**BENTHAM OPEN**

**CrossMark**

# The Open Psychology Journal
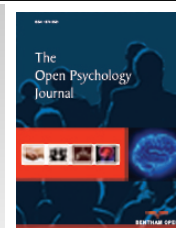
RESEARCH ARTICLE

# Priming Younger and Older Adults' Sentence Comprehension: Insights from Dynamic Emotional Facial Expressions and Pupil Size Measures

Maria Nella Carminati[1] and Pia Knoeferle[2,*]

[1]*Bielefeld University, Bielefeld, Germany*
[2]*Humboldt University Berlin zu Berlin, Institute for German Language and Linguistics, Berlin, Germany*

**Abstract:**

*Background:*

Prior visual-world research has demonstrated that emotional priming of spoken sentence processing is rapidly modulated by age. Older and younger participants saw two photographs of a positive and of a negative event side-by-side and listened to a spoken sentence about one of these events. Older adults' fixations to the mentioned (positive) event were enhanced when the still photograph of a previously-inspected positive-valence speaker face was (*vs.* wasn't) emotionally congruent with the event/sentence. By contrast, the younger adults exhibited such an enhancement with negative stimuli only.

*Objective:*

The first aim of the current study was to assess the replicability of these findings with dynamic face stimuli (unfolding from neutral to happy or sad). A second goal was to assess a key prediction made by socio-emotional selectivity theory, *viz.* that the positivity effect (a preference for positive information) displayed by older adults involves cognitive effort.

*Method:*

We conducted an eye-tracking visual-world experiment.

*Results:*

Most priming and age effects, including the positivity effects, replicated. However, against our expectations, the positive gaze preference in older adults did not co-vary with a standard measure of cognitive effort - increased pupil dilation. Instead, pupil size was significantly bigger when (both younger and older) adults processed negative than positive stimuli.

*Conclusion:*

These findings are in line with previous research on the relationship between positive gaze preferences and pupil dilation. We discuss both theoretical and methodological implications of these results.

**Keywords:** Cognitive effort, Emotional processing, Facial expressions, Positivity effect, Pupil dilation, Sentence processing, Speaker cues, Visual-world paradigm.

## 1. INTRODUCTION

The development of the visual world paradigm in recent years has made it possible for psycholinguists to investigate

\* Address correspondence to this author at the Humboldt University zu Berlin - Institute for German Language and Linguistics, Berlin, Germany; Tel: +493020939672; E-mail: pia.knoeferle@hu-berlin.de

the visual context as one important source of context effects on language processing. Overall, findings show that visual context information exerts a very rapid effect on language processing and influences, for instance, the resolution of syntactically ambiguous sentences [1, 2]. Findings also show that information such as an object's size, color [3], or shape [4], depicted clipart events [2], real-world action events [5], action affordances [6], and the spatial location of objects [7] are all rapidly integrated during sentence comprehension and can affect a listener's visual attention within a few hundred milliseconds (for a recent review, see [8]**)**.

A relatively new topic of interest in current research on language-vision interaction is how visible speaker-based cues such as eye gaze and gestures influence language comprehension and how this process varies depending on properties of the listener. Here, too, available evidence points to a rapid and incremental integration: When a speaker inspects and mentions objects, listeners rapidly align their gaze with the gaze of the speaker [9 - 12]. But some variation in the deployment of listeners' attention exists, for instance, as a function of their literacy [13 - 15].

## 1.1. Positivity Biases in Emotional Priming of Sentence Processing

In the context of speaker-based effects and listener variation [16], (henceforth C&K) investigated for the first time how a speaker's facial emotional expression affects incremental sentence processing as a function of participants' age, comparing groups of younger ($M_{age}$ = 24) and older participants ($M_{age}$ = 64 [16], see [17] for relevant research on facial expressions in emotion research). The motivation for the age manipulation came from the socio-emotional selectivity theory's interpretation of lifespan changes in emotion processing [18 - 20]. According to this theory, older people focus more on positive than negative information, a 'positivity effect' that is absent, or even reversed, in younger adults [19]. This developmental trend has been argued to be responsible for the so-called positivity effects, which have been found across a number of experimental paradigms and range of stimuli. For example, in studies of visual attention using dot-probe and eye-tracking paradigms, older adults displayed an attentional bias away from negative and towards happy facial expressions [18, 21]. Younger adults, on the other hand, show no preference [18], or prefer negative faces [22].

In C&K's study, the older and younger adults saw either a smiling or a sad face, which they were told was the face of the speaker (thus simulating a speaker-hearer scenario). The next screen replaced the speaker's face with two emotional pictures from the International Affective Picture System database (IAPS [23]), one positive and one negative, displayed side by side. Participants then heard the speaker's voice describe either the positive or the negative picture. During sentence presentation, participants' eye movements to the IAPS pictures were recorded. As is usually the case in 'visual world' language comprehension tasks, participants began to look more at the IAPS picture mentioned in the sentence and less at the other picture. C&K's main interest was, however, whether the previously seen emotional face would facilitate the processing of the sentence if the emotional facial prime was emotionally congruent (*vs.* incongruent) with the sentence and whether such facilitation would exhibit a positivity bias in older adults. In line with their expectations, facilitation through the facial prime emerged rapidly (as soon as the sentence referenced one or the other IAPS picture) and occurred for the older adults with positive sentences only. For the younger adults, facilitation emerged with negative sentences only, and/or (depending on the time region considered) to a lesser extent for positive sentences (see [24] for a description of possible patterns consistent with a positivity effect and see [25, 26] on gaze patterns reflecting processing facilitation).

These findings constitute the first evidence of listener positivity effects in the domain of incremental language processing. However, one possible concern with these results is that the (static) facial primes may not have been powerful enough to elicit priming across the board (*i.e.*, in younger adults also with positive and in older adults also with negative faces). Perhaps the observed positivity bias would disappear with more powerful facial stimuli. Furthermore, participants also verified whether (or not) the sentence matched the face by pressing a 'yes' or 'no' button on each experimental trial. Although positivity effects were expected, they did not emerge for these reaction times and accuracy. Instead, older adults' behavioural responses were similar to younger adults'. Response times were slowed for both older (as expected from their positivity effects in emotion processing) *and* younger adults when a negative face was paired with a negative sentence. Assessing the replicability of these online and offline results is a first goal of the present research.

## 1.2. Does the Positivity Bias Implicate Cognitive Effort?

In addition to more fully assessing the robustness and replicability of these gaze positivity preferences, we must better understand the mechanisms implicated in the observed gaze positivity preference. One prediction of socio-emotional selectivity theory is that, because goal processes are motivationally based, positivity effects should require

some cognitive effort in the older individual. For example, Mather and Knight 2005 [27] observed positivity effects for memory of emotional faces only among older individuals scoring high on a series of cognitive control/executive function tests Knight M *et al.*, [28] reported positivity effects (in eye movements to pairs of emotional faces and pictures), under conditions of full attention only. Positivity effects were, by contrast, absent under conditions of divided attention (when cognitive resources are presumably more limited), suggesting these effects require cognitive resources.

However, other evidence suggests that cognitive effort is not necessary for positivity effects to emerge. For example, Thomas RC *et al.*, [29] asked younger and older participants to perform a parity decision task about two numbers, while distracting emotional words appeared between the numbers. During an unannounced, post-experiment word memory task, a positivity effect emerged in that older participants recognized a higher proportion of positive relative to negative words, suggesting possibly automatic, effortless processing of emotional information. Moreover, when measuring fixation preferences to emotional-neutral and emotional-emotional image pairs, Rosler A *et al.*, [30] found no difference in positivity bias between a group of older healthy and a group of mildly cognitively impaired individuals: both preferred to fixate away from negative and towards neutral images as compared to young adults.

More recently, Allard ES *et al.*, [31] measured pupil dilation, which is commonly assumed to index cognitive effort, while groups of younger and older participants looked at emotional-neutral face pairs. They found that older people's positive gaze preferences (*i.e.*, a preference for a neutral over a negative face, and a preference for a positive over a neutral face) were *not* associated with an increase in pupil dilation: In fact, for both younger and older participants, pupil dilation was larger for fixations on sad faces than on positive, fearful and angry faces. Allard ES *et al.*, [31] also manipulated mood between participants (positive, negative, neutral) and interestingly, found that neutral mood, irrespective of age, triggered the biggest change in pupil dilation; however, crucially they found no significant Age x Mood x Emotion interaction and concluded that gaze acts as a rather effortless and economical regulatory tool during older individuals' processing of emotional information. In sum, the evidence for the role of cognitive effort in positivity effects is controversial. To the best of our knowledge, Allard ES *et al.*, [31] is the only research to date that has directly investigated the association between positivity effects - as manifested in fixation preferences, and cognitive load - as indexed by pupil dilation.

## 1.3. The Present Research

The present research thus examined the robustness and replicability of the positivity effects reported by C&K with stronger emotional facial primes. In addition, we examined whether the eye-gaze positivity effects implicate cognitive effort.

To examine the replicability of these effects, we strengthened the emotional facial primes by using dynamic instead of static stimuli. Although most of the research on the processing of emotional faces has used static faces (*i.e.*, still photographs), in everyday life, the facial display of emotion is a highly dynamic process. A crucial difference between static and dynamic faces is that the former lack information as to the possible direction and speed in which the emotional facial expression changes over time (for example, while moving from a neutral facial expression to a full-blown expression of happiness or sadness). This transitional information matters in the recognition of certain facial expressions [32, 33]. Compared with static faces, dynamic emotional expressions elicit greater responses in brain regions concerned with the processing of facial emotional information [34 - 38] and produce stronger emotional effects on their viewers [39, 40]. Particularly relevant for us is that older people's recognition of facial emotions improves with dynamic faces under certain circumstances [41 - 43]. In one study, older adults were as accurate as younger adults in recognizing (negative and positive) dynamic faces, but were less accurate with static faces [44]. In sum, dynamic faces should be more effective as primes, particularly for older adults.

