Aging, effort, and stereotyping: The evidence for the moderating role of self-involvement

Gabriela Czarnek, Małgorzata Kossowska, & Michael Richter

Word count (excluding abstract, references, and supplemental materials): 7,673 words

Author note

Gabriela Czarnek and Małgorzata Kossowska, Institute of Psychology, Jagiellonian University, Ingardena Street 6, 31-006 Kraków, Poland.

Michael Richter, School of Natural Sciences and Psychology, Liverpool John Moores University, Tom Reilly Building, Byrom Street, Liverpool, L3 3AF, UK.

This work was supported by the National Science Centre, Poland [MAESTRO grant no. DEC - 2011/02/A/HS6/00155] received by Małgorzata Kossowska. The funding body did not interfere with the preparation of this manuscript at any stage.

The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Correspondence concerning this article should be addressed to Gabriela Czarnek, gabriela.czarnek@uj.edu.pl.

Abstract

A study with young and older adults (*N*=91) investigated the effect of self-involvement on stereotyping tendency and effort mobilization. We hypothesized that the impact of self-involvement varies as a function of age: increased self-involvement should lead older adults to engage in more effortful information processing and decreased stereotyping, whereas increased self-involvement should have no impact on effort mobilization and stereotyping tendency in young adults. Young and older adults read narratives under low and high-self-involvement conditions before performing a recognition test that measured their stereotyping tendency. Effort was assessed as cardiovascular responses. We found that older adults in the high-self-involvement condition presented low stereotyping tendency (similar to that of young people) in comparison to older adults in the low-self-involvement condition. Furthermore, older adults in the high-self-involvement condition had decreased high-frequency heart rate variability in comparison to the other conditions, but only during the recognition test; this suggests increased effort mobilization. These findings suggest that self-involvement decreases older adults' stereotyping tendency, possibly through effort mobilization.

Keywords: aging; stereotyping; effort; self-involvement; motivational intensity theory; selective engagement theory

1. Introduction

This research aim is to examine whether the relationship between age and stereotyping might be modified through self-involvement and the associated increase in effort mobilization. Previous studies have shown that older adults are more prone than young adults to exhibiting stereotyping towards social groups (e.g., Firebaugh & Davis, 1988; Gonsalkorale, Sherman, & Klauer, 2009; Herek, 1994; Wilson, 1996). We suggest that these age-related changes in stereotyping tendency can be prevented by increased self-involvement. In particular, we posit that the effortful, piecemeal analysis of incoming information that is fostered by high self-involvement results in decreased stereotyping tendency among older adults.

The novelty of the presented study is threefold. First, the impact of motivation (self-involvement) on the relationship between aging and stereotyping has hardly been tested. Second, the study tested for the first time the predictions of selective engagement theory (Hess, 2014) and motivational intensity theory (Brehm & Self, 1989; Wright, 1996) in a stereotyping task. Third, it provides a test of the assumption that stereotyping results from effortless (vs. effortful) information processing (e.g., Fiske, Lin, Neuberg, 1999; Wight & Kirby, 2001). The presented research could thus advance theories of motivated social cognition and models of stereotyping in older age.

1.1. Stereotyping in Older Age

Older adults are less able to inhibit stereotypical associations and prejudiced reactions than young people (Henry, von Hippel, & Baynes, 2009; Stewart, von Hippel, & Radvansky, 2009; von Hippel, Silver, & Lynch, 2000). Research on the self-regulatory deficit in stereotyping in older age showed that stereotypes impact the inferences that people make: older adults were more prone than younger adults to making and remembering stereotypical inferences while reading a short story (Radvansky, Copeland, & von Hippel, 2010). This is because stereotypes – as a form of well-established schemata stored in long-term memory – color judgments and inferences (Dunning & Sherman, 1997). The task that Radvansky and colleagues (2010) used consists of two parts: in the first (a story period), a participant reads a short narrative; in the second (a test period), a participant is presented with a recognition test in which the task is to decide whether a presented sentence had appeared in the story they had read. Older adults, in comparison to young people, were more likely accept statements that were new but consistent with the stereotypic inferences they made while reading ambiguous information. Radvansky and colleagues (2010) demonstrated that while both young and older adults activated stereotypes at early stages of information processing, older adults were less likely than young adults to inhibit them during reading. Instead, the stereotypical information was integrated in their understanding of a story. Thus, the lack of stereotype inhibition when encoding explained the increased stereotyping tendency in a subsequent memory test. These results are in line with the suggestion that stereotypes impact information at early stages of information processing (e.g., von Hippel, Sekaquaptewa, & Vargas, 1995).

The majority of the research on aging and stereotyping suggests that the tendency of older adults to stereotype is a consequence of age-related cognitive decline (that is, the reduction in cognitive resources that are available for information processing). However, the life-span perspective suggests that age-related cognitive losses might be counteracted (e.g., Baltes, 1987). Thus, an intriguing question is if it is possible for older adults to compensate for the age-related decline in cognitive performance that results in stereotyping.

According to Hess (1994, 1999), older adults' representations of other people are more biased than those of young adults; this is not only due to a reduction in the available cognitive resources, but also because of the selective allocation of these resources. Hess and colleagues demonstrated that cognitive decline-related biases in person perception could be compensated by higher resource allocation to tasks with high self-involvement (Hess, Follett, & McGee, 1998, Study 1; Hess, Germain, Swaim, & Osowski, 2009; Hess, Rosenberg, & Waters, 2001). Older adults are also able to regulate their attitudes toward stigmatized groups but need to invest more effort than young adults to overcome these biases (Krendl, Heatherton, & Kensinger, 2009; Krendl & Kensinger, 2016). Furthermore, our own research (Czarnek, Kossowska, & Sędek, 2015) showed that older adults are capable of inhibiting this stereotyping tendency under certain conditions. We replicated the aforementioned results of Radvansky and colleagues (2010), showing that older adults are indeed more prone to incorporating stereotypical inferences in their thinking; however, we also demonstrated that the relationship between age and stereotyping was moderated by self-involvement: when older adults read a self-involving story about their peers, their stereotyping tendency was low and similar to that of young adults. In contrast, when older adults read about people of other ages and the story was therefore less self-relevant to them, they showed a higher stereotyping tendency than young adults. We hypothesized that the mechanisms underlying this effect were the differences in effort mobilization triggered by different levels of self-involvement induced by the stories. However, this mechanism remains to be tested and the current research project aims to provide more direct evidence that effort mobilization moderates the relationship between age and stereotyping.

1.2. Effort Mobilization

Effort is defined as mobilization of resources to carry out instrumental behavior (Gendolla & Wright, 2009). According to motivational intensity theory, people invest only as much effort as needed to successfully perform a task; therefore, effort is directly determined by task difficulty (Brehm & Self, 1989; Wright, 1996). However, effort is determined not only by the objective difficulty of a task but also by its subjective perception (Wright, 1998). Individuals who perceive that they have limited abilities in a domain perceive a task as being more difficult and consequently invest more effort. In contrast, individuals who perceive themselves as having high ability perceive a task as easier and invest less effort. The proportional relationship between task difficulty and effort investment is, however, limited. The upper limit of effort mobilization is defined by potential motivation: a person mobilizes effort as long as she or he believes that success is possible and the effort required for success is justified. Potential motivation is affected by, for example, rewards (e.g., Richter, 2012) or personality traits (e.g., Richter, Baeriswyl, & Roets, 2012; Szumowska, Szwed, Kossowska, & Wright, 2017), but also by abstract goals such as personal interests, improving or maintaining self-esteem, or aiming for self-definition, namely self-involvement (Gendolla, 2004; see Gendolla & Richter, 2010 for an overview).

Research on the active coping hypothesis (Obrist, 1981), motivational intensity theory (Brehm & Self, 1989; Wright, 1996), the model of autonomic space (Berntson, Cacioppo, & Quigley, 1991), and studies on physical effort (Rowell, 1986; White & Raven, 2014) have demonstrated that effort is reflected in the activity of the autonomic nervous system (ANS) on the heart. Native heart activity, which is generated spontaneously by the sinoatrial node, is modified by ANS (Klabunde, 2011). The parasympathetic branch of the ANS (PNS) operates as a brake on the heart, i.e., it decreases its activity, whereas the sympathetic branch (SNS) acts as an accelerator, i.e., it increases its activity. At rest, the heart is mainly under an inhibiting parasympathetic influence. To support the metabolic requirements of bodily activity, the ANS decreases PNS activity (it releases the “vagal brake”) and increases SNS activity (Porges, 2007). Such a response, with a release of the heart’s “vagal brake” and an increase in SNS activity, occurs in response to both physical (Rowell, 1986; White & Raven, 2014) and mental task challenges (Van Roon, Mulder, Althaus, & Mulder, 2004).

In the presented study, we measured PNS activity with heart-rate variability (HRV). HRV refers to the variation in the interval between successive heartbeats (Berntson et al., 1997; Task Force, 1996). Frequency domain measures of HRV decompose the total variance into several power bands: very low (below 0.04 Hz), low (0.04–0.15 Hz), and high (0.15–0.40 Hz) frequency ranges (Berntson et al., 1997; Task Force, 1996). While very low and low-frequency HRV reflects a mixture of SNS and PNS activity, high-frequency HRV (HF-HRV) reflects vagal sinoatrial control and is thus a selective measure of parasympathetic activity (Berntson et al., 1997; Task Force, 1996). Importantly, it was shown that HF-HRV decreases in response to increased task difficulty (De Rivecourt, Kuperus, Post, & Mulder, 2008; Hjortskov et al. 2004; Mizuno, Tajima, Watanabe, & Kuratsune, 2014; Taelman, Vandeput, Vlemincx, Spaepen, & Van Huffel, 2011). For example, Taelman and others (2011) showed that power in the high-frequency range was reduced in task periods in comparison to rest; it was also further reduced in a difficult mental task than in an easy task; finally, it was reduced when a difficult task was performed for a first vs. second time. Moreover, the power in the high-frequency range decreases under mentally taxing conditions in executive function tasks which are time-constrained (Byrd, Reuther, McNamara, DeLucca, & Berg, 2015). Furthermore, in comparison to mid-frequency HRV, blood pressure, and several other cardiovascular indices, HF-HRV accounts for the largest proportion of variance in performance in an attention task (Duschek, Muckenthaler, Werner, & del Paso, 2009). Overall, these findings suggest that effort mobilization is related to the decrease in PNS activity.

SNS activity was assessed with systolic blood pressure (SBP). SBP is strongly influenced by the force with which the heart contracts, which directly reflects SNS activity (Wright, 1996). Total peripheral resistance, which is not systematically related to SNS activity, is the second determinant of SBP, but SBP reliably reflects beta-adrenergic sympathetic activity when parallel changes in total peripheral resistance are negligible. Although SBP is not a perfect measure of sympathetic activity; however, it is more valid than diastolic blood pressure (DBP), which is less influenced by cardiac contractility and more by total peripheral resistance, or heart rate (HR), which is dominated by the parasympathetic influence. Although there are more direct indices of heart contractility and thus SNS activity (such as pre-ejection period or RZ-interval), SBP is easier to measure. Most important, there is a great body of evidence from research on motivational intensity theory that used SBP to examine the theory's effort-related predictions and that demonstrated that SBP responds to changes in task difficulty and success importance (potential motivation) (see Gendolla, Wright, & Richter, 2012; Wright & Kirby, 2001, for overviews).

