage of older adults ($Q_M(1) = 9.22, R^2 = 62.79\%, p < 0.01$) emerged as significant factors that could mostly explain the heterogeneity of true effect sizes. For other emotions, none of the moderators could account for the heterogeneity of true effect sizes (see Supplemental Sheet 3 for detailed information).

4. Discussion

The present meta-analysis aimed to achieve two primary objectives. The first objective was to assess the extent of age-related decline in APR and to investigate whether the emotion-specific pattern observed by Ruffman et al. (2008) remains consistent with more recent research findings. The second objective was to explore potential contributions of demographic characteristics and methodological factors to the heterogeneity of results.

Regarding the first objective, this meta-analysis significantly built upon and expanded the prior work of Ruffman et al. (2008) (which included findings from only five studies) by incorporating findings from a much larger number of studies ($N = 19$; for overall performance estimate, $N = 13$; for emotion-specific performance analysis, $N = 14$, and numbers differed across different emotions). Our meta-analysis results revealed that (i) older adults did exhibit significantly lower accuracy in overall APR, (ii) the emotion-specific pattern observed by Ruffman et al. (2008) was not consistently replicated, and (iii) older adults demonstrated a positivity bias in recognizing affective prosody. In the emotion-specific analysis, fear, disgust, anger, and sadness exhibited greater declines in older adults compared to happiness and surprise, with neutral serving as the baseline.

Aligning with the second objective, the systematic investigation of participants' demographic features and methodological factors showed that (i) participants' years of education contributed to the heterogeneity of the overall performance results, and (ii) the heterogeneity in the results of happiness was influenced by the male percentage and mean age of older adults. These findings provide valuable insights for future research directions in the study of APR in the elderly population as discussed in the following sections.

4.1. Evidence for APR decline in older adults (Aim 1a)

The overall performance estimate, based on a comprehensive analysis of 13 studies encompassing 378 older adults, revealed a very large effect size (-1.21), indicating significant challenges in APR among the elderly. This aligns with the conclusions of the previous meta-analysis (Ruffman et al., 2008) and the recent systematic review (Baglione et al., 2023), both highlighting the APR difficulties experienced by older adults. Importantly, the effect size estimate of the meta-analysis remained largely unchanged even after correcting for publication bias (Hedges' $g = -0.93$ or -0.99), with the 95% CI still indicative of the medium magnitude. This suggests that publication bias did not significantly impact the overall effect size estimate, reinforcing the notion that older adults indeed exhibit impaired APR abilities compared to their younger counterparts.

Various factors have been proposed to explain the APR decline in older adults, including age-related changes in cognitive functions, hearing loss, and alterations in brain mechanisms. The subsequent analysis will delve into a comprehensive examination and discussion of the impact of these three distinct perspectives on older adults' APR.

Roles of cognitive functions

The connection between age-related cognitive decline and the identification of emotions has been a subject of investigation (Mitchell, 2007). Most reviewed articles considered participants' general cognitive status before experiments, employing various assessments, including standardized

tests such as Mini-Mental State Examination (MMSE; Folstein et al., 1975), Montreal Cognitive Assessment test (MoCA; Nasreddine et al., 2005), Wechsler Adult Intelligence Scales-III (WAIS-III; Wechsler, 1997), and non-standardized tests such as short interview and self-reports (see Supplemental Sheet 1 for more detailed information).

Despite efforts to link older adults' impaired APR to general cognitive decline, studies, such as the one by Orbelo et al. (2005), concluded that standard neuropsychological tests may not fully explain impairments in APR for healthy older adults. The recognition of emotions encompasses a variety of cognitive abilities, including processing speed, intelligence, and working memory. Researchers have explored potential cognitive mediators. For example, Mitchell (2007) found that only frontal lobe load and verbal IQ were linked to the decline in older adults' APR after controlling for hearing loss and some cognitive features. They thus postulated that while some cognitive features may exaggerate age-related difficulties, APR challenges in older adults were, to some extent, primary in origin.

Due to the diverse measurement methods for assessing mental status across different studies, it is not possible to quantitatively analyze the mental status of the participants. To investigate whether cognitively normal participants also exhibited age-related decline in APR, we conducted an additional meta-analysis focusing on studies that employed standardized tests to screen the general cognitive ability of older participants. By synthesizing the available data, we again obtained a large and statistically significant effect size (Hedges' $g = -1.87$, 95% CI -2.52 to -1.22 , $p < 0.01$), suggesting a notable decline in APR ability even among older adults with normal general cognitive ability. Therefore, it is crucial to acknowledge that the age-related deterioration of cognition may not entirely explain the discrepancies reported in the capacity to perceive and interpret affective prosody. However, since the number of studies included in this additional analysis was relatively small ($N = 4$), further inclusion of more studies is needed to strengthen the statistical power for a robust conclusion. Moreover, the variation in age-related decline across different emotions, as seen in emotion-specific performance, implies that cognitive abilities may not entirely account for the observed APR decline among older participants.