Concerning the mechanisms implicated in the positivity effects, we analysed the association between positivity effects (as manifested in fixation preferences) and cognitive load (as indexed by pupil dilation) for both the dynamic facial prime study and the data from C&K. Most research aimed at demonstrating that pupil size reflects cognitive effort [45 - 47] has ignored the empirical question of how pupil size relates to fixation preferences and fixation durations. Although the assumption appears to be that longer and/or more frequent fixations (*e.g.*, on a face or while reading the words of a sentence) are positively correlated with pupil size [31], to our knowledge there is no research that has looked systematically at the relationship between these two psychophysical variables. We believe that the current study contributes towards the understanding of this relationship.

If older people's positive gaze preferences require the exertion of cognitive effort, increases in pupil dilation should

be observed correspondingly. In our particular case, older people's looks to a positive picture (referenced by a positive sentence) were enhanced when the face was positive, but there was no corresponding enhancement of looks to a negative picture (referenced by a negative sentence) when the face was negative. Thus, a straightforward prediction is that for older people pupil size should be greater when they looked at the positive than the negative picture.

However, another outcome is also possible. Because positivity effects have been defined both as an increased focus on positive information and as a decreased focus on negative in older adults [24], it may be that cognitive effort is required both in enhancing attention to positive information and/or in inhibiting or decreasing attention to negative information. For our experiments, this would predict no difference in pupil size for older adults between looks to the positive and negative picture. But because the necessary requirement of a positivity effect is a Valence x Age interaction, what is most crucial is that the pattern of pupil size as a function of the valence of the fixated picture should differ for young and older adults (a significant Age x Picture interaction). If, on the other hand, young and older adults exhibit the same pattern of behavior, then we cannot make any firm conclusions about cognitive effort having been exerted by older people with regard to the positivity effects found in the gaze measure. Below we first describe the methods of a new experiment with dynamic facial prime stimuli (N=16 in each age group) and subsequently outline how we analysed the joint data from this new study and from C&K for potential associations between pupil dilation and positivity preferences in the fixation pattern to assess the role of cognitive effort in the positivity gaze preference.

## 2. METHOD

### 2.1. Participants

The participants were 16 older and 16 younger adults; their demographic characteristics are given in Table **1**. Older adults with no history of neuropsychiatric disorders were recruited through advertisements posted in the university and in other public places of the city of Bielefeld, Germany. The younger participants were students at the University of Bielefeld. Younger and older participants received a monetary reward for their participation in the experiment and gave informed consent. None of the participants had taken part in the previous study with static faces.

**Table 1. Cognitive test results and demographic characteristics of younger and older adults.**

| Characteristic | Younger | Older |
|---|---|---|
| Age range | 18-30 | 60-80 |
| Mean age in years | 23.8 (2.3) | 67.8 (5.3) |
| Animal naming[a] | 30.81(5.96) | 26.81 (7.15) |
| Picture completion[b] | 3.81 (.98) | 4.25 (1.00) |
| Digit Symbol[b] | 90.12 (8.88) | 62.31 (11.99) * |
| Word naming [c] | 13.06 (4.34) | 13.87 (3.72) |
| Digit span[b] | 18.75 (3.04) | 16.31 (2.15) * |
| Similarities[b] | 13.81 (1.32) | 13.81 (2.4) |
| BMIS scores[d] | 8.18 (5.95) | 10.68 (6.35) |
| Male/female (*n*) | 4/12 | 5/11 |

*Note*. Standard deviations are in parentheses
[a] Task: Name as many animals as possible, time allowed: 1 min
[b] latest German version of the Wechsler Adult Intelligence Scale (WAIS) [70, 71].
[c] Task: Name as many words as possible starting with the letter 'l', time allowed: 1 min
[d] Brief Mood Introspection Scale questionnaire (BMIS [72],) translated
into German. The mood rating for each participant was obtained by subtracting the scores
for negative mood from those of positive mood.
* Younger and older adults' means differ significantly (*t* test, $p < .05$)

As the University of Bielefeld did not yet have an official institutional Ethical Review Board, at the time of grant application, we raised the issue of ethics approval with our sponsor (The German Research Foundation, DFG). They replied that under their standard procedure for psycholinguistic research, obtaining ethics approval would not be necessary for our research (a copy of the communication is available from the authors). We took all the necessary steps to conduct the research following the guidelines laid down in the Declaration of Helsinki. Before the experiment all participants read an information sheet in which they were informed about the experiment and the tasks involved, about the potential risks and discomforts (none known, since these were non-invasive behavioral experiments), and about data treatment. After reading the information sheet and receiving any further clarifications from the experimenter,

participants signed a written informed consent form. The form clarified that participants could discontinue the study at any time if they wished to do so. During the experiment the experimenter ensured that participants were feeling comfortable and well. After the experiment participants were debriefed and received an answer to any further questions. Copies of the Information Sheet and the Informed Consent form are available from the authors.

The categorization of participants by age was exclusively for the purpose of the experimental manipulation of our study. We controlled for key cognitive variables with the administration of cognitive tests.

## 2.2. Materials

Materials (28 experimental and 56 filler items) consisted of video clips of emotional facial expressions, emotional pictures and auditorily presented sentences. Each experimental item consisted of a video clip showing a transition from a neutral to a happy or sad facial expression, a display showing a positive and a negative picture taken from the IAPS database [23] and a sentence describing either the positive or negative picture. Speaker facial emotion and sentence valence were fully crossed such that the sentence could have the same or the opposite emotional valence as the initially presented speaker face (so the design was 2 (Face: sad *vs*. happy) x 2 (Sentence: negative *vs*. positive) x 2 (Picture (negative *vs*. positive)).

All materials, except for the video clips, were the same as in the previous study by C&K. For details about the sentences and IAPS pictures, we refer the reader to the Methods section of C&K (see Fig. **1** for an example of a negative/positive experimental sentence; a list of the sentences and IAPS pictures is available in the **Supplementary Information** of C&K). The same face models were associated with the experimental items in C&K and the current experiment. During the recording of the video clips for the dynamic faces, models were instructed to naturally change their expression from neutral into happy or sad. We also recorded them simulating steady-state neutral expressions (these were used in a proportion of the filler items). To maximize visual similarity, all recordings took place in the same room, against a white background, with the model at the same distance from the camera and in similar light conditions. After recording, we selected the best video clips for each model in the happy, sad and (steady-state) neutral condition. Each clip was then cut to a length of 5 seconds. In the happy/sad dynamic videos the initial neutral expression was maintained for about 1.3 seconds; then the expression changed and the model maintained the full-blown happy or sad expression until the end of the clip. Example video clips are available in the **Supplementary Information**.
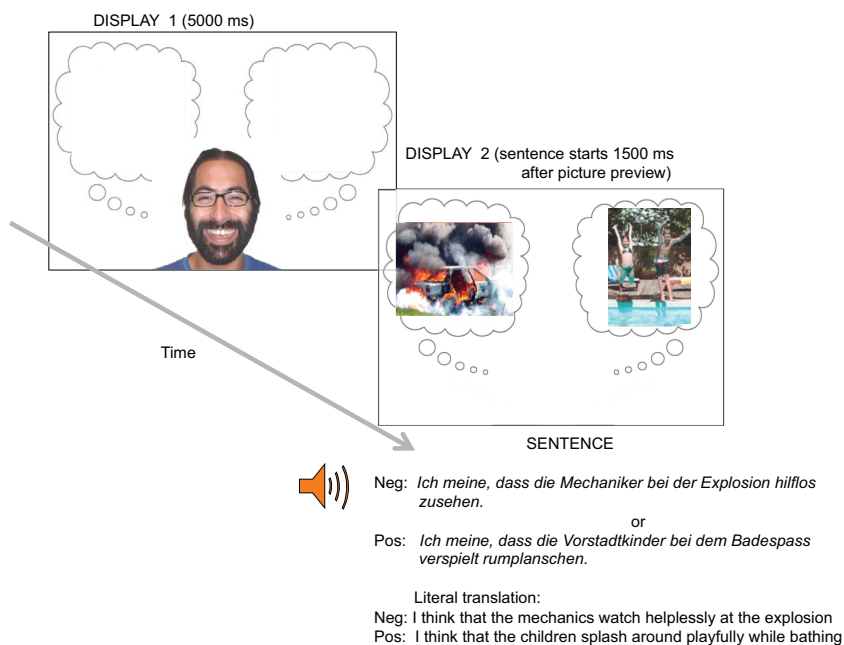


**Fig. (1).** Sequence of events in an experimental trial. Display 1 shows the snapshot of a full-blown neutral-to-happy face change. For copyright restrictions the IAPS pictures in this figure have been replaced with similar pictures downloaded from the web.

## 2.3. Procedure

The procedure was the same as in C&K. The experimental session started with the collection of participants' demographic details, followed by the administration of some cognitive tests and of a mood questionnaire (see the notes to Table **1**). Eye movements were recorded using an Eyelink 1000 head-stabilised Desktop eye tracker (SR Research, Mississauga, Ontario, Canada). The sequence of events in an experimental trial is illustrated in Fig. (**1**).

A trial started (see Fig. **1**) with the video clip of the dynamic face (the prime face, see Display 1) and empty thought bubbles on either side (duration of the video clip: 5000 ms, Display 1 in Fig. (**1**)). Next, two pictures one depicting a positive and the other a negative event respectively appeared on the screen side by side, inside the thought bubbles (Fig. **1**, Display 2). 1500 ms after the onset of Display 2 a sentence about one of the two event pictures was played out over the loudspeakers. Display 2 remained on the screen for 1500 ms after sentence end for young participants, and for 3000 ms for the older participants, after which it disappeared and the trial ended.

Participants were told that the study investigated language comprehension in relation to a visual display on the computer screen: They would first see the face of a person who was thinking about something (future thoughts about an event were given a place holder in the form of the empty thought bubbles) and was about to speak, and after that they would hear him/her utter a sentence which described one of two pictures shown on the screen. The task was to look, listen and understand the sentence, and decide as quickly and accurately as possible whether the valence of the face matched the valence of the sentence ("Does the face match the sentence?") by pressing one of two buttons. The timeout for answering the question was the clearing of Display 2. The experimental session, comprising the administration of the cognitive tests and the eye-tracking experiment, lasted on average 45 minutes for the younger adults, and one hour or longer for the older participants.