1.3. Effort in older age

Drawing on motivational intensity theory, especially the analysis of the influence of ability on effort, Hess (2014) suggested that older adults, due to cognitive constraints, need to exert more effort to support performance on a particular level in comparison to young people. It indeed turned out that older adults need to mobilize more effort to perform a task of any difficulty, but they also disengage from a task at lower difficulty levels in comparison to young adults (Ennis, Hess, & Smith, 2013; Hess & Ennis, 2012; Smith & Hess, 2015). Most importantly in the current context, self-involvement manipulations had a positive impact on effort mobilization and its impact was larger on older adults relative to younger adults (Smith & Hess, 2015). Furthermore, there is research that used SBP to examine effort mobilization in older adults and supported motivational intensity theory (e.g., Ennis et al., 2013; Hess & Ennis, 2012; Hess, Smith, & Sharifian, 2016; Smith & Hess, 2015; Stewart, Wright, & Griffith, 2016). Likewise, it was shown that HF-HRV decreases in response to a task challenge among older adults (Uchino, Uno, Holt-Lunstad, & Flinders, 1999). The effects on SBP and HF-HRV reactivity were observed despite the fact that older adults tend to have higher baseline blood pressure, lower heart-rate, reduced autonomic regulation (e.g., Bertel, Bühler, Kiowski, & Lütold, 1980; De Meersman & Stein, 2007; Lakatta, 1993), and higher SBP sensitivity to task challenge (Uchino, Birmingham, & Berg, 2010). Thus, despite the physiological changes in the cardiovascular system in older age, the evidence presented above suggests that effort in both young and elderly individuals is reflected by an increase in sympathetic activity (higher SBP) and a withdrawal of parasympathetic activity (decreased HF-HRV).

1.4. Present Research

The aim of the current study is to investigate the impact of effort mobilization on stereotyping in older age. Drawing on selective engagement theory (Hess, 2014) and motivational intensity theory (Brehm & Self, 1989; Wright, 1996), we expect an interactive impact of age and self-involvement on both effort mobilization and stereotyping tendency in a text comprehension task (similar to the task used by Radvansky and others, 2010, discussed earlier). In particular, we hypothesized that younger participants would mobilize low effort and show a low stereotyping tendency independently of their level of self-involvement. Older participants should invest high effort under conditions of high self-involvement and this should allow them to have a low stereotyping tendency. Under conditions of low-self-involvement, older adults should show low effort (because they disengage) and a strong stereotyping tendency. We hypothesize that the pattern of effort mobilization would occur during a story period, as suggested in previous research (e.g., Radvansky et al., 2010; Von Hippel et al., 1995). However, the evidence presented by Radvansky and colleagues that stereotypes only impact encoding (2010) is partial: they found evidence for the impact of stereotypes at encoding but did not present evidence that it occurs exclusively at those early stages of information processing. Thus, we also tested the patterns of effort mobilization during later stages of information processing, i.e. in the test period, using the same hypothesis as for the story period. The graphical representation of these hypotheses is shown in Figure 1 for stereotyping tendency (Panel A) and effort mobilization (Panel B).

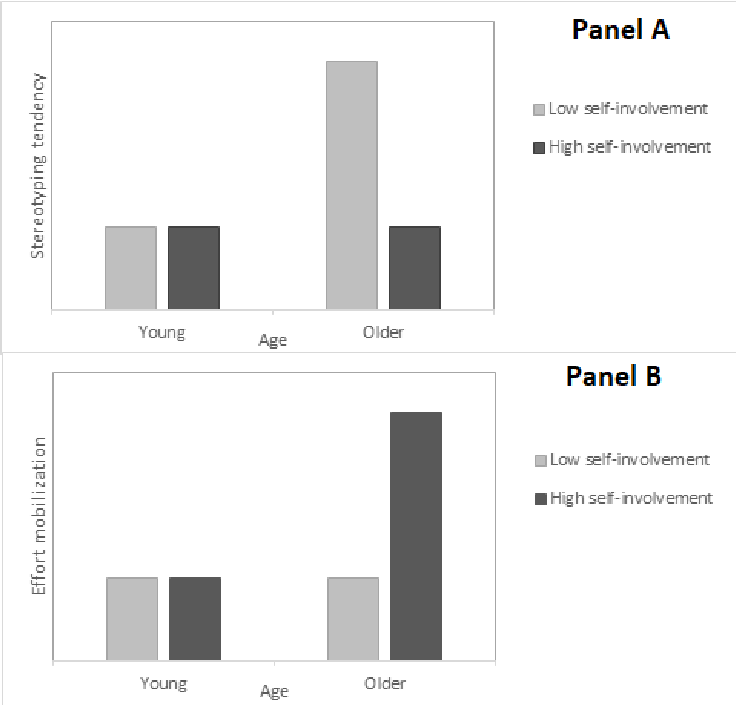

Figure 1. Hypothesized pattern of stereotyping tendency (Panel A) and effort mobilization (Panel B) as a function of age and self-involvement.

2. Method

2.1. Participants

Forty-eight young adults (39 women, 9 men; aged 19–27 years, *M* = 22.43, *SD* = 2.16) and 50 older adults (43 women, 7 men; aged 60–80 years, *M* = 66.20, *SD* = 4.66) took part in this study. The average number of years of education was *M* = 14.45, *SD* = 1.89 among young adults, and *M* = 15.91, *SD* = 1.69 among older adults. Young participants were recruited through online advertisements; older adults were recruited through the Universities of the Third Age in Kraków, online advertisement, and information distributed via study participants. Young adults were paid 10 PLN (3 USD) and older adults were paid 40 PLN (10 USD) for their participation. The difference in reimbursement was due to the higher costs that older participants incur (longer time spent on the tasks, cognitive screening, the need to commute to the university area). Importantly, the reimbursement difference did not affect the reported motivation to perform well in the tasks (see the results in the *Subjective Ratings* section). All subjects took part in the study voluntarily. This research was in accordance with the Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects and was accepted by the Ethical Committee of the Institute of Psychology (Jagiellonian University).

None of the young participants reported health problems related to the cardiovascular system. In contrast, 31 older adults reported cardiovascular diseases (the majority of these issues were hypertension) which were being treated with medication. Participants whose average SBP for any of the two baseline periods exceeded 160 mmHg were excluded from the sample; this was the case for three older adults. Furthermore, two older participants were excluded from the analysis because of problems with the cardiovascular measurements. Finally, two young adults were excluded from the sample: one participant decided to withdraw during the experiment and one participant struggled to understand the instructions because he was not a native Polish speaker. In sum, a total of five older adults and two young adults were excluded from the final dataset. The final sample thus consisted of 91 subjects, with 22 young and 24 older adults in the low-self-involvement conditions and 24 young and 21 older adults in the high-self-involvement conditions.

A priori power analysis for effort mobilization indices was conducted using *GPower* (ver. 3.1.9.2; Faul, Erdfelder, Buchner, & Lang, 2009). We expected to find a large effect size for our effort-related cardiovascular measures (such an effect size was observed in a previous study on the effects of self-involvement; Gendolla, Richter, & Silvia, 2008). Analyses revealed that a sample of 52 participants would be needed to obtain a power of .80 (Cohen, 1988). We also calculated a priori power analysis for the stereotyping tendency measure with the PANGEA application (Westfall, Judd, & Kenny, 2015). This showed that we needed 88 participants to detect a moderate-size effect in a 2 x 2 ANOVA design with 2 replicates. However, we decided to collect a 10% larger sample to secure the final sample size against the outliers. Our final sample, N = 91, was thus sufficiently powerful to detect the effects that we expected.

2.2. Measures and Equipment

2.2.1. Screening

Older participants were tested with the Addenbrooke's cognitive examination – III (ACE-III; Hsieh et al., 2013; Polish version of Senderecka et al., 2015). The maximum of scores in the global index of ACE-III is 100 points and scores below the cut-off points (88 and 82 points) suggest cognitive impairment. The average score of the current sample of older adults was M = 95.76, SD = 3.44, and all participants scored above the cut-offs. As is a standard practice in aging research, we measured participants' processing speed, need for closure (NFC), and affect. As those measures are not directly relevant for the current study we report them in the Supplemental materials (in section *S1. Individual differences and affect measures*).

2.2.2. Text Comprehension Task

We used a slightly modified version of the text comprehension task used by Czarnek and colleagues (2015). This task consists of reading a short story and then recalling story elements in a recognition test. In the current version of the task we included minor changes in wording, imposed a time-constraint in the story and test periods, and added more verbatim probes to the recognition tests.

To manipulate self-involvement, participants read one of two stories: a story about their peers (high self-involvement) or about people of a different age group (low self-involvement). One of the stories was about young drivers; the other was about older drivers. Both stories had 38 sentences (a 385-word story about young drivers and a 411-word story about older drivers). The stories were presented as individual sentences displayed on a computer screen; each sentence was presented for five seconds and it was not possible to repeat a sentence. In the recognition test, participants had to recall elements of the stories. The test was based on the task introduced by Schmalhofer and Glavanov (1986) and later adapted by Radvansky and colleagues (2010) to measure stereotyping tendency. For the test, 26 original sentences were taken from the stories and modified to various degrees to create test probes. Roughly half of the test probes were related to the stereotypical depiction of young or older drivers (e.g., Joanisse, Gagnon, & Voloaca, 2012; Williams, 2003) and the rest were neutral (they related to the general population of drivers or related not to any social group but to, e.g., road conditions). There were four types of test probes: verbatim probes (6 stereotypical and 6 neutral probes), paraphrases (2 neutral and 2 stereotypical probes), correct inferences (2 stereotypical and 2 neutral probes), and wrong inferences (2 neutral and 2 stereotypical probes)[1]. Participants were presented with one probe at a time, each presented for 4 seconds, after which they were asked to decide whether the presented probe had been part of the original story. Participants were warned that some sentences might be very similar to the original ones, but their task was to decide whether the sentences were identical. To measure stereotyping tendency, we analyzed the proportion of "yes" responses to the stereotypical wrong inferences (the exact wording is presented in Supplemental materials in section *S2. Stereotypical wrong inferences probes*). This is because stereotypes influence the inferences people make (Dunning & Sherman, 1997). In particular, stereotypes guide interpretations of ambiguous or vague information. For example, stereotypes could lead to an inference that the cause of the accident was reckless driving when informed that a young driver was involved in this accident. Conversely, being informed that an older driver was involved in an accident could lead to a stereotypical inference that the potential reason of this accident was poor vision or other health-related issues.