4.1.2. Roles of hearing sensitivity

Some researchers have postulated that age-related hearing loss could be a factor influencing APR in older adults (Lambrecht et al., 2012; Picou, 2016). Picou (2016) examined the differences in APR between older adults with varying degrees of hearing loss and middle-aged adults with normal hearing. Results indicated that hearing loss, rather than age, was the primary factor influencing emotional valence rating. Lambrecht et al. (2012) found a mediation effect of hearing loss at the frequency of 4000 Hz on APR, proposing that interpreting nonverbal affective prosody may require auditory signals from the 4000 Hz band.

Contrastingly, other studies reached different conclusions. Conventional audiometric pure-tone thresholds did not correlate with APR difficulties in individuals with age-related hearing loss, especially those with mild to moderate hearing loss (Mitchell, 2007; Orbelo et al., 2005). According to Orbelo et al. (2005), older adults with mild to moderate hearing impairment exhibited APR impairments, but the variability within older adults was associated with aging effects on the right hemisphere of the brain, not the degree of hearing loss. In a study by Dupuis (2015) focusing on normal-hearing older adults, acoustical features (vowel F0 difference limens, gap detection, and intensity difference limens) did not significantly impact APR, suggesting that hearing sensitivity might not be a prominent determinant of APR in older adults with normal hearing.

The heterogeneity in assessing hearing sensitivity, mirroring the varied measurement of cognitive ability, precluded a quantitative analysis. Nonetheless, an additional meta-analysis focusing on studies rigorously screening participants' hearing sensitivity showed an aggregated effect size of -1.05 (95% CI -1.43 to -0.66 , $p < 0.01$), still indicating a large and significant decline in APR. While limited by a small number of studies ($N = 4$), this underscores the need for more research to enhance generalizability. Importantly, since all reviewed studies included participants with

normal hearing, age-related hearing decline may not entirely explain observed APR challenges in older adults.

4.1.3. Roles of age-related changes in neural mechanisms

The age-related decline in APR can be linked to the changes in brain mechanisms involved in emotion processing. Emotion processing, whether visual or auditory, relies on specific brain regions and networks like the frontal and temporal lobes, which undergo significant declines in functional activity with age (Raz et al., 2005; Ruffman et al., 2008). Notably, age-related alterations in these brain regions may contribute to the challenges in APR for older individuals. While cognitive functions and hearing sensitivity alone cannot fully explain the age-related decline in APR, we also consider that aging brain mechanisms play a role in this decline. This perspective aligns with the findings of a previous meta-analysis (Ruffman et al., 2008), which suggested that changes in emotion recognition with age might be attributed to alterations in the brain.

However, the neuroanatomical mechanisms underlying APR deficits have received limited attention (Baglione et al., 2023). Existing literature has predominantly focused on behavioral and cognitive aspects of APR decline among older individuals, with few studies directly exploring the neural substrates. Among the included studies, only one investigated the relationship between neuro-stimulation targeting the inferior frontal cortex and APR in older adults (Maltezos-Papastilianou et al., 2022). Nevertheless, their results indicated that neither transcranial direct current stimulation (tDCS) nor high-frequency transcranial random noise stimulation (tRNS) facilitated APR in the elderly group. Further research is needed to delve into the precise nature of age-related alterations and their direct impact on APR in older adults.

4.2. Patterns of emotion-specific APR performance in older adults (Aim 1b)

This study conducted meta-analyses to assess the effect sizes of specific emotions. The included studies predominantly reported results for sadness ($N = 14$), happiness ($N = 11$), and anger ($N = 11$). About one-third of the studies provided results for fear ($N = 8$), disgust ($N = 7$), and surprise ($N = 6$).

4.2.1. Comparing findings with previous meta-analysis: Divergence and convergence

In comparison to Ruffman et al.'s (2008), our study yielded some convergent findings. Similar to their observations that older adults performed worse than younger adults in recognizing happy, sad, and angry prosody, our study also revealed challenges for the elderly group in these affective prosodies (Hedges' $g = -0.56$ for happiness, -0.74 for sadness, and -1.15 for anger).

Contrary to Ruffman et al.'s (2008) findings that older adults performed similarly to younger adults in identifying fearful, disgusted, and surprised prosody, our results indicated challenges for older adults in these affective prosodies as well (Hedges' $g = -0.96$ for fear, -1.03 for disgust, -0.62 for surprise). Thus, our findings did not support the notion that recognition abilities for the other three affective prosodies were preserved in older adults. Instead, we suggested that all affective prosodies exhibited varying degrees of decline, potentially due to the inclusion of a greater number of studies in the current analysis.

Our results indicated that the elderly group has difficulties in APR across all six basic emotions. The largest effect size was observed in anger (-1.15 , favoring the younger group), with happiness showing the smallest effect size (-0.56 , favoring the younger group). Considering the observed decline in all emotions, we chose the effect size of neutral prosody (Hedges' $g = -0.65$) as the baseline to assess the extent of the decline in basic emotions. Compared to the neutral prosody, older adults exhibited greater difficulties in recognizing angry, fearful, disgusted, and sad prosody, while showing relatively preserved recognition abilities for happy and surprised prosody. Notably, surprise was not an inherently positive emotion since it can also convey negative emotion in an unpleasant manner (Hunter et al., 2010; Ruffman et al., 2008), but most studies included in this meta-synthesis employed pleasant surprise (4 out of 6). When we specifically examined positive surprise,

the effect size diminished (Hedges' $g = -0.39$), indicating an even stronger recognition ability for positive surprise prosody among older adults.