## 3. ANALYSES

Below we first describe the analysis procedure for the pre-test, the eye tracking, and the verification task data which was the same as for C&K. Subsequently we explain how we analysed the pupil dilation measures for both the present experiment and the data from the static-face version of the experiment reported in C&K (pupil dilation had not been analysed in C&K).

### 3.1. Replication of C&K

#### 3.1.1. Pre-test Data

For each participant, we computed scores for the mood and cognitive tests according to test-specific instructions (see Table **1**). *T*-tests for independent samples were carried out to test for differences between young and older adults on these variables.

#### 3.1.2. Eye-movement Data

Just as for the static-face study, we analyzed eye movements during the presentation of Display 2 (see Fig. **1**). As in that study, one fixation measure was the mean log ratio of the probability of looking at the picture of the negative event over the probability of looking at the picture of the positive event ($ln(p$(negative picture)$/p$(positive picture)) [25]). This measure expresses the inspection bias towards the negative relative to the positive picture. A positive value indicates a bias for the negative over the positive picture, a negative value a bias for the positive picture, and zero indicates no bias. Based on C&K, the time region of interest was the post-NP1 onset region, which started with the onset of NP1, *i.e.* the noun phrase in the sentence which disambiguated towards the positive or the negative picture (referred to as 'long region', see Table **2**) and ended with the end of the sentence. This region was further subdivided into NP1, NP2, Adverb and Verb regions (see Table **2** for details). Note that a sentence effect (*i.e.*, fixations on the pictures as a function of the sentence being heard), as well as a facilitation effect from the face on sentence processing (*i.e.*, face-sentence priming) can only be expected after referential disambiguation (*i.e.*, after NP1 onset). Following C&K, we in addition analyzed the duration of the first fixation after the onset of NP1, an eye tracking measure of early processing.

**Table 2. Critical regions and their duration (*SD* in parentheses).**

| Critical Region | Start | End | Avg duration in ms |
|---|---|---|---|
| Pre-NP1 | Approx 300 ms after appearance of Display 1 | NP1 onset | 3000 (0) |
| Post-NP1 | | | |

*(Table 4) contd.....*

| Critical Region | Start | End | Avg duration in ms |
|---|---|---|---|
| NP1<br>*die Mechaniker* | NP1 onset | NP2 onset | 1300 (226) |
| NP2<br>*bei der Explosion* | NP2 onset | Adverb onset | 1100 (267) |
| Adverb<br>*hilflos* | Adverb onset | Verb onset | 900 (137) |
| Verb<br>*zusehen.* | Verb onset | Verb offset<br>(end of sentence) | 716 (108) |
| Long region<br>*die Mechaniker bei der Explosion hilflos zusehen* | NP1 onset | Verb offset | 4016 (456) |

Note. The onsets of NP1, NP2 and the Adverb within an item were the same for the negative and the positive sentence.

### 3.1.3. Reaction Times and Accuracy

Apart from looking at the displays and listening to the sentence, participants had to decide with a button press whether the face matched the sentence in valence or not. Reaction times (RTs) to answer the question ("Does the face match the sentence?") were measured from the onset of NP1 (*i.e.*, the start of the disambiguating region) until the time of the button press. We removed trials with timeouts (N=12 for older and N=8 for younger adults), and incorrectly-answered trials (N=72 for older and N=26 for younger adults). The thus-filtered RTs were submitted to ANOVAs by participants and items. The fixed factors were Prime face (positive *vs*. negative), Sentence (positive *vs*. negative) and Age (young *vs*. old). Prime face and Sentence were within-participants and -items factors in the participants and items analysis, while Age was between in the participants and within in the items analysis.

Accuracy scores were computed on a total of 876 observations (N=440 / 98% for young and N=436 / 97% for old) after removing timeouts (N=20). A logistic linear mixed effect (LME) model was fitted to the binary (*i.e.*, correct *vs*. incorrect) response data [48]. In this model the predicted outcome was the response and the predictors were face and sentence valence, each with two levels (negative *vs*. positive), and age (young *vs*. old). Participants and items, with their intercepts and slopes and the intercept x slope interactions, were included in the random effects part of the model. For the predictors we transformed the fixed effect coding into a numerical value and centered it so as to have a mean of 0 and a range of 1 [49]. Centering reduces collinearity of the predictor variables, and allows the coefficients of the regression to be interpreted as the main effects in a standard ANOVA [50].

### 3.2. Analyses of Pupil Dilation

In addition to eye fixations and saccades, our eye tracking equipment (Eyelink 2000, SR Research) also measures pupil diameter. Pupil size has been shown to vary as a function of cognitive load [51] and can thus be used to test the hypothesis that positivity effects, as defined by socio-emotional selectivity theory, require cognitive effort.

We restricted analysis of pupil size to the conditions where positivity effects were observed, *i.e.*, in which face and sentence matched in valence (negative face + negative sentence, and positive face + positive sentence; mismatching conditions in which no positivity gaze preferences emerged were omitted). Following the same logic, for positive congruent conditions only fixations on the positive picture were included and for the negative congruent conditions only fixations on the negative picture were included. Because both experiments (with static and dynamic faces respectively) yielded the same pattern and timing of positivity effects, the present analysis on pupil size was done on the merged data of the two experiments (C&K and the present data, yielding *N*=48 for each age group).

We analysed pupil dilation during the part of the sentence that disambiguated towards the positive or the negative picture in the display (*i.e.*, when positivity preferences had been observed). We used two measures of pupil size, the absolute pupil size, provided for each fixation by our eye processing software, and a "normalized" pupil size derived through a transformation from the absolute pupil size, similar to the measure used by [31]. The second measure was to control for individual and age differences in pupil reactivity (older adults tend to have a smaller pupil size and a smaller range in pupil dilation).

For each fixation, the corresponding absolute pupil size was recorded and a mean pupil size per participant, item, condition, word region and fixated areas of interest (*i.e.*, negative or positive picture) was computed. To compute the corresponding mean normalized pupil size, we first extracted the absolute smallest and greatest recorded pupil size for each participant (computed based on fixations to the pictures in the display over the whole course of the experiment). Note that the pictures included experimental (*N*= 28) as well as filler (*N*=56) trials, and that in filler trials the IAPS pictures showed mostly neutral events. The minimum pupil size was then subtracted from the maximum pupil size

[maximum pupil size - minimum pupil size], giving the range of pupil dilation for each participant, this range being an estimate of the participant's pupil reactivity to the pictures during the experiment. The mean normalized pupil size was calculated according to the following formula [current absolute pupil size - minimum absolute pupil size]/ [maximum pupil size - minimum pupil size], the result of this ratio representing the percentage change in pupil dilation [31, 47]. We entered the pupil size values into mixed regression models with participants and items as random effects. One advantage of this method is that random effects for both items and participants can be jointly modelled [52].

## 4. RESULTS

### 4.1. Replication of C&K with Dynamic Facial Primes

#### 4.1.1. Matching Participants for Mood and Cognitive Processing (pre-test)

Please see Table **1** for a summary of the pre-test results and demographic details. Table **1** shows that older people performed significantly worse than young adults only in the Digit Symbol and the Digit Span tasks of the Wechsler Adult Intelligence Test. In all the other tests, including the BMIS mood-rating test, the two groups did not differ significantly from each other. This suggests that participants in the two age groups were well matched for mood but differed somewhat in their verbal storage capacity (digit span) and in basic processing operations (digit symbol).

#### 4.1.2. Do We Replicate the Descriptive Positivity and Negativity Gaze Preferences from C&K (Time Course Graphs)?

Fig. **(2 A and B)** plots the time course of the fixation pattern in the post-NP1 region for each age group. This graph is based on mean log ratios computed on successive 20 ms time slots as a function of prime face and sentence valence. Fig. **(2 A and B)** illustrates that participants made more fixations to the mentioned than the unmentioned picture from about 500 ms after NP1 onset (sentence effect): The red lines for the two negative sentence conditions rise above zero (a preference for the negative picture) while the black lines for the positive sentence conditions go in the opposite direction (indicating a preference for the positive picture).
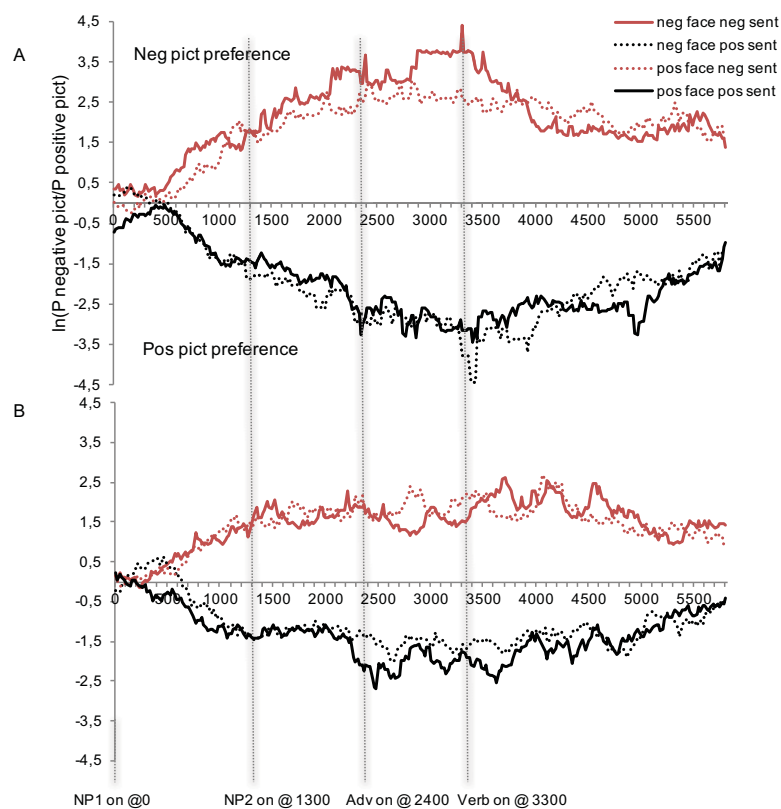


**Fig. (2).** Mean log gaze probability ratios for (**A**) young and (**B**) older participants in the Post -NP1 onset region (N=32, dynamic face study).