[1] There were also 2 additional stereotypical probes included in the test to reflect the specific definition of stereotypical wrong probes of Radvansky et al (2010). For the discussion of this topic see Czarnek et al. (2015). The responses for these probes were not included in the analysis.

2.2.3. Subjective Ratings of Tasks

After story and test periods, participants were given the NASA-TLX (Hart & Staveland, 1988; Polish version of Zieliński & Biernacki, 2010) to assess the level of workload of the task. In this measure, participants answered six questions regarding the perceived mental workload, physical workload, time pressure, effort, performance, and frustration related to a task. All responses were given using a 21-point slider. Later, participants were asked to indicate which aspect of the workload was more strongly felt in a task: e.g., mental vs. physical workload, or mental workload vs. time pressure (there are 15 such pairwise comparisons of workload aspects). The number of times a workload aspect is chosen serves as a weight. For each aspect of workload, the index is calculated by multiplying the rating on a 21-point scale by the corresponding weight and then dividing by 15.

We also asked participants about task difficulty ("How difficult was the task?", on a scale from ”very easy" to "very difficult"), tiredness ("Was the task tiring for you?"), and motivation to do well in the tasks ("Did you care about performing well in the task?"). Additionally, after the test period, we asked about anxiety related to their performance ("Were you afraid that the task would show that you have low abilities?") and the experimenter ("Did the presence of a person running this study bother you?"), as well as their perception of the possibility of succeeding in a task ("Did you think that good performance in the task was possible?"). Participants answered these questions using the 21-point slider (with anchor labels for all the questions, but with difficulty ratings from "definitely not" to ”definitely yes").

2.2.4. Physiological Measures

During the experiment, we collected continuous measures of blood pressure, HR, and inter-beat intervals (IBI) with a Finometer MIDI (Finapres Medical Systems, Amsterdam, The Netherlands). The Finometer MIDI device uses Peňáz's method (1973) to measure beat-to-beat blood pressure. Cardiovascular measurements were taken from the middle or ring finger of a subject's non-dominant hand using a cuff. BeatScope Easy software, which is provided with the Finometer device, was used to extrapolate pressure readings from the finger to represent brachial artery pressure. The validity of this technology has previously been shown (Bogert & van Lieshout, 2005).

2.3. Procedure

Participants were tested individually. All the tasks and questionnaires, except the ACE-III and processing speed task, were programmed in Inquisit 4 software (Millisecond Software, Seattle) and presented to participants on a 19-inch LCD monitor. After signing the consent agreement, participants took the processing speed task and responded in questionnaires (reported in the Supplemental materials in section *S1. Individual differences and affect measures*). Later, participants' cardiovascular activity was assessed during an 8-minute baseline period while participants watched a relaxing movie.

Next, participants were assigned to the low- or high-self-involvement condition. The self-involvement was manipulated as in Czarnek and colleagues (2015), namely with the relevance of the story: in the low-self-involvement condition, participants read about the outgroup (for older participants it was the story on young drivers; for young subjects it was the story on older drivers). In the high-self-involvement condition, participants read about their ingroup (for older participants it was a story on older drivers; for young participants the story was on young drivers). After reading one story, participants completed the NASA-TLX and answered three subjective rating questions (task difficulty, tiredness, and motivation to do well) concerning the story. A second 8-minute cardiovascular baseline period followed, during which participants again watched a relaxing movie.

Later, participants were given a recognition test and again completed NASA-TLX; they also answered six subjective rating questions concerning the test (task difficulty, tiredness, motivation to do well, possibility of succeeding in a task, anxiety related to test results and experimenter) and rated their affect. At the end of the session, older adults were tested with ACE-III. Lastly, participants were thanked, paid, and debriefed.

2.4. Cardiovascular Data Processing

SBP, DPB, and HR values assessed during the two baseline periods and the two task periods were screened for potential outliers. Samples deviating by more than 8 units from both the preceding and the following sample were excluded (e.g., Schmidt, Richter, Gendolla, & Van der Linden, 2010). The remaining values were used to compute the first baseline score by calculating the arithmetic mean of data collected over the last 180 seconds of the first baseline period. For the second baseline score, we used the last 150 seconds of measurement. For the task periods, we calculated cardiovascular response for the first 180 seconds of a story and 150 seconds of the test period[2]. Cardiovascular change scores (delta) were computed for SBP, DBP, and HR for each participant in both task periods by subtracting baseline scores from the associated task scores (Llabre, Spitzer, Saab, Ironson, & Schneiderman, 1991).

[2] We only used 150-second periods for the test period and the associated baseline period because the fastest participant took only 153 seconds to perform the test. We also conducted an additional analysis of the full periods that participants spent on responding to a test; this is available in the Supplementary materials (in section *S5. Cardiovascular responses in full test period*). This analysis led to a similar conclusion as the analysis for the first 150 seconds of the task.

IBI data was used to calculate the frequency measure of HRV. As with the analyses of blood pressure and IBI data, we used a 180-second measurement of IBI from the first baseline and story period, and a 150-second IBI measurement from the second baseline and the test period. We calculated HF-HRV in normalized units (n.u.; nHF-HRV) in HRVAS software (Ramshur, 2010). Spectral decomposition was performed with the autoregressive method. The power values for the high-frequency band were defined according to established guidelines as 0.15–0.40 Hz (Task Force, 1996). We identified ectopic beats as IBIs whose values differed by more than 20% or 3 standard deviations in comparison to the preceding IBI (e.g., Akhter, Gite, Tharewal, & Kale, 2015). Beats identified as ectopic were removed from analysis (Lippman, Stein, & Lerman, 1994). We interpolated the IBI signal with a rate of 10 Hz. Furthermore, we used a wavelet-based method for trend removal (Thuraisingham, 2006). We did not measure respiratory parameters. The nHF-HRV change scores (delta) were computed for each participant by subtracting baseline period scores from respective task period scores.

In the analysis we focus on SBP and nHF-HRV change scores as we have specific hypotheses for these measures. While positive values of SBP change scores are interpreted as sympathetic activation, negative values of nHF-HRV change scores imply a decrease in parasympathetic activity. The analyses of DBP and HR are presented in the Supplemental Materials (in section *S4. Cardiovascular responses: DBP and HR*).

2.5. Statistical analysis

Since prediction of the joint effect of age and self-involvement on participants' stereotyping tendency and effort mobilization are not captured well by the 2 x 2 ANOVA (Rosenthal & Rosnow, 1985), we tested specific contrasts that reflect our theory-driven hypotheses. For stereotyping tendency, the contrast weights were +3 for older adults in the low-self-involvement condition and -1 for all the other three groups. For effort mobilization the contrast weights were +3 for older adults in the high-self-involvement condition and -1 for all the other three groups. We analyzed effort mobilization in both the story and the test periods.

The contrast weights correspond to the theoretical predictions presented in Figure 1. In the planned contrast analysis, we used a one-tailed test of significance[3] and a conventional significance level, i.e., $\alpha = 0.05$. We also calculated Likelihood Ratios (LR), which tests the null hypothesis against the hypothesized models, and we used the correction of Hurvich and Tsai (1989) for unequal numbers of parameters. LR is a measure that compares the probability of results under one hypothesis with the

[3] We used one-tailed testing only to test our directional hypotheses. Otherwise, two-tailed tests are used.

probability of these results under a competing hypothesis (Glover & Dixon, 2004; Richter, 2016). In particular, we compared the unexplained sum of squares of the null model against the unexplained sum of squares of our hypothesized model. Royall (1997) proposed that LR below 8 should be interpreted as weak evidence, 8–32 as moderate evidence, and above 32 as strong evidence. Data analysis was carried out in IBM SPSS Statistics 24 (IBM Corporation, 2016) and Open Office Calc software (The Apache Software Foundation).

3. Results

3.1. Stereotyping Tendency

The planned contrast was significant for the stereotyping tendency measure, $F(1,87) = 17.43$, $MSE = 0.08$, $p < .001$, $\eta^2 = .17$; in comparison to the other three conditions, the stereotyping tendency was highest among older participants in the low-self-involvement condition, $M = 0.50$, $SE = 0.07$. Average stereotyping tendency among young adults in the high-self-involvement condition was $M = 0.27$, $SE = 0.06$; among young adults in the low-self-involvement condition it was $M = 0.11$, $SE = 0.05$; among older adults in the high-self-involvement condition it was $M = 0.29$, $SE = 0.06$. Figure 2 presents this pattern. Corrected LR for this comparison was $\lambda = 912.64$, which provides very strong evidence for the hypothesized joint impact of age and self-involvement manipulation on stereotyping tendency. In fact, the observed data are 912 times more likely under our hypothesized model in comparison to the null hypothesis. Importantly, the planned contrast remained significant and the LR was still high after inclusion of control variables: responses to neutral wrong inferences or responses to stereotypical correct inferences (the details of this analysis as well as analysis of probes other than stereotypical wrong inferences are presented in the Supplemental materials in section *S3. Stereotyping tendency: additional analysis*). These results suggest that the stereotyping tendency of older adults, i.e., acceptance of stereotypical wrong inferences, occurs above and beyond age differences in the tendency to accept neutral wrong inferences and stereotypical correct inferences.

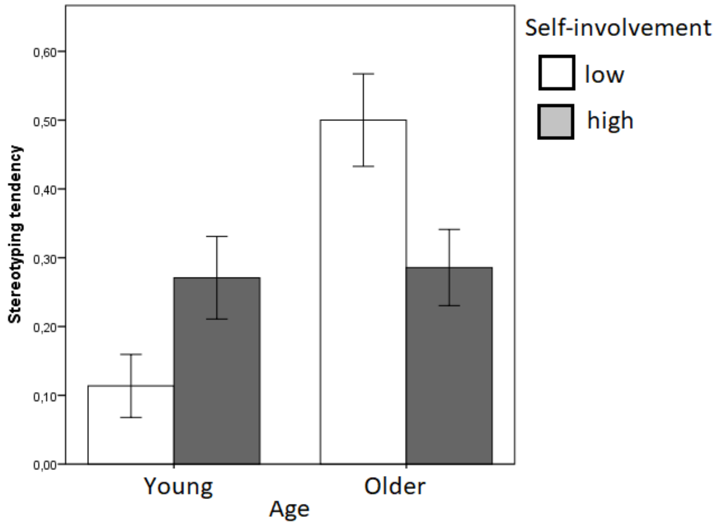

Figure 2. Mean stereotyping tendency. Error bars represent +/- 1 standard error of the mean.

3.2. Cardiovascular Responses

3.2.1. Baseline Periods

Resting cardiovascular baseline scores are presented in Table 1. As we did not have precise predictions of the pattern of baseline scores, we submitted them to a 2 (age: young vs. older adults) x 2 (self-involvement: low vs. high) ANOVA. For SBP, there were significant effects of age in the first, $F(1,87) = 9.30$, $p = .003$, $\eta^2_p = .10$, and second, $F(1,87) = 11.11$, $p = .001$, $\eta^2_p = .11$, baseline periods; older adults had higher scores than younger counterparts. There were no effects of age or condition on nHF-HRV. Instead, there was a correlation between nHF-HRV for the first baseline and story period scores, $r(90) = -.49$, $p < .001$, and between the second baseline and test period scores, $r(90) = -.53$, $p < .001$. No correlation between baseline and change scores emerged for SBP. Because of the significant differences in several baseline scores and correlations between baseline and reactivity values, we included baseline scores as covariates in the analyses of cardiovascular indices.