4.2.2. Positivity bias in older adults

One important finding of our meta-analysis is that the effect sizes for positive prosody were lower than those for negative prosody. This is in contrast to the prior meta-analysis (Ruffman et al., 2008) that did not show a positivity bias in the auditory domain of emotion recognition. It is important to note that the positivity bias in older adults did not imply superior or equivalent performance compared to younger individuals. Instead, it suggested a relative preservation of recognition abilities for positive prosody compared to neutral stimuli, albeit still inferior to the younger adults.

Some previous studies suggested that the elderly group exhibits an asymmetric APR pattern in favor of positive prosody (Carstensen & Mikels, 2005; Hunter et al., 2010). There are several potential explanations. Firstly, limited future time views may motivate older adults to prioritize positive social interactions and emotion control skills for improved emotional well-being (SST; Carstensen et al., 1999; Carstensen & DeLiema, 2018). Secondly, the complexity of negative information might make it more challenging for older adults to integrate negative emotions into cognitive-affective processes than positive information (DIT; Labouvie-Vief et al., 2010; Labouvie-Vief, 2003). Thirdly, older adults might enhance their emotion perception accuracy when more familiar with experiential knowledge of positive emotions (SAIV; Charles, 2010). Lastly, age-related decline in certain brain regions associated with negative emotions might contribute to this positivity bias (Martzoukou et al., 2022; Sowell et al., 2003). However, since our study did not directly investigate these theories or models, we refrained from delving into further discussion on these explanations in this context.

It is important to note that although we found superior performance for positive prosody compared to negative prosody among the elderly group, it does not imply its presence in all scenarios. Within the six basic emotions, the number of positive emotions was limited, with happiness being the only strictly positive emotion. Even with the inclusion of surprise, there were only two positive emotions represented. Therefore, in most studies, participants might be more inclined to employ strategies that categorize all positive emotions as happiness or surprise, since there were no other competing options to identify, leading to higher accuracy rates in the recognition of positive emotions (Cortes et al., 2021). Future research can incorporate the balance of valence as a consideration and further investigate how valence influences the age-related APR performance.

4.3. Moderator variable effects on APR in older adults (Aim 2)

To elucidate the heterogeneity observed between studies, we conducted meta-regression analyses of demographic characteristics and methodology-related factors for all meta-analyses. The results of meta-regression analyses provided insights into the association between the moderator variables and the effect estimate.

4.3.1. Participant-related moderators

For the overall performance effect estimate, years of education were investigated as a potential moderator for heterogeneity ($p = 0.004$) across studies. Most studies included in the review provided information on participants' education, revealing a range from 10.4 to 16.3 years for older adults. The findings indicated that studies with participants having fewer years of education showed larger effect sizes, suggesting that older adults with higher educational levels exhibit smaller differences in APR compared to younger adults.

While higher education is typically associated with improved cognitive function and reduced dementia risk (Lövdén et al., 2020), it has been debated whether cognitive abilities directly impact APR in healthy elderly individuals (Mitchell, 2007; Orbelo et al., 2005). The present study suggested that education level may exert a direct influence on APR ability in later life, potentially independent of cognitive mediation. This aligns with Demenescu et al.'s (2014) findings, revealing a significant

correlation between higher education and better recognition of fearful and happy voices. However, in our study, this impact was observed at a general recognition level rather than being specific to individual emotions. While the existing research indicates a connection between education and APR ability in older adults, it lacks conclusive data, necessitating further exploration.

In addition, our meta-analysis also examined gender distribution as a moderator. Results showed a significant influence of gender distribution on between-study heterogeneity for happiness ($p = 0.002$). Notably, a higher percentage of males in the elderly group was linked to a reduced performance gap in identifying happy prosody, indicating that older males might excel in recognizing happy prosody compared to older females. However, Sen et al. (2018) discovered that older females exclusively outperformed older males in happy vocal emotion, indicating that the female advantage in older adults' APR performance was emotion-specific. Additionally, an overall trending association ($p = 0.053$) between gender distribution and effect size variation in overall APR performance was observed. However, this time, the effect was inverse. Elderly groups with a higher proportion of males exhibited a larger performance discrepancy, suggesting that older male adults performed worse than older females in overall APR tasks. This aligns with existing literature demonstrating a female advantage in overall APR performance in older adults (Keshtiari & Kuhlmann, 2016; Lausen & Schacht, 2018; Paulmann & Uskul, 2014; Scherer et al., 2001) and in older adults in particular (Demenescu et al., 2014; Ross & Monnot, 2011).