Crucially, a clear facilitatory effect of the (valence-matching) prime face on sentence processing replicated in the

relative distance between the solid and the dotted line within each sentence condition: For the young adults, face-sentence priming happens in the negative sentence conditions, while for the older adults it emerges in the positive sentence conditions (especially during the adverb, see Fig. **2**).

### 4.1.3. Do We Replicate the Reliable Positivity and Negativity Gaze Preferences from C&K? (ANOVAs on Regions)

Inferential analyses (repeated-measures ANOVAs with participants and items as random effects, see [53]) were then performed on the mean log ratios for the (long) post-NP1 region and the individual word regions after NP1 onset (NP1, NP2, Adverb, Verb), see Table **2**. In these ANOVAs, the fixed factors were Prime face (positive *vs.* negative), Sentence (positive *vs.* negative) and Age (young *vs.* old). Prime face and sentence were within-participants and -items factors in the participants and items analysis, while age was between in the participants and within in the items analysis.

The ANOVA results are summarized in Table **3**. As expected, there was a significant sentence effect in all regions in that participants looked more at the mentioned than the unmentioned picture post-NP1. There was further a significant Sentence by Age interaction, however, this result is irrelevant to our research question (see C&K, **Supporting Information S3**).

**Table 3. Summary of ANOVA results. Degrees of freedom: $F1(1,30)$, $F2 (1,27)$. $P$ values are under the corresponding $F$-values.**

| Effects and interactions | NP1 | | NP2 | | Adverb | | Verb | | Long region | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $F1$ | $F2$ | $F1$ | $F2$ | $F1$ | $F2$ | $F1$ | $F2$ | $F1$ | $F2$ |
| | $p1$ | $p2$ | $p1$ | $p2$ | $p1$ | $p2$ | $p1$ | $p2$ | $p1$ | $p2$ |
| Age | < 1 | < 1 | 1.87 .10 | < 1 | < 1 | < 1 | < 1 | 1.76 .19 | < 1 | 1.93 .18 |
| Face (face-picture priming) | 7.7 .00* | 1.44 .24 | < 1 | < 1 | 1.98 .17 | < 1 | < 1 | < 1 | 5.73 .03* | 2.75 .11 |
| Sentence (sentence-picture congruence) | 98.56 .00* | 67.64 .00* | 167.51 .00* | 216.41 .00* | 174.45 .00* | 473.23 .00* | 186.39 .00* | 575.75 .00* | 365.66 .00* | 241.21 .00* |
| Face x sentence (face-sentence priming) | < 1 | 1.23 .30 | < 1 | < 1 | < 1 | < 1 | < 1 | < 1 | < 1 | < 1 |
| Face x age | < 1 | < 1 | < 1 | 1.56 .22 | < 1 | < 1 | < 1 | < 1 | < 1 .38 | 4.34 .04* |
| Sentence x age | 2.09 .16 | 3.06 .09 | 11.41 .00* | 26.32 .00* | 10.13 .00* | 18.95 .00* | 7.13 .01* | 21.01 .00* | 9.48 .00* | 4.34 .04* |
| Face x sentence x age (face-sentence priming x age) | 7.72 .00* | 1.74 .20 | < 1 | 1.04 .32 | 7.05 .01* | 4.87 .03* | < 1 | < 1 | 13.94 .00* | 5.18 .03* |

*P*-values not given if $F < 1$. The asterisk (*) denotes significance at $p < .05$. See Table **2** for details of the critical regions.

The face effect (*i.e.*, face-picture priming) was significant by participants in the NP1 and the long region: Participants looked longer at the picture that was emotionally congruent with the face than at the other picture. There was no significant Face x Sentence interaction (*i.e.*, face-sentence priming). However, crucially the 3-way Face x Sentence x Age interaction (*i.e.*, face-sentence priming modulated by age) was fully reliable in the Adverb and in the long region, and in the NP1 region in the analysis by participants. This 3-way interaction is important and suggests that the effect of the face on sentence processing was modulated by age.

**Table 4. Results of planned pairwise-comparisons. Degrees of freedom: for participants ($t1$) =16, for items ($t2$) = 27: $p$-values in parentheses. The asterisk (*) denotes significance at $p < .05$. The symbol § indicates significance at $p < .05$ in the static-face study.**

| Region | Younger | | | | Older | | | |
|---|---|---|---|---|---|---|---|---|
| | Neg sent neg face *vs.* Neg sent pos face | | Pos sent neg face *vs.* Pos sent pos face | | Neg sent neg face *vs.* Neg sent pos face | | Pos sent neg face *vs.* Pos sent pos face | |
| | $t1 (p1)$ | $t2 (p2)$ | $t1 (p1)$ | $t2 (p2)$ | $t1 (p1)$ | $t2 (p2)$ | $t1 (p1)$ | $t2 (p2)$ |
| NP1 | 2.4 (.027)*§ | 1.4 (.14) | -.24 (.81) | -.62 (.54) | .44 (.67) | 1.16 (.25) | 3.3 (.005)* § | .73 (.47)§ |
| Adv | 2.2 (.04)*§ | 3.14 (.004)*§ | -1.3 (.22) | -.75 (.45) | -.14(.89) | -.5(.62) | 1.6 (.13)§ | 1.3 (.21) |
| Long | 2.2 (.04)*§ | 1.4 (.16)§ | -.62(.54) | -.15(.14) | .37 (.71) | .66 (.52) | 3.3 (.005)*§ | 2.4 (0.2)*§ |

Follow-up analyses presented in Table **4** clarify that the comparisons on younger participants yielded significances in the two negative sentence conditions - in the Adverb region by participants and items, and in the NP1 and long region by participants. By contrast, comparisons on older participants were significant only in the positive sentence conditions

in the long region, and in the NP1 region by participants. These results show that for the younger adults a negative (*vs.* positive) prime face significantly enhanced looks to a mentioned negative picture; a positive (*vs.* negative) face did not enhance the processing of a positive sentence. For the older group however, the opposite occurred: A negative face had no effect on the processing of a negative sentence, but a positive (*vs.* negative) face triggered more looks to the positive picture when it was mentioned.

Table **4** also marks (using "§") the *t*-tests that yielded significances in the static-face study. As one can verify, the dynamic-face results faithfully replicate those by C&K with static prime faces, and with only half the number of participants of the original study (16 *vs.* 32 participants per age group). Like the previous results, they show an asymmetry between older and younger participants in their integration of emotional face information into sentence processing. This asymmetry is consistent with a positivity effect, whereby older people respond more to positive faces and younger ones to negative faces.

### 4.1.4. Do We Replicate Early Positivity and Negativity Biases (in First Fixation Duration after NP1 Onset)?

In eye-tracking research on language processing, first-fixation duration is usually taken to index the very early stages of processing a word [54]. In the static-face study, we had found a significant Sentence x Age interaction in first fixation duration. For the younger participants, first fixation durations after the onset of NP1 were numerically longer when the sentence was negative than positive. By contrast, for the older adults, they were significantly longer when the sentence was positive than negative. This interaction suggests that for older people, positive (*vs.* negative) sentences act as a trigger to inspect the visual display longer and in more depth in first fixations (a positivity bias), while for younger participants the opposite held (a negativity bias).

In the dynamic-face study, a marginal Sentence x Age interaction emerged in the participants' ANOVA on the first fixation durations ($F1(1,30) = 3.16$, $p = .085$)[1]. Interestingly, the interaction pattern in duration of first fixation for dynamic faces is opposite that for static faces: For older adults first-fixation durations were longer with negative (*vs.* positive) sentences (274 *vs.* 249 ms), while for younger adults the opposite held (287 *vs.* 297). Planned pairwise comparisons on the means for each age group showed the difference to be significant for the older ($t1(15) = 2.24$, $p1 = .040$), but not the younger ($t1(14) = -.87$, $p1 = .40$) participants. Thus, with dynamic facial primes, we failed to replicate the positivity bias in the first-fixation durations at NP1. We return to this in the discussion.

### 4.1.5. Do the Post-sentence Responses Reveal Negativity and Positivity Biases (Reaction Times and Accuracy)?

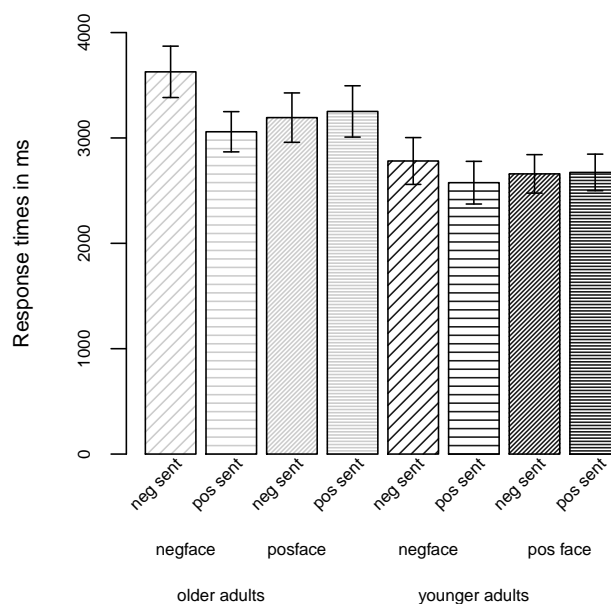The descriptive results for the RT data are presented in Fig. (**3**), and the corresponding ANOVA results in Table **5**.



**Fig. (3).** RTs by face and sentence valence for each age group (N=32, dynamic face study). Error bars indicate SE of the mean.

---

[1] The ANOVA by items was not performed, due to a high number of missing data points.

**Table 5. Summary of ANOVA results on RTs. Degrees of freedom: *F*1(1,30), *F*2 (1,27).**

| Effects and interactions | *F*1 *p*1 | *F*2 *p*2 |
|---|---|---|
| Age | 14.67 .001* § | 238.38 .000* § |
| Face | <1 § | 2.53 § .12 |
| Sentence | 10.72 .003* § | 14.48 .003* § |
| Face x sentence | 4.32 .046 * § | 33.95 .000* § |
| Age x face | < 1 | 1.9 .17 |
| Age x sentence | 1.32 .26 | 5.54 .026* § |
| Age x face x sentence | < 1 § | 5.23 .030* § |

*P*-values are under the corresponding *F*-values. *P*-values are not given if *F* < 1. The asterisk (*) denotes significance at $p < .05$. The symbol § indicates significance at $p < .05$ in the static-face study.