Table 1
Cell means and standard errors of the SBP and nHF-HRV baseline scores

			Baseline 1		Baseline 2	
Age	Self-involvement		SBP	nHF-HRV	SBP	nHF-HRV
Young	Low	*M*	114.67	0.58	115.26	0.54
		SE	3.40	0.04	3.38	0.04
	High	*M*	114.33	0.55	117.00	0.54
		SE	2.38	0.05	1.95	0.05
Older	Low	*M*	122.91	0.64	126.79	0.57

		SE	2.81	0.03	3.49	0.03
	High	*M*	124.14	0.55	126.33	0.54
		SE	3.26	0.04	3.52	0.04

Note. SBP values are in mmHg; nHF-HRV values are in n.u.

3.2.2. Story Period

In the story period, the planned contrast was significant for SBP, $F(1,86) = 7.48$, $MSE = 57.82$, $p = .004$, $\eta^2_p = .08$, but not significant for nHF-HRV change scores, $F(1,86) = 2.09$, $MSE = 0.03$, $p = .076$, $\eta^2_p = .02$. Furthermore, comparing the evidence of the null hypothesis against the evidence of the hypothesized pattern resulted in corrected LR: $\lambda = 4.66$ for SBP and $\lambda = 1.00$ for nHF-HRV. This provides weak evidence in favor of the joint impact of age and self-involvement on effort mobilization for SBP, but no evidence for nHF-HRV. Cell means and standard errors of cardiovascular change scores in the story period can be found in Table 2. The analyses of DBP and HR change scores are presented in the Supplemental Materials (in section *S4. Cardiovascular responses: DBP and HR*).

Table 2
Cell means and standard errors of the SBP and nHF-HRV change scores in the story and test periods

			Story period		Test period	
Age	Self-involvement		SBP	nHF-HRV	SBP	nHF-HRV
Young	Low	*M*	2.14	0.02	2.04	0.08
		SE	1.69	0.03	1.48	0.04
	High	*M*	3.36	0.03	4.33	0.06
		SE	1.14	0.04	1.32	0.03
Older	Low	*M*	14.60	-0.07	12.81	-0.01
		SE	1.88	0.05	2.08	0.05
	High	*M*	11.14	-0.05	12.89	-0.07
		SE	1.73	0.04	2.11	0.04

Note. SBP values are in mmHg; nHF-HRV values are in n.u.

3.2.3. Test Period

In the test period, the planned contrasts were significant for SBP, $F(1,86) = 10.63$, $MSE = 71.14$, $p = .001$, $\eta^2_p = .11$, and nHF-HRV change scores, $F(1,86) = 8.74$, $MSE = 0.03$, $p = .002$, $\eta^2_p = .09$. Comparing the evidence for the null hypothesis against the evidence of the predicted pattern resulted in corrected LR: $\lambda = 25.99$ for SBP and $\lambda = 25.50$ for nHF-HRV. This means that the data is 25 times more likely under our hypothesized model than under the null hypothesis model for SBP as well as for nHF-

HRV. This provides moderate evidence in favor of the joint impact of age and self-involvement on effort mobilization for SBP and nHF-HRV. Cell means and standard errors of cardiovascular change scores in the test period can be found in Table 2, Figure 3, and Figure 4. The analyses for DBP and HR change scores are presented in the Supplemental Materials (in section *S4. Cardiovascular responses: DBP and HR*).

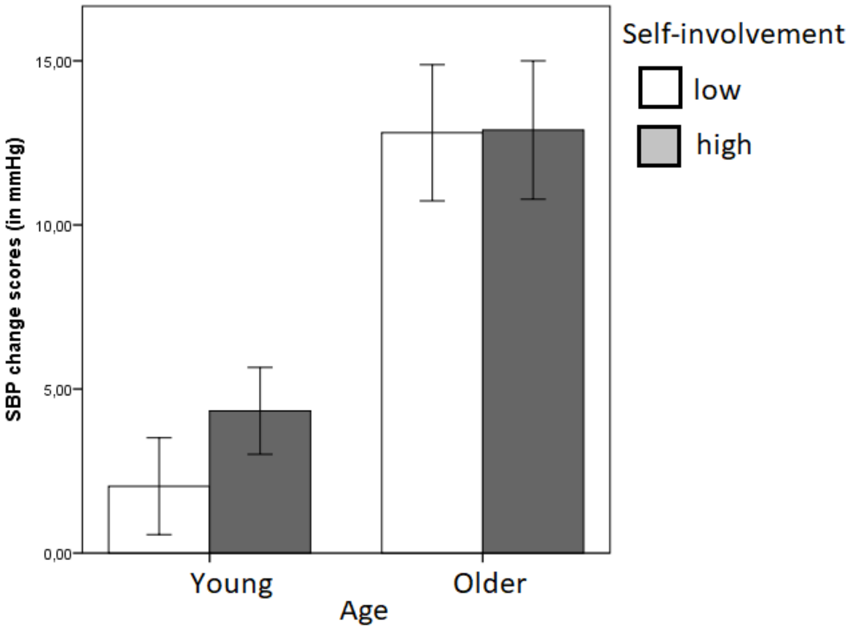

Figure 3. Mean SBP change scores in the test period. Error bars represent +/- 1 standard error of the mean.

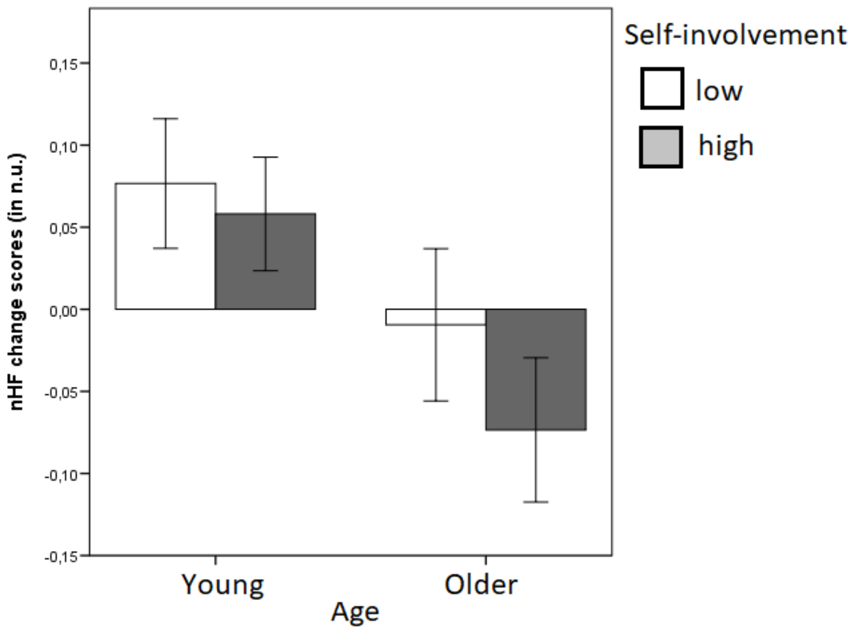

Figure 4. Mean nHF-HRV change scores in the test period. Error bars represent +/- 1 standard error of the mean.

3.3. Subjective Ratings

We submitted subjective ratings of story and test periods to a 2 (age: young vs. older adults) x 2 (self-involvement: low vs. high) ANOVA. Subjective ratings of the task were measured with NASA-TLX and a set of additional questions. For the current purposes, we focus on ratings of mental workload, effort, task difficulty, and motivation to do well, as these ratings are most relevant to the current study. The rest of the subjective ratings are presented in the Supplemental materials (in section *S6. Subjective ratings*).

Story Period. Cell means with standard errors for workload ratings in the story period are provided in Table 3. We did not find significant effects for any rating of interest. Importantly, ratings of motivation to do well in a task were high among all the four groups.

Table 3

Cell means and standard errors of the workload and task perception ratings in the story and test periods

			Young adults		Older adults	
			Low self-involvement	High self-involvement	Low self-involvement	High self-involvement
Mental demand	Story	*M*	2.91	3.18	2.89	2.27
		SE	0.27	0.30	0.38	0.35
	Test	*M*	2.96	3.51	3.57	3.07
		SE	0.28	0.33	0.36	0.36

Effort	Story	*M*	1.34	1.47	1.04	1.00
		SE	0.21	0.30	0.18	0.22
	Test	*M*	2.37	1.71	1.28	1.31
		SE	0.26	0.24	0.20	0.30
Difficulty	Story	*M*	9.64	8.83	7.71	7.05
		SE	0.90	1.09	0.97	0.96
	Test	*M*	13.68	11.38	10.71	10.40
		SE	0.81	0.98	1.29	1.05
Motivation to do well	Story	*M*	16.73	17.04	18.33	16.90
		SE	0.65	0.60	0.65	1.05
	Test	*M*	17.18	17.08	17.83	15.65
		SE	0.64	0.66	0.69	1.06

Test Period. Cell means with standard errors for the workload ratings of interest in the test period are provided in Table 3. We found a main effect of age for the perception of effort, $F(1,86) = 8.91$, $p = .004$, $\eta^2_p = .09$, with young adults reporting higher levels of effort in comparison to older adults. We did not find significant differences for other ratings of interest. Importantly, all experimental groups perceived succeeding in a task as possible and worthwhile.

4. Discussion

This study examined if stereotyping tendency in older age might be decreased through self-involvement and the associated increased effort mobilization. The results presented above support our hypothesis regarding decreased stereotyping tendency among older adults in the high- versus low-self-involvement condition. Importantly, we found evidence that older adults in the high-self-involvement condition presented stereotyping tendency on a similarly low level as young adults. In contrast, older adults in the low-self-involvement condition presented a higher stereotyping tendency than people in the three other conditions (both groups of young adults and older adults in the high-self-involvement condition). These results replicate our previous findings (Czarnek et al., 2015) that the relationship between age and stereotyping tendency is moderated by self-involvement. They are also in line with research on the impact of motivation on person perception among older adults (Hess et al., 1998, Study 1; Hess et al., 2009; Hess et al., 2001), but they provide more direct evidence regarding effort mobilization during the task.

In particular, we found moderate evidence for our effort-related hypotheses for SBP and nHF-HRV reactivity in the test period. The nHF-HRV reactivity displayed the predicted pattern but the SBP response in older adults did not show the expected self-involvement effect. Although contrast analysis is a powerful technique, it has weaknesses (see Abelson & Prentice, 1997; Richter, 2016). Accordingly, the positive statistical result for SBP reactivity should be treated with caution.