While some studies have found that there were no significant gender effects (Hyde, 2014; Lima et al., 2014; Paulmann et al., 2008), very few studies have found evidence to suggest that males were better than females at emotion recognition. The discrepancy between the current findings and the previous research might be influenced by two influential studies included in the syntheses, where gender distribution was intertwined with variables such as mean age and educational level. The study with the highest proportion of older males (60%) had the youngest mean age (66.96 years old) and the highest level of education (15.08 years), while the other study that had the lowest proportion of older males (17.9%) had the highest mean age (75.2 years old). Exclusion of these studies resulted in non-significant results for gender distribution in the happiness analysis ($p = 0.94$). This underscores the need for additional research to unravel the interplay of variables contributing to gender distribution's impact on APR outcomes and to establish the true extent of its influence.

Furthermore, the analysis revealed a significant association between the mean age of older participants and effect size variation in happiness ($p < 0.001$), with a trending association noted in surprise ($p = 0.069$). Older participants with higher mean age showed significantly lower accuracy than their younger counterparts specifically in happiness recognition. However, no significant relationship between mean age and effect size variation was observed for other emotions. A study by Amorim et al. (2021) using a lifespan approach reported variations in age-related decline patterns across different emotion categories in APR. This diversity in age sensitivity across affective prosody categories may explain why the overall effect size variations in APR performance were not linked to the mean age of the older group. The aggregation of results from different affective prosodies might counteract the influence of mean age in the overall analysis.

4.3.2. Methodology-related moderators

In addition to participant characteristics, we explored methodological variability as a potential contributor to result heterogeneity. Specifically, we examined the impact of material type (for overall performance) and task difficulty (for both overall and emotion-specific performance) on result heterogeneity. Our investigation into material type aimed to discern whether the presence of semantic information played a role in result variability. Task difficulty was gauged through three variables: the number of speakers, the number of trials, and the number of answer options. Additionally, we explored the impact of positivity on the heterogeneous results related to surprise.

Our analysis revealed no significant association between material type and result heterogeneity. It is important to note that, as described in the Methods section, studies employing materials with verbal semantic content were also included. The meta-regression results suggested that material type did not contribute significantly to result variability. However, caution is warranted in interpreting

these findings, as they only indicate that the inclusion of semantic information did not have a significant impact on result variability across studies. It does not necessarily imply that semantic content had no influence on older adults' APR performance. Some previous research tentatively suggests that older adults might lean toward making emotional judgments based on semantic content (Ben-David et al., 2019). However, since this review specifically focuses on older adults' ability to accurately identify affective prosody of basic emotions, the interaction effect between affective prosody and content (neutral or emotion-laden) on older adults falls outside the scope of the current examination, warranting more comprehensive investigation.

Apart from material type, the influence of task difficulty, including the number of speakers, trials, and answer options, was investigated, and none of these moderators emerged as significant sources of heterogeneity across studies in either overall or emotion-specific performance. Among the 19 included studies, more than half ($N = 10$) presented stimuli using one or two speakers, and only one study used ten speakers. The gender distribution of speakers was not balanced, as only six studies demonstrated an equal ratio of male and female speakers, while five studies had entirely male speakers, and the rest exhibited a male speaker proportion of less than 50% or lacked male representation. The limited number and unbalanced gender distribution of speakers may be attributed to time and resource constraints faced by researchers, highlighting the need for a standardized affective prosody database with diverse speakers and balanced gender representation (Hearnshaw et al., 2019; M. Zhang et al., 2022).

Similarly, the number of trials and answer options did not have a significant effect on effect size heterogeneity in either overall or emotion-specific performance. In the studies analyzed, the number of trials ranged from 12 to 980, with more than half ($N = 11$) using less than 50 trials. The number of answer options ranged from two to eight, with more than half ($N = 12$) providing participants with no fewer than six options. Despite studies in psychiatric disorders suggesting that task complexity may increase cognitive demands and affect perception difficulties (Hoekert et al., 2007; Tang et al., 2022; M. Zhang et al., 2022), our results indicated that neither the number of trials nor the number of answer options had association with between-study heterogeneity. It is noteworthy that these null findings do not definitively indicate the absence of a relationship between moderators and effect size variation (Hedges & Pigott, 2004). The limited number of eligible studies may have constrained statistical power in tests investigating moderators (Borenstein, 2009). While it is plausible that the covariates were associated with the effect size, but the restricted number of qualifying studies may have hindered the conclusive identification of a relationship between these moderators and older adults' APR performances.

Positivity in surprise did not show significance in the meta-regression, suggesting no association with between-study heterogeneity in surprise performance. Caution is needed as there were only five studies reported the positivity of their surprise stimuli, with only one including negative surprise. The statistical power may be compromised, potentially attenuating the impact of positivity on older adults' APR performance due to the limited number of studies included (Borenstein, 2009). Given this limitation, further research is needed to include more eligible studies and conduct a more comprehensive analysis.