Overall, the RT results did not mirror the eye-tracking results, and this discrepancy replicated the pattern of results from the study with static faces (in Table **5** "§" indicates significant results with static faces). As Table **5** shows, the main effect of age was significant, with older participants displaying overall much longer RTs (3241 *vs*. 2656 ms). This was expected based on prior results (*e.g.*, [16, 55, 56]). Furthermore, the sentence effect was significant, with longer RTs for negative than positive sentences (3038 *vs*. 2858 ms). The sentence effect was qualified by an Age x Sentence interaction (significant by items). This interaction was due to older people's longer reaction times with negative (*vs*. positive) sentences, an effect that was reduced in younger adults (see Fig. **3**). We also observed a Face x Sentence interaction, which was *not* due to facilitation when prime and sentence were emotionally congruent (*vs*. non congruent); participants responded faster when face and sentence had opposite valence (see Fig. **3**). Furthermore, one can note in Fig. (**3**) that the negative congruent condition (the leftmost bar) triggers the longest RTs in both age groups. The Face x Sentence interaction was qualified by a 3-way interaction with Age, albeit significant by items only. Despite the overall similar pattern for young and old (see Fig. **3**), the 3-way interaction occurs because older adults experience greater difficulty than younger adults with the face-negative sentence condition.

The proportions of accurate answers as a function of face and sentence valence are illustrated in Fig. (**4**) for the two age groups. Overall, the accuracy results did not mirror the eye-tracking results, and this discrepancy replicated the results from the study with static faces.
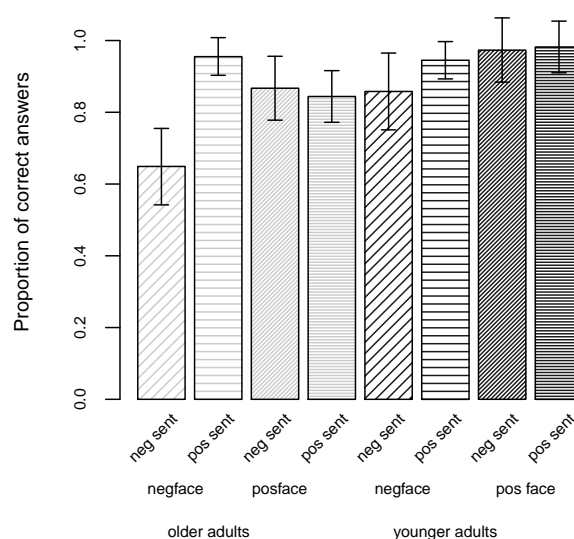


**Fig. (4).** Proportions of correct answers as a function of face, sentence and age (N=32, dynamic face study). Error bars represent the *SE* of the mean.

One can see in Fig. (**4**) that the particular difficulty observed in RTs with the negative congruent condition is also reflected in the accuracy scores; in fact, this condition is associated with the lowest score for both age groups. Table **6** presents the results of the corresponding logistic regression model.

**Table 6. Linear mixed-effect model results for accuracy scores. The asterisk (\*) denotes significance at *p* < .05. The symbol §
indicates significance at *p* < .05 in the static-face study.**

|  | *Coefficient* | *SE* | *z-Value* | *P* |
|---|---|---|---|---|
| *Intercept* | 3.08 | .26 | 11.81 | .000 * § |
| *Face valence* | .39 | .19 | 2.09 | .036 * § |
| *Sentence valence* | .52 | .18 | 2.88 | .003 * § |
| *Age* | .66 | .25 | 2.61 | .008 * § |
| *Face x Sent* | -.58 | .23 | -2.53 | .011 * |
| *Face x Age* | .40 | .18 | 2.25 | .024 * |
| *Sent x Age* | -0.14 | .18 | -0.79 | .431 |
| *Face x Sent x Age* | .36 | .23 | 1.57 | .117 |

As can be noted in Table 6, the three main effects (Face, Sentence and Age) were significant. Both the face and the sentence effects are due to higher accuracy for the positive face and positive sentence condition than their negative counterparts (see Fig. **4**). The expected age effect occurred because of older people's overall lower accuracy rate [55, 56]. Furthermore, both the Face x Age and the Face x Sentence interaction achieved significance. The Face x Sentence interaction was due to low scores in both groups for the negative congruent condition (negative face + negative sentence, see Fig. **4**). The Face x Age interaction occurred because the relative difficulty with negative (*vs*. positive) faces was greater for younger than older adults. This latter result is *not* consistent with the prediction of a positivity effect. In fact, according to this prediction, one would expect the opposite pattern, *i.e.*, older participants' accuracy to be more adversely affected by negative faces (compared to positive ones) than younger adults'.

In sum, when considering only the results for the dynamic-face study, the statistical analyses on the accuracy scores generally corroborated the findings from the RTs in that, for both groups, decisions involving negative sentences were particularly difficult, with the negative face-negative sentence condition appearing the most challenging. The analyses on the accuracy scores further revealed a general difficulty with negative (*vs*. positive) faces (independent of sentence valence), which was slightly more pronounced for younger adults (see Fig. **4**).

## 4.2. Does Cognitive Effort (As Measured by Pupil Dilation) Correlate with Eye-Gaze Positivity/Negativity Biases?

Fig. (**5**) illustrates the mean normalized pupil size by region for old Fig. (**5A**) and younger participants Fig. (**5B**). As one can see, the pattern looks similar, in that both ages display a bigger pupil size when they look at the negative rather than the positive picture. From the graphs it can also be seen that pupil size increases as participants process the words of the sentence incrementally.

The mixed regression models were run using the R software and the package lme4. The dependent variable was pupil size and the fixed effects were age (young *vs*. old), picture (negative *vs*. positive) and region (n1, n2, adverb, verb). To minimize collinearity, we used effect coding by transforming the fixed effects into a numerical value and centering them so as to have a mean of zero and a range of 1 [52]. Effect coding has the further advantage of allowing the coefficients of the regression to be interpreted as the main effects in a standard ANOVA. Following the recommendations in [57], the random part of the model (for participants and items) included the intercept and the slope of the fixed effects and their interaction (*i.e.*, the maximal model, according to [57]); note that Age was a between variable with respect to the random effects for participants, so the slope of Age or its interaction with the other predictors were not included in the random part of the model for participants as per [57]. The regression analysis yields coefficients, standard errors and *t*-values for each fixed effect and interaction. A coefficient was considered to be significant at alpha = 0.05 when the absolute value of *t* was greater than 2 [52].
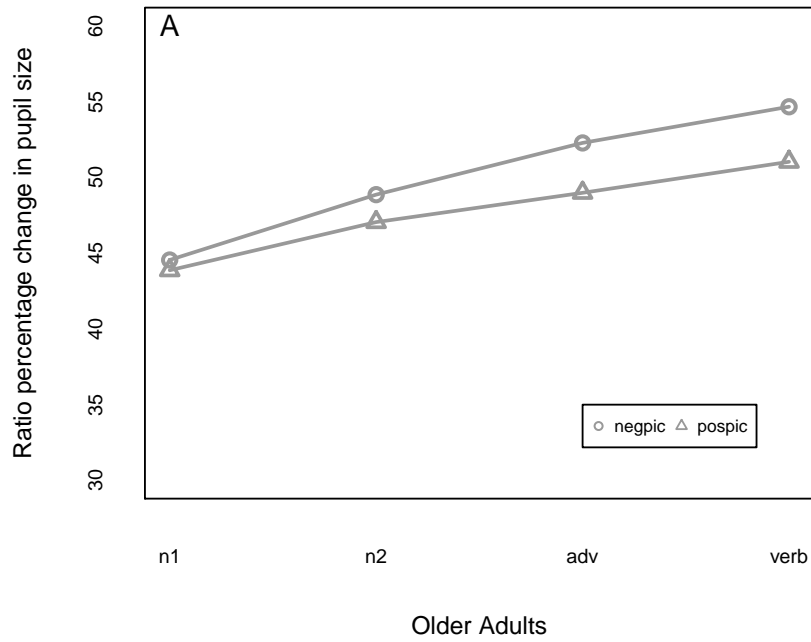
**Fig. (5A).** A Older adults' mean normalized pupil size by word region (N=48, data from both the static and dynamic face studies).
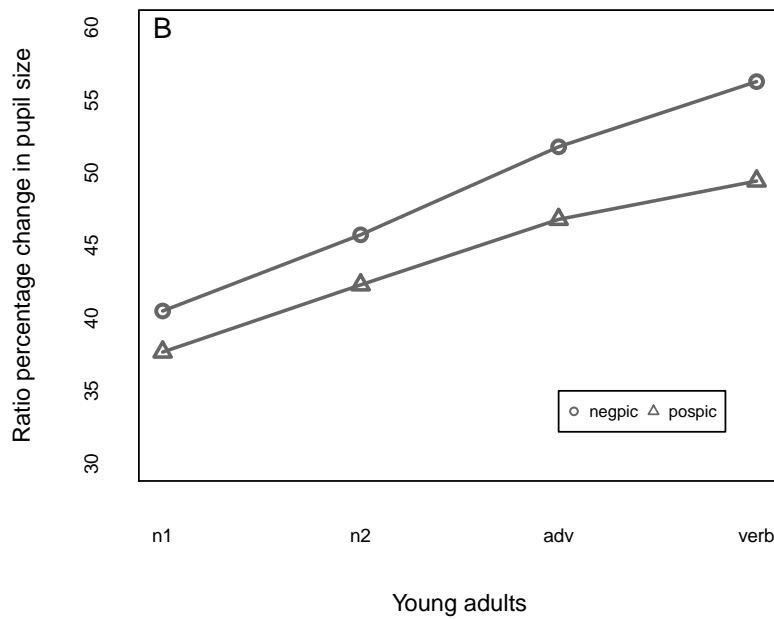


**Fig. (5B).** B Younger adults' mean normalized pupil size by word region (N=48, data from both the static and dynamic face studies).

The results of the model using the normalized mean pupil size as dependent variable are given in Table **7**.