However, the pattern of parasympathetic response closely followed the predicted pattern: the decrease in nHF-HRV was most pronounced in the high-self-involvement group of older participants. This is in line with models of effort that suggest that the cardiac response to low task demands is mainly driven by a decrease in parasympathetic activity (Rowell, 1993; Ogoh, et al., 2005; Van Roon et al., 2004; White & Raven, 2014). The combination of evidence for parasympathetic withdrawal with a lack of evidence for increased sympathetic activity could thus suggest that overall effort investment in the text processing task was relatively low. An alternative interpretation of the data would be that the nHF-HRV response was mainly driven by the implementation of cognitive control to overcome the stereotyping tendency. Corroborating this idea, Mathewson, Dywan, Snyder, Tays, and Segalowitz (2011) demonstrated that HF-HRV is related to error monitoring and Byrd and others (2015) suggested that HF-HRV is sensitive to increasing demands in executive tasks. However, future research would be needed to disentangle these alternative explanations. Interestingly, the hypothesized pattern of parasympathetic withdrawal occurred only in the test period and not in the story phase. This raises the question of whether stereotypes influence comprehension during encoding, as was proposed by Radvansky and colleagues (2010). Our results indicate that reduced stereotyping tendency coincides with increased effort mobilization during information retrieval (test period), but not during encoding (story period). This could suggest that inhibition of stereotyping tendency occurs in the later stages of information processing (e.g., Devine, 1989; Kundra & Sinclair, 1999; Sherman, Macrae, & Bodenhausen, 2000). Additional research is needed to answer the question of the relative importance of stereotype inhibition at the level of encoding or retrieval. For example, a study employing a lexical decision task (e.g., Radvansky et al., 2010, Study 2) or a study in which effort mobilization is instigated only during a reading or test phase could help to answer this question. Furthermore, we believe that our results highlight the importance of assessing the impact of both branches of the autonomic nervous system in resource mobilization. This also corroborates with previous studies which used HRV as a measure of parasympathetic withdrawal among older adults (e.g., Smith et al., 2009; Uchino et al., 1999).

It is also worth mentioning that subjective perceptions of effort did not mimic the actual effort investments measured with cardiovascular reactivity, despite the fact that the measure of mental effort (NASA-TLX) has been successfully employed in previous aging studies (e.g., Hess et al., 2016). In the majority of ratings, young and older adults did not differ. We found only that younger adults perceived that they mobilized more effort during the test period in comparison to older adults. This effect is at odds with the observed pattern of cardiovascular activation; however, it is in line with research showing that self-report measures of effort do not necessarily reflect actual task engagement (e.g., Harper, Eddington, & Silvia, 2016; Silvia & Gendolla, 2001; Smith & Hess, 2015, Stewart et al., 2016). It is possible that self-report measures of effort are influenced by age-related differences in strategies and standards (Uchino et al., 2010). To elaborate on the relationship between subjective effort and cardiovascular measures, research with self-involvement manipulation and varied difficulty levels across age groups would be needed.

Furthermore, older adults presented higher SBP in both task periods in comparison to young people. There are several possible explanations of this effect: the higher reactivity of older adults, the impact of total peripheral resistance, the difference in the reimbursement participants were offered, and underpowered analysis. Firstly, our SBP findings are in line with research indicating that older adults are on average more physiologically reactive in comparison to their young counterparts (Uchino et al., 2010). This reactivity concerns only blood pressure to some degree, not HR. Such an increased blood pressure reactivity might be problematic in research comparing different age groups because it may not be related to the processes of interest. On the other hand, age-related increased SBP reactivity might reflect an active coping response (Hess & Ennis, 2014). Our results indeed indicate that, on average, older adults are more reactive in comparison to their young counterparts; however, in all of the analyses of cardiovascular responses, we used a baseline value as a covariate. This precludes the possibility that the age-related effects that we found reflect only a correlation between baseline values and reactivity scores. Furthermore, baseline measurements were taken twice throughout the experiment to reduce the potential carry-over effect between story and test periods. Thus, we believe that the cardiovascular change scores presented above reflect effort mobilization processes. This elevated responsivity among older people might also be reflective of inefficient resource use due to self-regulatory failure (Hess & Ennis, 2014).

Secondly, the elevated SBP reactivity among older people is surprising given the fact that many of them were suffering from hypertension and took medication to treat this condition. Consequently, we could instead expect diminished SBP (or sympathetic) reactivity among older participants because of beta-blocker drugs, but we observed the opposite. Yet, for our hypothesis the crucial differences are within age groups, not between them. Furthermore, as already mentioned, SBP might be influenced by total peripheral resistance, and we did not control for this in the current study. Thus, future research should employ more direct measures of sympathetic activity to examine hypothesis of low effort mobilization in stereotyping in older age. In particular, the most sensitive non-invasive index of beta-adrenergic activation of the heart is the pre-ejection period, which is a measure of heart contractility (Kelsey, 2012).

On the other hand, the higher overall SBP reactivity might be related to the difference in reimbursement for younger and older adults. In particular, young participants were paid considerably less than older adults. This was related to the higher costs of taking part in the study for older adults. However, the current study design does not allow for a rebuttal of the hypothesis that increased SBP reactivity among older people was related to the higher reimbursement they were offered. To test this, a study which manipulates the reimbursement *within* age groups would be needed. For the current purposes, however, it is important to note that younger and older adults did not differ in their reported motivation to do well in tasks, which suggests that there was no effect of different levels of reimbursement.

Finally, it seems that the effect of self-involvement on effort mobilization (both for SBP and HF-HRV) in older age might be smaller in size than what we expected. Thus, future studies might want to employ larger sample sizes, especially if their aim is to investigate whether effort mobilization mediates the relationship between age, self-involvement, and stereotyping (Rucker, Preacher, & Hayes, 2007; Schoemann, Boulton, & Short, 2017). Moreover, future research might focus on increasing the signal to noise ratio not only by employing larger sample sizes, but also by employing stronger manipulations of self-involvement or by adding more probes to the measure of stereotyping tendency (in the current study there were two probes measuring stereotyping tendency).

4.1. Conclusions

Older adults present a higher stereotyping tendency towards many social groups in comparison to young adults. However, we demonstrated that these age-related changes might be, we believe, compensated for by self-involvement. Secondly, we showed that selective engagement theory (Hess, 2014) and motivational intensity theory (Brehm & Self, 1989; Wright, 1996) are useful in predicting effort mobilization and performance in social cognitive tasks. Third, we have provided initial support for differential effort mobilization in stereotyping. In particular, we showed that stereotyping among older people is accompanied by lower levels of engagement evinced by the lack of parasympathetic withdrawal. This corroborates the assumption that stereotyping is relatively effortless and serves as an efficient mental shortcut (e.g., Fiske et al., 1999; Macrae, Milne, & Bodenhausen, 1994). Thus, this research demonstrates that the metaphor of social perceivers as "motivated tacticians" (Fiske & Taylor, 1991) is especially relevant in explaining the stereotyping tendencies of older adults. An older adult is, or can be, as Fiske and Taylor (1991) put it, "a fully engaged thinker who has multiple cognitive strategies available and chooses among them based on goals, motives, and needs” (p. 13).

References

Abelson, R. P., & Prentice, D. A. (1997). Contrast tests of interaction hypothesis. *Psychological Methods*, *2*(4), 315-328.

Akhter, N., Gite, H., Tharewal, S., & Kale, K. V. (2015). Computer based RR-interval detection system with ectopy correction in HRV data. *International Conference on Advances in Computing, Communications and Informatics, IEEE* (pp. 1613-1618).

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology*, *23*(5), 611-626.

Berntson, G. G., Bigger, J. T., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., . . . Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology, 34*(6), 623-648. doi: 10.1111/j.1469-8986.1997.tb02140.x

Bertel, O., Bühler, F. R., Kiowski, W., & Lütold, B. E. (1980). Decreased Beta-adrenoreceptor responsiveness as related to age, blood pressure, and plasma catecholamines in patients with essential hypertension. *Hypertension*, *2*(2), 130-138.

Bogert, L. W., & van Lieshout, J. J. (2005). Non-invasive pulsatile arterial pressure and stroke volume changes from the human finger. *Experimental Physiology*, *90*(4), 437-446. doi: 10.1113/expphysiol.2005.030262

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology, 40*, 109–131. doi: 10.1146/annurev.psych.40.1.109

Byrd, D. L., Reuther, E. T., McNamara, J. P., DeLucca, T. L., & Berg, W. K. (2015). Age differences in high frequency phasic heart rate variability and performance response to increased executive function load in three executive function tasks. *Frontiers in Psychology*, *5*, 1470. doi: 10.3389/fpsyg.2014.01470

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159. doi: 10.1037/0033-2909.112.1.155

Czarnek, G., Kossowska, M., & Sędek G. (2015). The Influence of Aging on Outgroup Stereotypes: The Mediating Role of Cognitive and Motivational Facets of Deficient Flexibility. *Experimental Aging Research, 41(3)*, 303-324. doi: 10.1080/0361073X.2015.1021647

De Meersman, R. E., & Stein, P. K. (2007). Vagal modulation and aging. *Biological psychology*, *74*(2), 165-173. doi: 10.1016/j.biopsycho.2006.04.008

De Rivecourt, M., Kuperus, M. N., Post, W. J., & Mulder, L. J. M. (2008). Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. *Ergonomics*, *51*(9), 1295-1319.Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5-18.

Dunning, D., & Sherman, D. A. (1997). Stereotypes and tacit inference. *Journal of Personality and Social Psychology, 73,* 459–471. doi: 10.1037/0022-3514.73.3.459

Duschek, S., Muckenthaler, M., Werner, N., & del Paso, G. A. R. (2009). Relationships between features of autonomic cardiovascular control and cognitive performance. *Biological Psychology, 81*(2), 110-117.

Ennis, G. E., Hess, T. M., & Smith, B. T. (2013). The Impact of age and motivation on cognitive effort: Implications for cognitive engagement in older adulthood. *Psychology and Aging, 28*, 495–504. doi: 10.1037/a0031255

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using GPower 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41,* 1149 – 1160. doi:10.3758/BRM.41.4.1149

Firebaugh, G., & Davis, K. E. (1988). Trends in antiblack prejudice, 1972-1984: Region and cohort effects. *American Journal of Sociology*, *94*, 251-272. doi:10.1086/228991

Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231-254). New York, NY: Guilford Press.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition, 2nd Edition*. New York, NY: McGraw-Hill.

Gendolla, G. H. E. (2004). The intensity of motivation when the self is involved: An application of Brehm's theory of motivation to effort-related cardiovascular response. In R. A. Wright, J. Greenberg, & S. S. Brehm (Eds.), *Motivational analyses of social behavior* (pp. 205–244). Mahwah, NJ: Erlbaum.