4.4. Limitations and implications

This study has several limitations that should be acknowledged. First, while a minimum of two empirical studies is theoretically adequate for a meta-analysis, the relatively small number of included studies in this synthesis may limit the robustness of the findings (Pigott, 2012; Valentine et al., 2010). A larger number of studies would enhance statistical power, particularly for emotion-specific analyses. Second, potential heterogeneity in the cognitive ability and hearing sensitivity of older adults across the included studies poses a limitation. Despite efforts to include studies reporting participants with normal cognitive and hearing abilities, the lack of standardized quantification criteria makes it challenging to control for these factors. While some studies reviewed suggest cognitive status and hearing acuity cannot fully explain the decline in APR ability in healthy older adults (Mitchell, 2007; Orbelo et al., 2005), future research should aim for more detailed and

standardized reporting of cognitive and hearing abilities to identify potential moderating variables more precisely.

The findings of this systematic review with meta-analysis highlight several directions for future research in the field. First, there is a notable gap in data specifically addressing tonal language users in APR studies for older adults. Given that all studies included in the current meta-analysis were conducted in non-tonal languages, the potential effects of language and cultural differences remain unexplored (Lin et al., 2020). Tonal languages have been suggested to exhibit restrictions in pitch range in emotional speech due to lexical tones (Ross et al., 1986; Wang et al., 2018), and aging-related decline in the right hemisphere (Kausler, 1994; Krzyżak, 2021) may impact spoken language comprehension processes, affecting pitch sensitivity for older adults (Rasmus & Błachnio, 2021). Investigating whether older adults in tonal language contexts face greater challenges in APR would be valuable. Cross-cultural research could contribute to a more comprehensive understanding of APR issues in older adults. Second, in line with Baglione et al. (2023), there is a lack of research directly investigating the neuroanatomical and neurophysiological mechanisms underlying APR decline. Researchers are encouraged to employ neuroimaging techniques to explore the neural mechanisms that underpin observed behavioral results.

5. Conclusion

This study conducted a comprehensive review and meta-analysis spanning 28 years of research with a focus on the age-related decline in affective prosody recognition (APR). The results indicate a significant age-related decline in recognizing basic emotions through affective prosody, with a valence-related asymmetry observed across emotions. Older adults demonstrated relatively preserved recognition of happiness and surprise compared to angry, sad, fearful, and disgusted prosody. Educational level was found to impact overall APR performance to some extent, and the mean age and gender distribution influenced performance in happy prosody. Future research needs to provide more precise information on participants' cognitive ability and hearing acuity. Additionally, researchers are encouraged to include participants with diverse linguistic and cultural backgrounds, and employ neuroimaging techniques for a comprehensive understanding of the neurobiological underpinnings of cognitive aging in connection with affective prosody processing. The results provide insights for conducting interdisciplinary studies by bridging psychology, speech-language-hearing sciences, and gerontology as well as for developing interventions and therapeutic strategies to enhance emotion perception and communication skills in older adults, contributing to improvements in their social and emotional well-being.

Supplementary Materials: The following supporting information can be downloaded at: Preprints.org.

Acknowledgments: This study was funded by grants from the Major Program of the National Social Science Foundation of China (18ZDA293). Yang Zhang received additional support from University of Minnesota's Brain Imaging Grant and Seed Grant.

Declaration of Conflicting Interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Amorim, M., Anikin, A., Mendes, A. J., Lima, C. F., Kotz, S. A., & Pinheiro, A. P. (2021). Changes in vocal emotion recognition across the life span. *Emotion, 21*(2), 315–325. <https://doi.org/10.1037/emo0000692>
2. Baglione, H., Coulombe, V., Martel-Sauvageau, V., & Monetta, L. (2023). The impacts of aging on the comprehension of affective prosody: A systematic review. *Applied Neuropsychology: Adult, 1*–16. Academic Search Complete.
3. Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614–636. <https://doi.org/10.1037//0022-3514.70.3.614>
4. Ben-David, B. M., Gal-Rosenblum, S., van Lieshout, P. H. H. M., & Shakuf, V. (2019). Age-related differences in the perception of emotion in spoken language: The relative roles of prosody and semantics. *Journal of Speech, Language, and Hearing Research, 62*(4S), 1188–1202. https://doi.org/10.1044/2018_JSLHR-H-ASCC7-18-0166

5. Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 358(1431), 561–572. <https://doi.org/10.1098/rstb.2002.1220>
6. Borenstein, M. (Ed.). (2009). *Introduction to meta-analysis*. John Wiley & Sons.
7. Brosgole, L., & Weisman, J. (1995). Mood recognition across the ages. *The International Journal of Neuroscience*, 82(3–4), 169–189. <https://doi.org/10.3109/00207459508999800>
8. Carstensen, L. L., & DeLiema, M. (2018). The positivity effect: A negativity bias in youth fades with age. *Current Opinion in Behavioral Sciences*, 19, 7–12. <https://doi.org/10.1016/j.cobeha.2017.07.009>
9. Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, 54(3), 165–181. <https://doi.org/10.1037/0003-066X.54.3.165>
10. Carstensen, L. L., & Mikels, J. A. (2005). At the intersection of emotion and cognition: Aging and the positivity effect. *Current Directions in Psychological Science*, 14(3), 117–121.
11. Chaby, L., Luherne-du Boullay, V., Chetouani, M., & Plaza, M. (2015). Compensating for age limits through emotional crossmodal integration. *Frontiers in Psychology*, 6, 691. <https://doi.org/10.3389/fpsyg.2015.00691>
12. Charles, S. T. (2010). Strength and vulnerability integration: A model of emotional well-being across adulthood. *Psychological Bulletin*, 136(6), 1068–1091. <https://doi.org/10.1037/a0021232>
13. Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Routledge. <https://doi.org/10.4324/9780203771587>
14. Cortes, D. S., Tornberg, C., Bänziger, T., Elfenbein, H. A., Fischer, H., & Laukka, P. (2021). Effects of aging on emotion recognition from dynamic multimodal expressions and vocalizations. *Scientific Reports*, 11(1), 2647. <https://doi.org/10.1038/s41598-021-82135-1>
15. Demenescu, L. R., Mathiak, K. A., & Mathiak, K. (2014). Age- and gender-related variations of emotion recognition in pseudowords and faces. *Experimental Aging Research*, 40(2), 187–207. <https://doi.org/10.1080/0361073X.2014.882210>
16. Dupuis, K., & Pichora-Fuller, M. K. (2015). Aging Affects Identification of Vocal Emotions in Semantically Neutral Sentences. *Journal of Speech, Language, and Hearing Research*, 58(3), 1061–1076. https://doi.org/10.1044/2015_JSLHR-H-14-0256
17. Duval, S., & Tweedie, R. (2000). Trim and Fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
18. Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
19. Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
20. Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198. [https://doi.org/10.1016/0022-3956\(75\)90026-6](https://doi.org/10.1016/0022-3956(75)90026-6)
21. Gaeta, L., & Brydges, C. R. (2020). An examination of effect sizes and statistical power in speech, language, and hearing research. *Journal of Speech, Language, and Hearing Research*, 63(5), 1572–1580. https://doi.org/10.1044/2020_JSLHR-19-00299
22. Gurera, J. W., & Isaacowitz, D. M. (2019). Emotion regulation and emotion perception in aging: A perspective on age-related differences and similarities. In *Progress in Brain Research* (Vol. 247, pp. 329–351). Elsevier. <https://doi.org/10.1016/bs.pbr.2019.02.007>
23. Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters* (Fourth edition). Advances Analytics, LLC.
24. Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing Meta-Analysis with R: A Hands-On Guide*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003107347>
25. Hearnshaw, S., Baker, E., & Munro, N. (2019). Speech perception skills of children with speech sound disorders: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, 62(10), 3771–3789. https://doi.org/10.1044/2019_JSLHR-S-18-0519
26. Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9(4), 426.
27. Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *American Psychologist*, 58(1), 78–79. <https://doi.org/10.1037/0003-066X.58.1.78>
28. Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539–1558. <https://doi.org/10.1002/sim.1186>
29. Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
30. Hoekert, M., Kahn, R., Pijnenborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research*, 96(1–3), 135–145. <https://doi.org/10.1016/j.schres.2007.07.023>