**Table 7. Linear mixed effect model results with normalized mean pupil size as dependent variable.**

|  | *Coefficient* | *SE* | *t-value* |
|---|---|---|---|
| *Intercept* | .4757 | .011 | 41.60 |
| *Picture* | -.0166 | .004 | - 3.94 * |
| *Age* | -.0094 | .009 | -1.06 |
| *Region* | .0135 | .002 | 7.69 * |
| *Picture x Age* | -.0049 | .003 | -1.66 |
| *Picture x Region* | -.0020 | .002 | -1.14 |
| *Age x Region* | .0030 | .002 | 1.70 |
| *Picture x Age x Region* | -.0014 | .002 | - 0.83 |

*p < .05.

As one can see from Table **7**, the only effects that reached significance were Picture and Region. These effects, already hinted at in our earlier comments on Fig. (**5**), are due to the bigger pupil size associated with the negative *vs*. positive picture, and to the gradual increase in pupil size as people process the regions of the sentence. The Picture x Age interaction, which would support the hypothesis that cognitive effort is required with a positivity effect, was not significant. The regression model in which the dependent variable was the (non-normalized) absolute pupil size yielded essentially the same results - there was no significant Picture by Age interaction ($t = 1.60$). As expected, in this model the effect of Age, and the Age x Region interaction reached significance ($t = 2.43$). The interaction is due to older people' slower rate of increase in pupil size as they process the words of the sentence. The table of the linear mixed effect model with the non-normalized pupil size as dependent variable is included in the **Supplementary Material**.

## 5. DISCUSSION

The motivation for the present research was two-fold. First, we wanted to ascertain that previously-observed positivity effects reported in 'C&K' [16] with static facial primes were genuine, and not the result of using facial primes not powerful enough to trigger equally strong priming for younger and older adults with negative and positive faces. We were, in particular, keen to replicate these effects, given they had only been observed once before during incremental language processing with the visual world methodology. Replication appeared desirable given the recent debate among the scientific community calling for more stringent statistical validation of findings with an emphasis on the replicability of results (*e.g.*, [58, 59]). The replication attempt further permitted us to verify the 'discrepancy' found in the static face experiment between the results of the eye movement and the RTs and accuracy scores. In addition to replication goals, we tested a prediction derived from socio-emotional selectivity theory - that positivity preferences in the older adults implicate measurable cognitive effort.

### 5.1. Replicating Positivity Effects on Sentence Comprehension

Regarding replicability, the results of the static face study replicated with dynamic faces, in the eye movements, but also in RTs and accuracy. In C&K (see General Discussion) we highlighted the importance of this result for various areas of research. Its replication in the current study with a new and smaller sample size testifies to the robustness of the observed effects. It shows that emotional priming of sentence processing through facial expressions takes place during the incremental interpretation of the sentence, and is not delayed. This is remarkable since, for emotional priming to occur incrementally in our experiments, information from three sources - the face, the sentence and the pictures - must be rapidly integrated. Our results show that this integration is relatively easy, and, importantly, it occurs as fast for older as for younger adults. In fact, for both age groups, priming was observed from the first word region in which it was considered possible (the NP1 region).

Crucially, these results also confirm that emotional priming affects young and older adults' visual attention differently and in line with the predictions of a positivity effect (see socio-emotional selectivity theory): For older adults, positive prime faces enhanced fixations to the positive pictures but negative prime faces did not have such enhancing effect on negative pictures. On the other hand, the opposite occurred for younger adults. To our knowledge, in combination with the findings from the static face experiment, these results are the first to show positivity effects in eye movements during incremental sentence processing.

Positivity effects in visual attention alone have been observed in previous studies using eye tracking [18, 21, 22]. Isaacowitz DM *et al*., [21, 22] showed that, when presented with pairs of pictures consisting of a neutral face and a positive (happy) or negative (sad or angry) face, older people spent less time looking at the negative than positive face; in other words, they displayed an attentional bias away from the negative and towards the happy facial expressions. In the same studies, younger people, on the other hand, displayed a preference for negative faces. Our two visual world studies with static and dynamic prime faces provide additional evidence for a positivity effect in visual attention using a new and seemingly more complex task, *i.e.*, a task requiring the processing of emotional visual stimuli in combination with language comprehension and within an emotional priming paradigm.

Not only did the eye movement data replicate the positivity effects found with static faces, but also the pattern in the verification response data (RTs and accuracy) was similar in both experiments, with no clear positivity effects. In fact, in RTs and accuracy (unlike in the eye-movement data) we observed the same overall difficulty in the congruent negative condition for both young and old; furthermore, in the accuracy results of the current experiment, the younger adults appeared to have more difficulty with negative (*vs*. positive) faces than the older adults, *i.e.* the opposite of a positivity effect.

Interestingly, the *only* measure that yielded discrepant results compared with the static face study is the first fixation duration analyses, an early measure that reflects visual attention in a very short time window, *i.e.*, the first 200-300 ms of processing the word that disambiguated reference towards one or the other picture. In this measure, older people displayed longer fixation durations for negative (*vs*. positive) sentences with dynamic faces, while with static faces the opposite occurred (for younger adults no statistical differences emerged in either experiment). A possible explanation for this is that older adults may have found the dynamic-face videos relatively complex to process (they contained in reality two facial expressions: the neutral expression at the beginning of the video clip and the final apex positive or negative expression). During the experiment participants also viewed steady-state videos of neutral facial expressions (*i.e.*, without change, in the filler items), for which they also had to provide a face-sentence verification response. Compared to neutral face videos, the dynamic neutral-to-positive or neutral-to-negative videos may have had an initial distracting effect on older adults as they involved identifying two different facial expressions, creating effectively a situation of divided attention. There is some evidence that when attention is divided, the positivity effect in attention is reversed [28]. The dynamic faces may thus have contributed to the unexpected reversal of the first-fixation-duration pattern in the present experiment. However, this was the only instance in which the results of the dynamic-face study differed from the static-face study, and, importantly, this result did not affect the remaining fixations in the NP1 (and the other) regions. In fact, both the face-picture priming effect and the Face x Sentence x Age interaction in the NP1 region replicated with dynamic faces.

Two further conclusions are in order. First, because the pattern of the verification responses did not substantially change from static to dynamic faces, we conclude that, against our expectations, the use of dynamic face primes did not result in better recognition of the facial expressions. Why then did we observe slower responses for negative than positive valence independent of age? In the verification task, participants had to decide by pressing a yes-no button whether 'the face matched the sentence'. To arrive at this decision an appraisal of the valence of the sentence, and not only of the face, was required. Perhaps then participants did not judge (some or all of) the negative sentences to be unambiguously negative, slowing RTs and increasing errors in the negative face-negative sentence condition. We think this is unlikely given the care with which we constructed the sentences (see C&K, Construction of sentences). We cannot exclude that the negative sentences were more difficult to understand than their positive counterpart, although this alone cannot accommodate the verification results pattern. A more likely explanation (also given the task) is that participants assessed the valence of the sentence *in combination* with the IAPS picture which the sentence described. In selecting the negative IAPS pictures for the study, we had avoided pictures with valence ratings towards the extreme negative end, as we feared these pictures might be too unpleasant to show to participants (on the other hand, positive pictures were nearer to the extreme positive end of the scale, see C&K, Materials section, IAPS pictures). It is possible that in some cases the combination of a negative sentence with a not-so negative picture may have generated confusion about its alleged negative valence, causing slower RTs and more errors compared with the positive face-sentence pairs.

A final question concerns the mismatch between the eye tracking and the verification response time results, which we replicated with dynamic faces. In C&K we suggested that the mismatch is probably due to eye movements and verification responses reflecting different stages of processing, with the former tapping initial processing of the sentence in relation to the face and the pictures, and the verification results reflecting in addition post-comprehension decision processes. While it is certainly possible for the outcomes of verification and comprehension processes to be consistent with each other (see, for *e.g.* [60, 61], we can expect cases in which complete agreement does not obtain [62]. In C&K (see General Discussion), we also touched on the implication of this explanation for the nature of the mechanisms underlying the positivity effects observed in our eye tracking data, and suggested that the effects are more likely to result from initial, automatic processing of emotional information, rather than later, controlled processing. Because these incremental effects replicated in the second study, this implication appears to remain valid.

## 5.2. Pupil Size and Cognitive Effort

Concerning the claim that positivity preferences in the older adults implicate measurable cognitive effort, as reflected in pupil size, the present results provide no evidence for this claim: Both young and older adults' pupil size displayed essentially the same pattern, with an increase when looking at negative relative to positive pictures. Interestingly, these results mirror those of Allard ES *et al*., [31], who found that for *both* younger and older people pupil size was larger when fixating on sad and fearful faces compared to happy ones.

One methodological question relates to the assumption that pupil dilation reflects cognitive effort or cognitive load of some sort. This was the initial assumption behind our analysis of pupil size and also behind the analyses in [31]. That

pupil dilation is positively correlated with cognitive load has been demonstrated over the years in many experimental studies performed on young adults (*e.g.* [51, 63].). A number of studies have also specifically investigated pupil size during language processing, typically in association with the processing of complex sentences, and found a corresponding increase in pupil diameter with these sentences [45, 46, 64] but again the participants in these studies were young people.

More recently, in a between-age experimental design, Piquado *et al.*, [47] have shown that the pupils of older adults, too, increase in size as a function of cognitive load. In this study, Piquado and colleagues tested younger and older adults on two tasks which have often been used in the past to establish an association between pupil size and cognitive effort, *i.e.*, the recall of a list of digits [51] and of complex sentences [45]. Piquado *et al.*, [47] found that, in digit recall, younger and older adults showed a similar progressive increase in pupil size as a function of the size of the memory load. For sentences, younger people's pupil dilation increased as a function of both sentence length and syntactic complexity. Older people's pupil size, by contrast, increased only as a function of sentence length. Thus, older adults' pupils do appear to respond in the same way as younger adults' (*i.e.*, by increasing in diameter size) to some tasks which are commonly assumed to involve a certain cognitive effort. In the case of our experiments, younger and older adults also showed the same trend in pupil dilation; however, our hypothesis was that older adults alone would show evidence of cognitive effort. One possible conclusion is that no cognitive effort (or minimal cognitive effort) was involved when older adults displayed positivity effects in gazing (see [31]).