Gendolla, G. H. E., & Richter, M. (2010). Effort mobilization when the self is involved: Some lessons from the cardiovascular system. *Review of General Psychology, 14*(3), 212-226. doi: 10.1037/a0019742

Gendolla, G. H. E., Richter, M., & Silvia, P. (2008). Self-focus and task difficulty effects on effort-related cardiovascular reactivity. *Psychophysiology, 45,* 653– 662. doi: 10.1111/j.1469-8986.2008.00655.x

Gendolla, G. H., Wright, R. A., & Richter, M. (2012). Effort intensity: Some insights from the cardiovascular system. *The Oxford handbook of human motivation*, 420-438.

Gendolla, G. H. E., & Wright, R. A. (2009). Effort. In D. Sander & K. R. Scherer (Eds.), *Oxford companion to the affective sciences* (pp. 134– 135). New York, NY: Oxford University Press

Glover, S., & Dixon, P. (2004). Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review,11,* 791-806. doi: 10.3758/BF03196706

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology, 45*, 410–414. doi: 10.1016/j.jesp.2008.11.004

Harper, K. L., Eddington, K. M., & Silvia, P. J. (2016). Perfectionism and Effort-Related Cardiac Activity: Do Perfectionists Try Harder? *Plos One, 11*(8), e0160340. doi: 10.1371/journal.pone.0160340

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139–183). Amsterdam, The Netherlands: North-Holland Press. doi: 10.1016/S0166-4115(08)62386-9.

Henry, J. D., von Hippel, W. V., & Baynes, K. (2009). Social inappropriateness, executive control and aging. *Psychology and Aging, 24*, 239-244. doi: 10.1037/a0013423

Herek, G.M. (1994). Assessing attitudes toward lesbians and gay men: A review of empirical research with ATLG scale. In B. Greene & G.M. Herek (Eds.), *Lesbian and Gay Psychology* (pp. 206-228). Thousand Oaks, CA: Sage Publications.

Hess, T. M. (1994). Social cognition in adulthood: Aging-related changes in knowledge and processing mechanisms. *Developmental Review, 14*, 373-412. doi: 10.1006/drev.1994.1015

Hess, T.M. (1999). Cognitive and knowledge-based influences on social representations. In T.M. Hess & F. Blanchard-Fields (Eds.), *Social cognition and aging* (pp. 237-264). San Diego: Academic Press.

Hess, T. M. (2014). Selective engagement of cognitive resources: Motivational influences on older adults' cognitive functioning. *Perspectives on Psychological Science, 9*(4), 388–407. doi: 10.1177/1745691614527465.

Hess, T. M., & Emery, L. (2012). Memory in context: The impact of age-related goals on performance. In M. Naveh-Benjamin & N. Ohta (Eds.), *Memory and aging: Current issues and future directions* (pp. 183-214). New York, NY, US: Psychology Press.

Hess, T. M., & Ennis, G. E. (2012). Age differences in the effort and cost associated with cognitive activity. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 67*, 447–455. doi: 10.1093/geronb/gbr129

Hess, T. M., & Ennis, G. E. (2014). Assessment of adult age differences in task engagement: The utility of systolic blood pressure. *Motivation and Emotion, 38*, 844–854. doi: 10.1007/s11031-014-9433-2

Hess, T. M., Follett, K. J., & McGee, K. A. (1998). Aging and impression formation: The impact of processing skills and goals. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *53*(3), 175-187. doi: 10.1093/geronb/53B.3.P175

Hess, T.M., Germain, C.M., Swaim, E.L., & Osowski, N.L. (2009). Aging and selective engagement: the moderating impact of motivation on older adults' resource utilization. *Journal of Gerontology: Psychological Sciences, 64B*(4), 447–456, doi: 10.1093/geronb/gbp020

Hess, T. M., Rosenberg, D. C., & Waters, S. J. (2001). Motivation and representational processes in adulthood: The effects of social accountability and information relevance. *Psychology and Aging, 16*, 629–642. doi: 10.1037//0882-7974.16.4.629

Hess, T. M., Smith, B. T., & Sharifian, N. (2016). Aging and effort expenditure: The impact of subjective perceptions of task demands. *Psychology and Aging*, *31*, 653-660. doi: 10.1037/pag0000127

Hjortskov, N., Rissén, D., Blangsted, A. K., Fallentin, N., Lundberg, U., & Søgaard, K. (2004). The effect of mental stress on heart rate variability and blood pressure during computer work. *European Journal of Applied Physiology*, *92*(1-2), 84-89. doi: 10.1007/s00421-004-1055-z

Hurvich, C. M., & Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika, 76,* 297–307. doi: 10.1093/biomet/76.2.297

Joanisse, M., Gagnon, S., & Voloaca, M. (2012). Overly cautious and dangerous: An empirical evidence of the older driver stereotypes. *Accident Analysis and Prevention*, *45*, 802–810. doi: 10.1016/j.aap.2011.11.002

Klabunde, R. (2011). *Cardiovascular physiology concepts*. Philadelphia (PA): Lippincott Williams & Wilkins.

Kreibig, S. D., Gendolla, G. H., & Scherer, K. R. (2012). Goal relevance and goal conduciveness appraisals lead to differential autonomic reactivity in emotional responding to performance feedback. *Biological Psychology*, *91*(3), 365-375. doi: 10.1016/j.biopsycho.2012.08.007

Krendl, A. C., Heatherton, T. F., & Kensinger, E. A. (2009). Aging minds and twisting attitudes: An fMRI investigation of age differences in inhibiting prejudice. *Psychology and Aging, 24*(3), 530-541. doi: 10.1037/a0016065

Krendl, A. C., & Kensinger, E. A. (2016). Does Older Adults' Cognitive Function Disrupt the Malleability of Their Attitudes toward Outgroup Members?: An fMRI Investigation. *PloS one*, *11*(4), e0152698. doi: 10.1371/journal.pone.0152698

Lakatta, E. G. (1993). Cardiovascular regulatory mechanisms in advanced age. *Physiological Reviews*, *73*, 413-467.

Lippman, N. E. A. L., Stein, K. M., & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. *American Journal of Physiology – Heart and Circulatory Physiology*, *267*(1), 411-418.

Llabre, M. M., Spitzer, S. B., Saab, P. G., Ironson, G. H., & Schneiderman, N. (1991). The reliability and specificity of delta versus residualized change as measure of cardiovascular reactivity to behavioral challenges. *Psychophysiology, 28,* 701–711. doi: 10.1111/j.1469-8986.1991.tb01017.x

Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Jetten, J. (1994). Out of mind but back in sight: Stereotypes on the rebound. *Journal of Personality and Social Psychology*, *67*(5), 808-817. doi: 10.1037/0022-3514.67.5.808

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences*, *9*(10), 496-502. doi: 10.1016/j.tics.2005.08.005

Mathewson, K. J., Dywan, J., Snyder, P. J., Tays, W. J., & Segalowitz, S. J. (2011). Autonomic regulation and maze-learning performance in older and younger adults. *Biological Psychology*, *88*(1), 20-27. doi: 10.1016/j.biopsycho.2011.06.003

Mizuno, K., Tajima, K., Watanabe, Y., & Kuratsune, H. (2014). Fatigue correlates with the decrease in parasympathetic sinus modulation induced by a cognitive challenge. *Behavioral and Brain Functions*, 10: 25- doi: 10.1186/1744-9081-10-25

Obrist, P.A. (1981). Cardiovascular psychophysiology: a perspective. New York: Plenum Press.

Ogoh, S., Fisher, J.P., Dawson, E.A., White, M.J., Secher, N.H., & Raven, P.B. (2005). Autonomic nervous system influence on arterial baroreflex control of heart rate during exercise in humans. *Journal of Physiology, 566*, 599–611. doi: 10.1113/jphysiol.2005.084541

Peňáz, J. (1973). Photoelectric measurement of blood pressure volume and flow in the finger. In A. Albert, W. Vogt, & W. Hellig (Eds). *Digest of the 10th International Conference on Medical and Biological Engineering.* Dresden, Germany: International Federation for Medical and Biological Engineering.

Preacher, K. J., Rucker, D. D., & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*, *42*(1), 185-227.

Radvansky, G., Copeland, D. E., & von Hippel, W. (2010). Stereotype activation, inhibition, and aging. *Journal of Experimental Social Psychology, 46,* 51-60. doi: 10.1016/j.jesp.2009.09.010

Ramshur, J. T. (2010). HRVAS: Heart Rate Variability Analysis Software. *University of Memphis, Department of Biomedical Engineering*.

Richter, M. (2010). Pay attention to your manipulation checks! Reward impact on cardiac reactivity is moderated by task context. *Biological Psychology, 84*(2), 279-289. doi: 10.1016/j.biopsycho.2010.02.014

Richter, M. (2012). Cardiovascular response to reward. In R. A. Wright, & G. H. E. Gendolla (Eds.), *How motivation affects cardiovascular response: Mechanisms and applications.* Washington, DC: American Psychological Association.

Richter, M. (2013). A closer look into the multi-layer structure of motivational intensity theory. *Social and Personality Psychology Compass, 7*, 1-12. doi: 10.1111/spc3.12007

Richter, M. (2016). Residual tests in the analysis of planned contrasts: Problems and solutions. *Psychological Methods*, *21*, 112–120. doi: 10.1037/met0000044

Richter, M., Baeriswyl, E., & Roets, A. (2012). Personality effects on cardiovascular reactivity: Need for closure moderates the impact of task difficulty on engagement-related myocardial beta-adrenergic activity. *Psychophysiology*, *49*(5), 704-707. doi: 10.1111/j.1469-8986.2011.01350.x

Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the analysis of variance*. Cambridge, UK: Cambridge University Press.

Rowell LB (1986). *Human Circulation Regulation during Physical Stress*. New York, NY: Oxford University Press.

Rowell, L.B. (1993). *Human Cardiovascular Control*. New York, NY: Oxford University Press.

Royall, R. M. (1997). *Statistical evidence. A likelihood paradigm*. London, UK: Chapman and Hall.

Schmalhofer, F., & Glavanov, D. (1986). Three components of understanding a programmer's manual: Verbatim, propositional, and situational representations. *Journal of Memory and Language, 25*, 279–294. doi:10.1016/0749-596X(86)90002-1

Schmidt, R. E., Richter, M., Gendolla, G. H., & Van der Linden, M. (2010). Young poor sleepers mobilize extra effort in an easy memory task: Evidence from cardiovascular measures. *Journal of Sleep Research*, *19*(3), 487-495. doi: 10.1111/j.1365-2869.2010.00834.x

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, *8*(4), 379-386.

Senderecka, M., Zabawa, J., Kluj-Kozłowska, K., Greń, M., Konkel, A., Kuklińska, M., Paprot, E., Sikorski, R., Barczak, A., & Sitek, E. (2015). *Polish version of Addenbrooke's Cognitive Examination – III.* Retrieved from: https://sydney.edu.au/brain-mind/resources-for-clinicians/dementia-test.html

Sherman, J. W., Macrae, C. N., & Bodenhausen, G. V. (2000). Attention and stereotyping: Cognitive constraints on the construction of meaningful social impressions. *European Review of Social Psychology*, *11*(1), 145-175.