31. Hunter, E. M., Phillips, L. H., & MacPherson, S. E. (2010). Effects of age on cross-modal emotion perception. *Psychology and Aging, 25*(4), 779–787. <https://doi.org/10.1037/a0020528>
32. Hyde, J. S. (2014). Gender similarities and differences. *Annual Review of Psychology, 65*(1), 373–398. <https://doi.org/10.1146/annurev-psych-010213-115057>
33. IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open, 6*(7), e010247. <https://doi.org/10.1136/bmjopen-2015-010247>
34. Kausler, D. H. (1994). *Learning and memory in normal aging*. Academic Press.
35. Keshtiari, N., & Kuhlmann, M. (2016). The effects of culture and gender on the recognition of emotional speech: Evidence from Persian speakers living in a collectivist society. *International Journal of Society, Culture & Language, 4*(2), 71.
36. Kmet, L. M., Lee, R. C., & Cook, L. S. (2004). *Standard quality assessment criteria for evaluating primary research papers from a variety of fields*. Alberta Heritage Foundation for Medical Research.
37. Krzyżak, A. (2021). Different pattern of auditory processing lateralization in musicians and non-musicians. *Acta Neuropsychologica, 19*, 105–119.
38. Labouvie-Vief, G. (2003). Dynamic integration: Affect, cognition, and the self in adulthood. *Current Directions in Psychological Science, 12*(6), 201–206. <https://doi.org/10.1046/j.0963-7214.2003.01262.x>
39. Labouvie-Vief, G., Grünh, D., & Studer, J. (2010). Dynamic integration of emotion and cognition: Equilibrium regulation in development and aging. In R. M. Lerner, M. E. Lamb, & A. M. Freund (Eds.), *The Handbook of Life-Span Development* (Vol. 2, pp. 79–115). Wiley. <https://doi.org/10.1002/9780470880166.hlsd002004>
40. Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion, 12*(3), 529–539. <https://doi.org/10.1037/a0026827>
41. Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition & Emotion, 28*(3), 452–469. <https://doi.org/10.1080/02699931.2013.837378>
42. Lausen, A., & Schacht, A. (2018). Gender differences in the recognition of vocal emotions. *Frontiers in Psychology, 9*, 882. <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00882>
43. Lima, C. F., Alves, T., Scott, S. K., & Castro, S. L. (2014). In the ear of the beholder: How age shapes emotion processing in nonverbal vocalizations. *Emotion, 14*(1), 145–160. <https://doi.org/10.1037/a0034287>
44. Lima, C. F., & Castro, S. L. (2011). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion, 11*(5), 1021–1031. <https://doi.org/10.1037/a0024521>
45. Lin, Y., Ding, H., & Zhang, Y. (2020). Prosody dominates over semantics in emotion word processing: Evidence from cross-channel and cross-modal Stroop effects. *Journal of Speech, Language, and Hearing Research, 63*(3), 896–912. https://doi.org/10.1044/2020_JSLHR-19-00258
46. Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and cognitive functioning across the life span. *Psychological Science in the Public Interest, 21*(1), 6–41. <https://doi.org/10.1177/1529100620920576>
47. Maltezou-Papastyliaou, C., Russo, R., Wallace, D., Harmsworth, C., & Paulmann, S. (2022). Different stages of emotional prosody processing in healthy ageing-evidence from behavioural responses, ERPs, tDCS, and tRNS. *PloS One, 17*(7), e0270934. <https://doi.org/10.1371/journal.pone.0270934>
48. Martzoukou, M., Nasios, G., Kosmidis, M. H., & Papadopoulou, D. (2022). Aging and the perception of affective and linguistic prosody. *Journal of Psycholinguistic Research*. <https://doi.org/10.1007/s10936-022-09875-7>
49. Mill, A., Allik, J., Realo, A., & Valk, R. (2009). Age-related differences in emotion recognition ability: A cross-sectional study. *Emotion, 9*(5), 619–630. <https://doi.org/10.1037/a0016562>
50. Mitchell, Rachel L. C. (2007). Age-related decline in the ability to decode emotional prosody: Primary or secondary phenomenon? *Cognition & Emotion, 21*(7), 1435–1454. Academic Search Complete.
51. Morris, J. S., Scott, S. K., & Dolan, R. J. (1999). Saying it with feeling: Neural responses to emotional vocalizations. *Neuropsychologia, 37*(10), 1155–1163. [https://doi.org/10.1016/S0028-3932\(99\)00015-9](https://doi.org/10.1016/S0028-3932(99)00015-9)
52. Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society, 53*(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
53. Orbelo, D. M., Grim, M. A., Talbott, R. E., & Ross, E. D. (2005). Impaired comprehension of affective prosody in elderly subjects is not predicted by age-related hearing loss or age-related cognitive decline. *Journal of Geriatric Psychiatry & Neurology, 18*(1), 25–32. Academic Search Complete.
54. Paulmann, S., Pell, M. D., & Kotz, S. A. (2008). How aging affects the recognition of emotional speech. *Brain and Language, 104*(3), 262–269. <https://doi.org/10.1016/j.bandl.2007.03.002>
55. Paulmann, S., & Uskul, A. K. (2014). Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners. *Cognition & Emotion, 28*(2), 230–244. <https://doi.org/10.1080/02699931.2013.812033>

56. Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2008). Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61(10), 991–996. <https://doi.org/10.1016/j.jclinepi.2007.11.010>
57. Picou, E. M. (2016). How hearing loss and age affect emotional responses to nonspeech sounds. *Journal of Speech, Language, and Hearing Research*, 59(5), 1233–1246. https://doi.org/10.1044/2016_JSLHR-H-15-0231
58. Pigott, T. (2012). *Advances in meta-analysis*. Springer Science & Business Media.
59. Rasmus, A., & Blachnio, A. (2021). Auditory processing disorders in elderly persons vs. linguistic and emotional prosody. *International Journal of Environmental Research and Public Health*, 18(12), 6427. <https://doi.org/10.3390/ijerph18126427>
60. Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., Dahle, C., Gerstorf, D., & Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex*, 15(11), 1676–1689. <https://doi.org/10.1093/cercor/bhi044>
61. Reed, A. E., Chan, L., & Mikels, J. A. (2014). Meta-analysis of the age-related positivity effect: Age differences in preferences for positive over negative information. *Psychology and Aging*, 29(1), 1–15. <https://doi.org/10.1037/a0035194>
62. Ross, E. D., Edmondson, J. A., & Seibert, G. B. (1986). The effect of affect on various acoustic measures of prosody in tone and non-tone languages: A comparison based on computer analysis of voice. *Journal of Phonetics*, 14(2), 283–302. [https://doi.org/10.1016/S0095-4470\(19\)30669-2](https://doi.org/10.1016/S0095-4470(19)30669-2)
63. Ross, E. D., & Monnot, M. (2011). Affective prosody: What do comprehension errors tell us about hemispheric lateralization of emotions, sex and aging effects, and the role of cognitive appraisal. *Neuropsychologia*, 49(5), 866–877. <https://doi.org/10.1016/j.neuropsychologia.2010.12.024>
64. Ruffman, T., Dittrich, W., & Sullivan, S. (2009). Older adults' recognition of bodily and auditory expressions of emotion. *Psychology and Aging*, 24(3), 614–622. <https://doi.org/10.1037/a0016356>
65. Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, 32(4), 863–881. <https://doi.org/10.1016/j.neubiorev.2008.01.001>
66. Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: Brain regions involved in processing anger prosody. *NeuroImage*, 28(4), 848–858. <https://doi.org/10.1016/j.neuroimage.2005.06.023>
67. Sbattella, L., Colombo, L., Rinaldi, C., Tedesco, R., Matteucci, M., & Trivilini, A. (2014). Extracting emotions and communication styles from prosody. In H. P. Da Silva, A. Holzinger, S. Fairclough, & D. Majoe (Eds.), *Physiological Computing Systems* (Vol. 8908, pp. 21–42). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45686-6_2
68. Schaffer, S. G., Wisniewski, A., Dahdah, M., & Froming, K. B. (2009). The comprehensive affect testing system-abbreviated: Effects of age on performance. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 24(1), 89–104. <https://doi.org/10.1093/arclin/acp012>
69. Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, 32(1), 76–92. <https://doi.org/10.1177/0022022101032001009>
70. Schmidt, J., Janse, E., & Scharenborg, O. (2016). Perception of emotion in conversational speech by younger and older listeners. *Frontiers in Psychology*, 7, 781. <https://doi.org/10.3389/fpsyg.2016.00781>
71. Seddoh, A., Blay, A., Ferraro, R., & Swisher, W. (2020). Prosodic perception in aging individuals: A focus on intonation. *Current Psychology*, 39(4), 1221–1233. <https://doi.org/10.1007/s12144-018-9806-x>
72. Sen, A., Isaacowitz, D., & Schirmer, A. (2018). Age differences in vocal emotion perception: On the role of speaker age and listener sex. *Cognition & Emotion*, 32(6), 1189–1204. <https://doi.org/10.1080/02699931.2017.1393399>
73. Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience*, 6(3), 309–315. <https://doi.org/10.1038/nn1008>
74. Tang, E., Zhang, M., Chen, Y., Lin, Y., & Ding, H. (2022). Recognition of affective prosody in bipolar and depressive conditions: A systematic review and meta-analysis. *Journal of Affective Disorders*, 313, 126–136. <https://doi.org/10.1016/j.jad.2022.06.065>
75. Thompson, B. (2007). Effect sizes, confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools*, 44(5), 423–432. <https://doi.org/10.1002/pits.20234>
76. Todorova, G. K., Hatton, R. E. M., & Pollick, F. E. (2019). Biological motion perception in autism spectrum disorder: A meta-analysis. *Molecular Autism*, 10(1), 49. <https://doi.org/10.1186/s13229-019-0299-8>
77. Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215–247. <https://doi.org/10.3102/1076998609346961>

78. Velikonja, T., Fett, A.-K., & Velthorst, E. (2019). Patterns of nonsocial and social cognitive functioning in adults with autism spectrum disorder: A systematic review and meta-analysis. *JAMA Psychiatry*, *76*(2), 135–151. <https://doi.org/10.1001/jamapsychiatry.2018.3645>
79. Wang, T., Lee, Y., & Ma, Q. (2018). Within and across-language comparison of vocal emotions in Mandarin and English. *Applied Sciences*, *8*(12), Article 12. <https://doi.org/10.3390/app8122629>
80. Wechsler, D. (1997). *The Wechsler Adult Intelligence Scale—3 Edition: Technical manual*. San Antonio: The Psychological Corporation.
81. Wildgruber, D., Pihan, H., Ackermann, H., Erb, M., & Grodd, W. (2002). Dynamic brain activation during processing of emotional intonation: Influence of acoustic parameters, emotional valence, and sex. *NeuroImage*, *15*(4), 856–869. <https://doi.org/10.1006/nimg.2001.0998>
82. Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., & Ackermann, H. (2005). Identification of emotional intonation evaluated by fMRI. *NeuroImage*, *24*(4), 1233–1241. <https://doi.org/10.1016/j.neuroimage.2004.10.034>
83. Wong, B., Cronin-Golomb, A., & Nearing, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology*, *19*(6), 739–749. <https://doi.org/10.1037/0894-4105.19.6.739>
84. Zhang, M., Xu, S., Chen, Y., Lin, Y., Ding, H., & Zhang, Y. (2022). Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. *Autism*, *26*(4), 798–813. <https://doi.org/10.1177/1362361321995725>
85. Zhang, X., Cheng, B., & Zhang, Y. (2022). A hands-on tutorial for systematic review and meta-analysis with example data set and codes. *Journal of Speech, Language, and Hearing Research*, *65*(9), 3217–3238. https://doi.org/10.1044/2022_JSLHR-21-00607

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.