Our results, together with those by Allard ES *et al.*, [31], suggest that the assumption that fixation preferences should be positively correlated with pupil size changes is too simplistic. Both our results and those of [31] revealed distinct fixation preferences for young compared with older adults (in line with positivity effects) but this pattern was not mirrored in age-related pupil size differences. Although it seems plausible to assume that fixation preferences and fixation duration should correlate positively with pupil diameter, most research aimed at demonstrating that pupil size reflects cognitive effort has looked only at pupil size, and has not directly investigated the relationship between pupil size and looking or reading times [45 - 47].

We know of only two studies that have, among others, considered this relationship. Jackson I *et al.*, [65] eye-tracked 8-month-old infants looking at physically impossible (one train entered, but another train exited the tunnel) *vs.* more conventional events (*e.g.*, the same train entering and exiting a tunnel). Pupil diameter increased for the impossible event, but, to the authors' surprise, no correlation was observed between looking duration and peak pupil diameter. Duque A *et al.*, [66], on the other hand, did find a positive correlation between gaze duration and pupil dilation in some conditions when testing a group of dysphoric and non-dysphoric young adults while they processed emotional faces (sad, happy, angry). Because of these conflicting results and the fact that these two studies investigated special populations, it is difficult to come to firm conclusion as to why the present positivity effects in gaze (determined on the basis of longer looking times in one condition compared to another) were not associated with increased pupil diameter. Clearly, the assumption that patterns of fixation preferences/times should map onto corresponding patterns of pupil size needs further investigation.

If we leave aside the relation between looking times and pupil size and focus only on pupil size, our result of a greater pupil size with negative *vs.* positive pictures for both age groups suggests a trigger for pupil dilation of a more general, emotional nature, *i.e.* arousal. In past pupillometry research, arousal as a result of processing emotional stimuli has been argued to be responsible for pupil size increase (*e.g.* [67 - 69]). Although the positive and negative IAPS pictures had been controlled for arousal (see C&K, Methods section), perhaps the negative picture and sentence together, rendered the negative items more arousing than the positive ones (another theoretical possibility is that negative sentences were more complex or longer than the positive ones; however this was not the case for our items, because negative sentences had the same grammatical structure as their positive counterpart and the same number of syllables).

## CONCLUSION

Positivity gaze preferences seem replicable, emerge incrementally, and do not appear to covary with measurable cognitive effort as indexed by pupil size increases. Instead, the pupil size of both young and older participants increased for negative *vs.* positive pictures. We have discussed some important methodological and conceptual implications of these findings that should be taken into consideration in future pupillometric studies.

## SUPPLEMENTARY MATERIAL

Supplementary material is available on the publishers Website along with the published article.

## CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

## ACKNOWLEDGEMENTS

## REFERENCES

[1]     Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. Science 1995; 268(5217): 1632-4.
[http://dx.doi.org/10.1126/science.7777863] [PMID: 7777863]

[2]     Knoeferle P, Crocker MW, Scheepers C, Pickering MJ. The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. Cognition 2005; 95(1): 95-127.
[http://dx.doi.org/10.1016/j.cognition.2004.03.002] [PMID: 15629475]

[3]     Sedivy JC, Tanenhaus MK, Chambers CG, Carlson GN. Achieving incremental semantic interpretation through contextual representation. Cognition 1999; 71(2): 109-47.
[http://dx.doi.org/10.1016/S0010-0277(99)00025-6] [PMID: 10444906]

[4]     Dahan D, Tanenhaus MK. Looking at the rope when looking for the snake: conceptually mediated eye movements during spoken-word recognition. Psychon Bull Rev 2005; 12(3): 453-9.
[http://dx.doi.org/10.3758/BF03193787] [PMID: 16235628]

[5]     Knoeferle P, Carminati MN, Abashidze D, Essig K. Preferential inspection of recent real-world events over future events: evidence from eye tracking during spoken sentence comprehension. Front Psychol 2011; 2: 376.
[http://dx.doi.org/10.3389/fpsyg.2011.00376] [PMID: 22207858]

[6]     Chambers CG, Tanenhaus MK, Magnuson JS. Actions and affordances in syntactic ambiguity resolution. J Exp Psychol Learn Mem Cogn 2004; 30(3): 687-96.
[http://dx.doi.org/10.1037/0278-7393.30.3.687] [PMID: 15099136]

[7]     Altmann GT. Language-mediated eye movements in the absence of a visual world: the blank screen paradigm. Cognition 2004; 93(2): B79-87.
[http://dx.doi.org/10.1016/j.cognition.2004.02.005] [PMID: 15147941]

[8]     Huettig F, Rommers J, Meyer AS. Using the visual world paradigm to study language processing: a review and critical evaluation. Acta Psychol (Amst) 2011; 137(2): 151-71.
[http://dx.doi.org/10.1016/j.actpsy.2010.11.003] [PMID: 21288498]

[9]     Richardson DC, Dale R. Looking to understand: the coupling between speakers and listeners eye movements and its relationship to discourse comprehension. Cogn Sci 2005; 29(6): 1045-60.
[http://dx.doi.org/10.1207/s15516709cog0000_29] [PMID: 21702802]

[10]    Hanna J, Brennan S. Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. J Mem Lang 2007; 57: 596-615.
[http://dx.doi.org/10.1016/j.jml.2007.01.008]

[11]    Richardson DC, Dale R, Kirkham NZ. The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. Psychol Sci 2007; 18(5): 407-13.
[http://dx.doi.org/10.1111/j.1467-9280.2007.01914.x] [PMID: 17576280]

[12]    Kreysa H, Knoeferle P. Peripheral speaker gaze facilitates spoken language comprehension: syntactic structuring and thematic role assignment in German. Proceedings of the European Conference on Cognitive Science. In: New Bulgarian University Press; Sofia 2011.

[13]    Mishra RK, Singh N, Pandey A, Huettig F. Spoken language-mediated anticipatory eye-movements are modulated by reading ability-Evidence from Indian low and high literates. J Eye Mov Res 2012; 5: 3-10.

[14]    Belke E, Meyer AS. Single and multiple object naming in healthy aging. Lang Cogn Process 2007; 22: 1178-211.
[http://dx.doi.org/10.1080/01690960701461541]

[15]    Zhang L, Knoeferle P. Visual Context Effects on Thematic Role Assignment in Children versus Adults: Evidence from Eye Tracking in German. In: Miyake N, Peebles D, Cooper RP, Eds. Proceedings of the Annual Meeting of the Cognitive Science Society. In: Cognitive Science Society; Austin, TX. 2012; pp. 2593-98.

[16]    Carminati MN, Knoeferle P. Effects of speaker emotional facial expression and listener age on incremental sentence processing. PLoS One 2013; 8(9): e72559.

[http://dx.doi.org/10.1371/journal.pone.0072559] [PMID: 24039781]

[17]   Wieser MJ, Brosch T. Faces in context: a review and systematization of contextual influences on affective face processing. Front Psychol 2012; 3: 471.
[http://dx.doi.org/10.3389/fpsyg.2012.00471] [PMID: 23130011]

[18]   Mather M, Carstensen LL. Aging and attentional biases for emotional faces. Psychol Sci 2003; 14(5): 409-15.
[http://dx.doi.org/10.1111/1467-9280.01455] [PMID: 12930469]

[19]   Scheibe S, Carstensen LL. Emotional aging: recent findings and future trends. J Gerontol B Psychol Sci Soc Sci 2010; 65B(2): 135-44.
[http://dx.doi.org/10.1093/geronb/gbp132] [PMID: 20054013]

[20]   Reed AE, Carstensen LL. The theory behind the age-related positivity effect. Front Psychol 2012; 3: 339.
[http://dx.doi.org/10.3389/fpsyg.2012.00339] [PMID: 23060825]

[21]   Isaacowitz DM, Wadlinger HA, Goren D, Wilson HR. Is there an age-related positivity effect in visual attention? A comparison of two methodologies. Emotion 2006; 6(3): 511-6.
[http://dx.doi.org/10.1037/1528-3542.6.3.511] [PMID: 16938091]

[22]   Isaacowitz DM, Wadlinger HA, Goren D, Wilson HR. Selective preference in visual fixation away from negative images in old age? An eye-tracking study. Psychol Aging 2006; 21(1): 40-8.
[http://dx.doi.org/10.1037/0882-7974.21.1.40] [PMID: 16594790]

[23]   Lang PJ, Bradley MM, Cuthbert BN. International Affective Picture System: Instruction manual and affective ratings. Technical Report A-4 1999. The Center for Research in Psychophysiology, University of Florida.