Silvia, P. J., Beaty, R. E., Nusbaum, E. C., Eddington, K. M., & Kwapil, T. R. (2014). Creative motivation: Creative achievement predicts cardiac autonomic markers of effort during divergent thinking. *Biological Psychology, 102*, 30-37. doi: 10.1016/j.biopsycho.2014.07.010

Silvia, P. J., & Gendolla, G. H. (2001). On introspection and self-perception: Does self-focused attention enable accurate self-knowledge? *Review of General Psychology, 5*(3), 241-269. doi: 10.1037//1089-2680.5.3.241

Smith, B. T., & Hess, T. M. (2015). The impact of motivation and task difficulty on resource engagement: Differential influences on cardiovascular responses of young and older adults. *Motivation Science*, *1*(1), 22-36. doi: 10.1037/mot0000012

Smith, T. W., Uchino, B. N., Berg, C. A., Florsheim, P., Pearce, G., Hawkins, M., ... & Olsen-Cerny, C. (2009). Conflict and collaboration in middle-aged and older couples: II. Cardiovascular reactivity during marital interaction. *Psychology and Aging*, *24*(2), 274-286.

Stewart, B. D., von Hippel, W., & Radvansky, G. A. (2009). Age, race, and implicit prejudice. Using process dissociation to separate the underlying components. *Psychological Science, 20*, 164–168. doi: 10.1111/j.1467-9280.2009.02274.x

Stewart, C. C., Wright, R. A., & Griffith, H. R. (2016). Difficulty, effort and cardiovascular response to a working memory challenge: Older adults with and without mild cognitive impairment. *International Journal of Psychophysiology*, *104*, 53-61. doi: 10.1016/j.ijpsycho.2016.04.005

Szumowska, E., Szwed, P., Kossowska, M., Wright, R. (in press). Multifaceted Effects of Need for Cognitive Closure on Effort: A Cardiovascular Response Study. *Motivation Science.* Advance online publication. doi: 10.1037/mot0000053.

Taelman, J., Vandeput, S., Vlemincx, E., Spaepen, A., & Van Huffel, S. (2011). Instantaneous changes in heart rate regulation due to mental load in simulated office work. *European Journal of Applied Physiology*, *111*(7), 1497-1505. doi: 10.1007/s00421-010-1776-0

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal, 17,* 354–381. doi: 10.1161/01.CIR.93.5.1043

Thuraisingham, R. A. (2006). Preprocessing RR interval time series for heart rate variability analysis and estimates of standard deviation of RR intervals. *Computer Methods and Programs in Biomedicine*, *83*(1), 78-82. doi: 10.1016/j.cmpb.2006.05.002

Uchino, B. N., Birmingham, W., & Berg, C. A. (2010). Are older adults less or more physiologically reactive? A meta-analysis of age-related differences in cardiovascular reactivity to laboratory tasks. *Journal of Gerontology: Psychological Sciences, 65B*, 154–162. doi: 10.1093/geronb/gbp127

Uchino, B. N., Uno, D., Holt-Lunstad, J., & Flinders, J. B. (1999). Age-related differences in cardiovascular reactivity during acute psychological stress in men and women. *Journal of Gerontology: Psychological Sciences*, *54B*, 339 - 346. doi: 10.1093/geronb/54B.6.P339

Van Roon, A. M., Mulder, L. J. M., Althaus, M., & Mulder, G. (2004). Introducing a baroreflex model for studying cardiovascular effects of mental workload. *Psychophysiology, 41*, 961–981. doi:1 0.1111/j.1469-8986.2004.00251.x

Von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. *Advances in Experimental Social Psychology*, *27*, 177-254. doi: 10.1016/S0065-2601(08)60406-2

Von Hippel, W., Silver, L.A., & Lynch, M.E. (2000). Stereotyping against your will: The role of inhibitory ability in stereotyping and prejudice among the elderly. *Personality and Social Psychology Bulletin, 26*, 523–532. doi: 10.1177/0146167200267001

Westfall, J., Judd, C. M., & Kenny, D. A. (2015). Replicating studies in which samples of participants respond to samples of stimuli. *Perspectives on Psychological Science, 10*(3), 390-399. doi: 10.1177/1745691614564879

White, D. W., & Raven, P. B. (2014). Autonomic neural control of heart rate during dynamic exercise: revisited. *The Journal of Physiology*, *592*(12), 2491-2500. doi: 10.1113/jphysiol.2014.271858

Williams, A. F. (2003). Teenage drivers: patterns of risk. *Journal of Safety Research*, *34*(1), 5-15. doi: 10.1016/S0022-4375(02)00075-0

Wilson, T. C. (1996). Cohort and Prejudice: Whites' attitudes Toward Blacks, Hispanics, Jews, and Asians. *Public Opinion Quarterly*, *60*(2), 253-274. doi: 10.1086/297750

Wright, R. A. (1996). Brehm's theory of motivation as a model of effort and cardiovascular response. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 424–453). New York, NY: Guilford Press.

Wright, R.A., 1998. Ability perception and cardiovascular response to behavioral challenge. In: Kofta, M., Weary, G., Sedek, G. (Eds.), Control in Action: Cognitive and Motivational Mechanisms (pp. 197–232). Plenum, New York. doi: 10.1007/978-1-4757-2901-6_8.

Wright, R. A., & Kirby, L. D. (2001). Effort determination of cardiovascular response: An integrative analysis with applications in social psychology. *Advances in Experimental Social Psychology, 33*, 255–307.

Zieliński, P., & Biernacki, M. (2010). Analiza psychometryczna polskiego przekładu narzędzia do subiektywnej oceny obciążenia zadaniowego NASA-TLX [Psychometric Analysis of Polish Version of NASA Task-Load Index]. *Polski Przegląd Medycyny Lotniczej, 3(16),* 219 – 239.

Supplemental Materials

S1. Individual differences and affect measures

We measured participants' processing speed, NFC, and affect. We used a 15-item NFC scale (Webster & Kruglanski, 1994; Polish version of Kossowska, Hanusz, & Trejtowicz, 2012; Cronbach's $\alpha = .80$) on which participants rate the degree to which they agree with a presented item using a 6-point Likert scale (1 = “strongly disagree” to 6 = “strongly agree”). This scale measures participants’ preference for order, predictability, intolerance of ambiguity, and closed-mindedness (subscale of decisiveness was excluded from the analysis; see Roets & Van Hiel, 2008). To measure processing speed, we used the Digit–Symbol Substitution task from the Wechsler Adults Intelligence Scale – Revised (WAIS-R; Wechsler, 1981), which is a standard measure of processing speed in aging research (Salthouse, 1996). In the Digit–Symbol Substitution task, participants are presented with a list of nine digit–symbol pairs and are asked to write down the corresponding symbols under each digit in the provided list. After 7 practice trials, participants are given a time limit of 60 seconds to write down as many symbols as possible. The number of correct symbols was used as a measure of processing speed. Furthermore, we measured participants' affect with PANAS (Watson, Tellegen, & Clark, 1988; Polish version of Brzozowski, 2010), which is a 20-item measure in which participants rate their affect using a 5-point Likert scale (1 = “very weakly” to 5 = “very strongly”). Ten of the items assess positive affect; the remaining 10 items assess negative affect. In the current study, participants responded in PANAS twice: at the beginning and at the end of the study. Both positive and negative affect subscales during pre- and post-experimental assessment showed high reliability. Cronbach's alphas were the following for the pre-experimental measurement: $\alpha = .80$ for positive affect, $\alpha = .83$ for negative affect. For the post-experimental measurement, Cronbach's alphas were the following: $\alpha = .83$ for positive affect; $\alpha = .91$ for negative affect.

We ran a 2 (age: young vs. older adults) x 2 (self-involvement: low vs. high) ANOVA on a series of individual differences and affect measures. We found age-related differences for processing speed, $F(1,87) = 150.05$, $p < .001$, $\eta^2_p = 0.63$, with young adults having higher processing speed than older adults. We also found a main effect of self-involvement manipulation on processing speed, $F(1,87) = 5.06$, $p = .027$, $\eta^2_p = 0.06$, with people in the low-self-involvement condition having higher processing speed in comparison to people in the high-self-involvement condition. Furthermore, we found effects of age on NFC, $F(1,87) = 16.31$, $p < .001$, $\eta^2_p = 0.16$, with older adults scoring higher than young adults. These patterns of results are typical in aging research (e.g., Hess, 2001; Kossowska, Jaśko, Bar-Tal, & Szastok, 2012; Kościelniak, Rydzewska, & Sędek, 2016; Salthouse, 1996).

Furthermore, we found a main effect of self-involvement manipulation for negative affect measured at the beginning of an experimental session, $F(1,87) = 4.69$, $p = .033$, $\eta^2_p = 0.05$. This effect was qualified by the interaction between age and self-involvement manipulation, $F(1,87) = 6.69$, $p = .011$, $\eta^2_p = 0.07$. Post hoc comparisons with Sidak correction revealed differences among older adults, $F(1,87) = 11.16$, $p = .001$, $\eta^2_p = 0.11$, with older adults in the high-self-involvement condition reporting more negative affect in comparison to those in the low-self-involvement condition. There were no differences between the low- and high-self-involvement conditions among young adults, $F(1,87) = 0.09$, $p = .765$, $\eta^2_p < 0.01$. The mean scores of the individual differences and affect measures for young and older participants in the low and high-self-involvement conditions are presented in Table S1.

Table S1

Cell means and standard errors of individual differences and affect measures

		Young adults		Older adults	
Measure		Low self-involvement	High self-involvement	Low self-involvement	High self-involvement
Digit–Symbol Substitution	*M*	47.59	44.71	31.46	28.38
task	*SE*	1.34	1.42	1.49	0.84
Need for Closure	*M*	3.64	3.50	3.97	4.16
	SE	0.14	0.13	0.11	0.11
PANAS: positive affect	*M*	3.28	3.20	3.17	3.27
(pre)	*SE*	0.07	0.13	0.09	0.09
PANAS: negative affect	*M*	1.62	1.58	1.41	1.86
(pre)	*SE*	0.09	0.08	0.07	0.13
PANAS: positive affect	*M*	3.04	3.09	3.12	3.28
(post)	*SE*	0.12	0.10	0.11	0.09
PANAS: negative affect	*M*	1.69	1.50	1.61	1.90
(post)	*SE*	0.12	0.11	0.13	0.16

S2. Stereotypical wrong inference probes

Below we present stereotypical wrong probes that were used in measuring stereotyping tendency.

Story regarding the young drivers

1. Original sentence: *"Groups posing the greatest threat are young, inexperienced drivers and those over 70."* (greatest threat to road safety)

   Test probe: "*The group posing the greatest threat is young drivers.*"

2. Original sentence: *"Six km later he hit 15-year-old Kasia Bolko."* (the preceding sentence also contained information that an 18 year-old driver was driving at a speed of 80 km/h).

   Test probe: *"The cause of the accident was driving too fast*."