[24]   Langeslag SJ, van Strien JW. Aging and emotional memory: the co-occurrence of neurophysiological and behavioral positivity effects. Emotion 2009; 9(3): 369-77.
[http://dx.doi.org/10.1037/a0015356] [PMID: 19485614]

[25]   Arai M, van Gompel RP, Scheepers C. Priming ditransitive structures in comprehension. Cognit Psychol 2007; 54(3): 218-50.
[http://dx.doi.org/10.1016/j.cogpsych.2006.07.001] [PMID: 16973144]

[26]   Carminati MN, van Gompel RP, Scheepers C, Arai M. Syntactic priming in comprehension: the role of argument order and animacy. J Exp Psychol Learn Mem Cogn 2008; 34(5): 1098-110.
[http://dx.doi.org/10.1037/a0012795] [PMID: 18763894]

[27]   Mather M, Knight M. Goal-directed memory: the role of cognitive control in older adults emotional memory. Psychol Aging 2005; 20(4): 554-70.
[http://dx.doi.org/10.1037/0882-7974.20.4.554] [PMID: 16420131]

[28]   Knight M, Seymour TL, Gaunt JT, Baker C, Nesmith K, Mather M. Aging and goal-directed emotional attention: distraction reverses emotional biases. Emotion 2007; 7(4): 705-14.
[http://dx.doi.org/10.1037/1528-3542.7.4.705] [PMID: 18039037]

[29]   Thomas RC, Hasher L. The influence of emotional valence on age differences in early processing and memory. Psychol Aging 2006; 21(4): 821-5.
[http://dx.doi.org/10.1037/0882-7974.21.4.821] [PMID: 17201502]

[30]   Rösler A, Ulrich C, Billino J, *et al.* Effects of arousing emotional scenes on the distribution of visuospatial attention: changes with aging and early subcortical vascular dementia. J Neurol Sci 2005; 229-230: 109-16.
[http://dx.doi.org/10.1016/j.jns.2004.11.007] [PMID: 15760628]

[31]   Allard ES, Wadlinger HA, Isaacowitz DM. Positive gaze preferences in older adults: assessing the role of cognitive effort with pupil dilation. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2010; 17(3): 296-311.
[http://dx.doi.org/10.1080/13825580903265681] [PMID: 19890752]

[32]   Kamachi M, Bruce V, Mukaida S, Gyoba J, Yoshikawa S, Akamatsu S. Dynamic properties influence the perception of facial expressions. Perception 2001; 30(7): 875-87.
[http://dx.doi.org/10.1068/p3131] [PMID: 11515959]

[33]   Sato W, Kochiyama T, Yoshikawa S, Naito E, Matsumura M. Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. Brain Res Cogn Brain Res 2004; 20(1): 81-91.
[http://dx.doi.org/10.1016/j.cogbrainres.2004.01.008] [PMID: 15130592]

[34]   Arsalidou M, Morris D, Taylor MJ. Converging evidence for the advantage of dynamic facial expressions. Brain Topogr 2011; 24(2): 149-63.
[http://dx.doi.org/10.1007/s10548-011-0171-4] [PMID: 21350872]

[35]   LaBar KS, Crupain MJ, Voyvodic JT, McCarthy G. Dynamic perception of facial affect and identity in the human brain. Cereb Cortex 2003; 13(10): 1023-33.
[http://dx.doi.org/10.1093/cercor/13.10.1023] [PMID: 12967919]

[36]   Recio G, Sommer W, Schacht A. Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. Brain Res 2011; 1376: 66-75.
[http://dx.doi.org/10.1016/j.brainres.2010.12.041] [PMID: 21172314]

[37]   Sato W, Yoshikawa S. The dynamic aspects of emotional facial expressions. Cogn Emotion 2004; 18: 701-10.
[http://dx.doi.org/10.1080/02699930341000176]

[38]     Trautmann SA, Fehr T, Herrmann M. Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. Brain Res 2009; 1284: 100-15.
[http://dx.doi.org/10.1016/j.brainres.2009.05.075] [PMID: 19501062]

[39]     Rubenstein AJ. Variation in perceived attractiveness: differences between dynamic and static faces. Psychol Sci 2005; 16(10): 759-62.
[http://dx.doi.org/10.1111/j.1467-9280.2005.01610.x] [PMID: 16181436]

[40]     Sato W, Yoshikawa S. Enhanced experience of emotional arousal in response to dynamic facial expressions. J Nonverbal Behav 2007; 31: 119-35.
[http://dx.doi.org/10.1007/s10919-007-0025-7]

[41]     Ambadar Z, Schooler JW, Cohn JF. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. Psychol Sci 2005; 16(5): 403-10.
[http://dx.doi.org/10.1111/j.0956-7976.2005.01548.x] [PMID: 15869701]

[42]     Harwood NK, Hall LJ, Shinkfield AJ. Recognition of facial emotional expressions from moving and static displays by individuals with mental retardation. Am J Ment Retard 1999; 104(3): 270-8.
[http://dx.doi.org/10.1352/0895-8017(1999)104<0270:ROFEEF>2.0.CO;2] [PMID: 10349468]

[43]     Fiorentini C, Viviani P. Is there a dynamic advantage for facial expressions? J Vis 2011; 11(3): 1-15.
[http://dx.doi.org/10.1167/11.3.17] [PMID: 21427208]

[44]     Krendl AC, Ambady N. Older adults decoding of emotions: role of dynamic *versus* static cues and age-related cognitive decline. Psychol Aging 2010; 25(4): 788-93.
[http://dx.doi.org/10.1037/a0020607] [PMID: 21186915]

[45]     Just MA, Carpenter PA. The intensity dimension of thought: pupillometric indices of sentence processing. Can J Exp Psychol 1993; 47(2): 310-39.
[http://dx.doi.org/10.1037/h0078820] [PMID: 8364533]

[46]     Engelhardt PE, Ferreira F, Patsenko EG. Pupillometry reveals processing load during spoken language comprehension. Q J Exp Psychol (Hove) 2010; 63(4): 639-45.
[http://dx.doi.org/10.1080/17470210903469864] [PMID: 20017058]

[47]     Piquado T, Isaacowitz D, Wingfield A. Pupillometry as a measure of cognitive effort in younger and older adults. Psychophysiology 2010; 47(3): 560-9.
[http://dx.doi.org/10.1111/j.1469-8986.2009.00947.x] [PMID: 20070575]

[48]     Jaeger TF. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. J Mem Lang 2008; 59(4): 434-46.
[http://dx.doi.org/10.1016/j.jml.2007.11.007] [PMID: 19884961]

[49]     Baayen RH. Analyzing Linguistic Data A Practical Introduction to Statistics Using R. Cambridge: Cambridge University Press 2011.

[50]     Barr DJ. Analyzing 'visual world' eyetracking data using multilevel logistic regression. J Mem Lang 2008; 49(4): 457-74.
[http://dx.doi.org/10.1016/j.jml.2007.09.002]

[51]     Kahneman D, Beatty J. Pupil diameter and load on memory. Science 1966; 154(3756): 1583-5.
[http://dx.doi.org/10.1126/science.154.3756.1583] [PMID: 5924930]

[52]     Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. J Mem Lang 2008; 59: 390-412.
[http://dx.doi.org/10.1016/j.jml.2007.12.005]

[53]     Clark HH. The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. J Verbal Learn Verbal Behav 1973; 12: 335-59.
[http://dx.doi.org/10.1016/S0022-5371(73)80014-3]

[54]     Inhoff AW, Radach R. Definition and computation of oculomotor measures in the study of cognitive processes. In: Underwood G, Ed. Eye guidance in reading and scene perception. Oxford: Elsevier 1998; pp. 29-54.
[http://dx.doi.org/10.1016/B978-008043361-5/50003-1]

[55]     Jensen AR. Clocking the mind: mental chronometry and individual differences. Oxford: Elsevier 2006.

[56]     Salthouse TA. Selective review of cognitive aging. J Int Neuropsychol Soc 2010; 16(5): 754-60.
[http://dx.doi.org/10.1017/S1355617710000706] [PMID: 20673381]

[57]     Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. J Mem Lang 2013; 68(3): 255-78.
[http://dx.doi.org/10.1016/j.jml.2012.11.001] [PMID: 24403724]

[58]     Asendorpt JB, Conner M, De Fruyt F, De Houwer J, Denissen JJ, Fiedler K. Recommendations for increasing replicability in psychology. Eur J Pers 2013; 27: 108-19.
[http://dx.doi.org/10.1002/per.1919]

[59]     Pashler H, Harris CR. Is the replicability crisis overblown? Three arguments examined. Perspect Psychol Sci 2012; 7(6): 531-6.
[http://dx.doi.org/10.1177/1745691612463401] [PMID: 26168109]

[60]    Altmann GT, Kamide Y. Incremental interpretation at verbs: restricting the domain of subsequent reference. Cognition 1999; 73(3): 247-64.
        [http://dx.doi.org/10.1016/S0010-0277(99)00059-1] [PMID: 10585516]

[61]    Singer M. Verification of text ideas during reading. J Mem Lang 2006; 54: 574-91.
        [http://dx.doi.org/10.1016/j.jml.2005.11.003]

[62]    Tanenhaus MK, Carroll JM, Bever TG. Sentence-picture verification models as theories of sentence comprehension: a critique of Carpenter
        and Just. Psychol Rev 1976; 83: 310-7.
        [http://dx.doi.org/10.1037/0033-295X.83.4.310]

[63]    Verney SP, Granholm E, Marshall SP. Pupillary responses on the visual backward masking task reflect general cognitive ability. Int J
        Psychophysiol 2004; 52(1): 23-36.
        [http://dx.doi.org/10.1016/j.ijpsycho.2003.12.003] [PMID: 15003370]

[64]    Hyönä J, Tommola J, Alaja AM. Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. Q J
        Exp Psychol A 1995; 48(3): 598-612.
        [http://dx.doi.org/10.1080/14640749508401407] [PMID: 7568993]

[65]    Jackson I, Sirois S. Infant cognition: going full factorial with pupil dilation. Dev Sci 2009; 12(4): 670-9.
        [http://dx.doi.org/10.1111/j.1467-7687.2008.00805.x] [PMID: 19635092]

[66]    Duque A, Sanchez A, Vazquez C. Gaze-fixation and pupil dilation in the processing of emotional faces: The role of rumination. Cogn
        Emotion 2014; 1-20.

[67]    Hess EH, Polt JM. Pupil size as related to interest value of visual stimuli. Science 1960; 132(3423): 349-50.
        [http://dx.doi.org/10.1126/science.132.3423.349] [PMID: 14401489]

[68]    Libby WL Jr, Lacey BC, Lacey JI. Pupillary and cardiac activity during visual attention. Psychophysiology 1973; 10(3): 270-94.
        [http://dx.doi.org/10.1111/j.1469-8986.1973.tb00526.x] [PMID: 4702521]

[69]    Bradley MM, Miccoli L, Escrig MA, Lang PJ. The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology
        2008; 45(4): 602-7.
        [http://dx.doi.org/10.1111/j.1469-8986.2008.00654.x] [PMID: 18282202]

[70]    Hautzinger M, Keller F, Kuhner Ch. Beck Depression Inventer. 2nd ed. Frankfurt am Main: Pearson Assessment & Information GmbH 2009.

[71]    Wechsler D. Wechsler Adult Intelligence Scale Revised Manual. New York: The Psychological Corporation 1981.

[72]    Mayer JD, Gaschke YN. The experience and meta-experience of mood. J Pers Soc Psychol 1988; 55(1): 102-11.
        [http://dx.doi.org/10.1037/0022-3514.55.1.102] [PMID: 3418484]