Story regarding the older drivers

1. Original sentence: *"The groups posing the greatest threat are young, inexperienced drivers and those over 70."* (greatest threat to road safety)

   Test probe: "*The group posing the greatest threat is drivers older than 70.*"

2. Original sentence: "*Six km later she hit 15 year-old Kasia Bolko."* (the preceding sentence also contained information that a 70-year-old lady was driving her car).

   Test probe: *"The cause of the accident was the poor vision of the older woman*."

S3. Stereotyping tendency: additional analysis

We repeated the analysis of the stereotyping tendency using control variables. As mentioned in the body of this article, after inclusion of the proportion of *yes*-responses to neutral wrong probes, the planned contrast remained significant, $F(1,86) = 15.81$, $MSE = 0.08$, $p < .001$, $\eta^2 = 0.16$. The LR remained high, $\lambda = 480.09$, and supports the hypothesized pattern of stereotyping tendency (see Figure 1). The caveat of this analysis is that due to an error in the programming of the procedure, some of the participants were presented with only one neutral wrong probe instead of two (which makes the assessment less reliable). Furthermore, after inclusion of another control variable (the proportion of *yes*-responses to stereotypical correct inferences), the contrast for the stereotyping tendency remained significant, $F(1,86) = 15.54$, $MSE = 0.08$, $p < .001$, $\eta^2 = 0.15$, and LR supported our hypothesized pattern of stereotyping tendency, $\lambda = 445.23$. These results provide very strong evidence for the hypothesized joint impact of age and self-involvement manipulation on stereotyping tendency after controlling for neutral wrong inferences or stereotypical correct inferences.

Additionally, we present analysis of responses to stereotypical probes other than the stereotypical wrong inferences that were our focus in the main analysis. For each analysis, we use responses to neutral inferences at the same level as the control variable, e.g., for the analysis of responses to verbatim stereotypical probes we enter responses to verbatim neutral probes as a covariate. For the verbatim stereotypical probes, the planned contrast was insignificant, $F(1,86) = 2.41$, *MSE* = 0.04, *p* = 0.062, η^2 = 0.03, and LR shows that the hypothesized pattern is not supported, λ = 1.18. For the stereotypical paraphrase probes, the planned contrast was significant, *F*(1,86) = 10.19, *MSE* = 0.13, *p* = 0.001, η^2 = 0.11, and LR supported the hypothesized pattern of results, λ = 38.62. Similarly for the stereotypical correct inferences, the planned contrast was significant, *F*(1,86) = 10.76, *MSE* = 0.09, *p* = 0.001, η^2 = 0.11, and LR supported the hypothesized pattern of results λ = 73.12. The effects for paraphrases and correct inference were, however, weaker than the effects for the wrong inferences. This suggests that the more the original probe is altered, the larger the joint impact of age and self-involvement. The average scores for each type of probe, both neutral and stereotypical, are presented in Table S2.

For more direct comparisons between the proportion of *yes*-responses to neutral vs. stereotypical wrong probes, we ran a repeated-measures ANOVA: 2 (age: young vs. old) x 2 (self-involvement: low vs. high) x 2 (type of a probe: neutral vs. stereotypical). We found a 3-way interaction between age, self-involvement, and the type of probe, *F*(1,87) = 27.88, *p* < .001, η^2 = .24. Pair-wise comparisons (with Sidak correction) showed that within the low-self-involvement condition, older adults had higher proportions of *yes*-responses to stereotypical probes in comparison to young adults, *F*(1,87) = 22.13, *p* < .001, η^2 = .20, but the difference between young and older people for neutral probes was insignificant, *F*(1,87) = 2.56, *p* = .113, η^2 = .029. In contrast, in the high-self-involvement condition, the difference between young and older people was insignificant for the stereotypical probes, *F*(1,87) = 0.03, *p* = .858, η^2 < .01, but older adults were more likely to accept neutral wrong probes in comparison to young people, *F*(1,87) = 20.50, *p* < .001, η^2 = .19.

Table S2

Cell means and standard errors of the responses to probes presented in the text comprehension task

			Neutral				Stereotypical			
Age	Self-involvement		Verbatim	Paraphrases	Correct Inferences	Wrong Inferences	Verbatim	Paraphrases	Correct Inferences	Wrong Inferences
Young	Low	*M*	0.81	0.73	0.32	0.27	0.69	0.41	0.34	0.11
		SE	0.03	0.07	0.12	0.10	0.04	0.08	0.07	0.05
	High	*M*	0.79	0.67	0.04	0.04	0.76	0.67	0.31	0.27
		SE	0.03	0.07	0.03	0.03	0.04	0.07	0.07	0.06
Older	Low	*M*	0.76	0.58	0.33	0.10	0.80	0.77	0.56	0.50
		SE	0.04	0.07	0.06	0.04	0.04	0.07	0.05	0.07
	High	*M*	0.81	0.60	0.90	0.52	0.72	0.43	0.36	0.29

SE	0.04	0.07	0.15	0.11	0.04	0.08	0.06	0.06

S4. Cardiovascular responses: DBP and HR

For ease of comprehension, in the main analyses of effort mobilization we presented only the results for SBP and nHF-HRV as we had specific hypotheses for these measures. Below we present analysis of DBP and HR baseline and task change scores.

Baseline periods. Resting cardiovascular baseline scores are presented in Table S3. We submitted baseline scores to a 2 (age: young vs. older adults) x 2 (self-involvement: low vs. high) ANOVA. There was a significant effect of age on HR in the first, $F(1,87) = 17.96$, $p < .001$, $\eta^2_p = .17$, and in the second baseline, $F(1,87) = 20.98$, $p < .001$, $\eta^2_p = .19$, with older adults having lower HR baseline scores than young adults. For the HR measure there was also an interaction effect between age and condition in the first baseline period, $F(1,87) = 4.59$, $p = .035$, $\eta^2_p = .05$. Post hoc comparisons with Sidak correction showed that there were differences between young and older adults in the high-self-involvement condition, $F(1,87) = 20.11$, $p < .001$, $\eta^2_p = .19$, but not in the low-self-involvement condition, $F(1,87) = 2.23$, $p = .149$, $\eta^2_p = .03$. A similar interaction occurred in the second baseline period for HR, $F(1,87) = 4.13$, $p = .045$, $\eta^2_p = .05$, with young and older adults differing from each other in the high-self-involvement condition, $F(1,87) = 21.59$, $p < .001$, $\eta^2_p = .20$, but not in the low-self-involvement condition, $F(1,87) = 3.29$, $p = .073$, $\eta^2_p = .04$. There were no effects of age or experimental condition on DBP. Correlation between baseline measures and respective change scores emerged between the first baseline and DBP change score in the story period, $r(90) = -.20$, $p = .056$, and between the second baseline and HR change scores in the test period, $r(90) = -.33$, $p = .002$. Because of the significant differences in several baseline scores and correlations between baseline and reactivity values, we included baseline scores as covariates in the analysis of DBP and HR.

Table S3

Cell means and standard errors of the DBP and HR baseline scores

			Baseline 1		Baseline 2	
Age	Self-involvement		DBP	HR	DBP	HR
Young	Low	*M*	69.69	78.01	70.47	78.33
		SE	2.12	1.91	2.23	1.87
	High	*M*	68.30	80.53	69.40	81.11
		SE	1.55	1.55	1.38	1.63
Older	Low	*M*	66.79	74.08	68.49	73.41

	SE	1.29	2.07	1.44	2.12
High	*M*	65.43	68.54	66.15	68.36
	SE	1.86	1.94	1.68	2.06

Note. DBP values are in mmHg; HR values are in bpm.

Story period. The planned contrast was significant for DBP change scores, $F(1,86) = 2.93$, $MSE = 12.37$, $p = 0.045$, $\eta^2_p = 0.03$, but insignificant for HR, $F(1,86) < 0.01$, $MSE = 5.87$, $p = 0.487$, $\eta^2_p < 0.01$, during the story period. Furthermore, comparing the evidence of the null hypothesis against the evidence for the hypothesized pattern resulted in corrected LR: $\lambda = 1.15$ for DBP and $\lambda = 0.34$ ($1/\lambda = 2.92$) for HR. This provides no evidence in favor of the joint impact of age and self-involvement on effort mobilization against the null model in the story period for DBP and HR. Cell means and standard errors of cardiovascular change scores in the story period can be found in Table S4.

Table S4

Cell means and standard errors of DBP and HR change scores in the story and test period

			Story period		Test period	
Age	Self-involvement		DBP	HR	DBP	HR
Young	Low	*M*	0.97	0.98	0.96	-0.53
		SE	0.83	0.56	0.67	1.02
	High	*M*	1.17	0.82	1.53	-0.66
		SE	0.68	0.52	0.61	0.62
Older	Low	*M*	5.53	2.89	5.58	1.65
		SE	0.82	0.56	0.83	0.63
	High	*M*	4.25	1.56	5.17	0.62
		SE	0.59	0.31	0.92	0.58

Note. DBP values are in mmHg; HR values are in bpm.

Test period. The planned contrast was significant for DBP, $F(1,86) = 6.66$, $MSE = 13.24$, $p = 0.005$, $\eta^2_p = 0.07$, but insignificant for HR change scores, $F(1,86) = 0.31$, $MSE = 11.31$, $p = 0.291$, $\eta^2_p < 0.01$. Comparing the evidence of the null hypothesis against the evidence of the predicted pattern resulted in corrected LR: $\lambda = 5.30$ for DBP and $\lambda = 0.34$ ($1/\lambda = 2.91$) for HR. Thus, there is weak evidence in favor of our hypothesis in comparison to the null model for DBP and no evidence for HR. Cell means and standard errors of cardiovascular change scores in the test period can be found in Table S4.

S5. Cardiovascular responses in full test period

In the main analyses of effort mobilization for the test period, we presented the results calculated for the first 150 seconds of the test period. Below we present analysis of all cardiovascular indices, i.e., SBP, nHF-HRV, DBP and HR change scores calculated for the whole test period. For this analysis we subtracted task scores from the last 150 seconds of the baseline period (the baseline scores were included as covariates).

The planned contrasts were significant for SBP $F(1,86) = 9.27$, $MSE = 68.09$, $p = .002$, $\eta^2_p = 0.10$, and nHF-HRV change scores in the test period, $F(1,86) = 9.2$, $MSE = 0.02$, $p = .002$, $\eta^2_p = 0.10$. The contrasts were not significant for DBP, $F(1,86) = 5.28$, $MSE = 12.75$, $p = .012$, $\eta^2_p = 0.06$, nor for HR, $F(1,86) = 0.28$, $MSE = 10.85$, $p = .300$, $\eta^2_p < .01$. Comparing the evidence for the null hypothesis against the evidence of the predicted pattern resulted in corrected LR: $\lambda = 14.69$ for SBP and $\lambda = 31.17$ for nHF-HRV. This means that the data is 14 times more likely under our hypothesized model than under the null hypothesis model for SBP and 31 times more likely for nHF-HRV. This provides moderate evidence in favor of the joint impact of age and self-involvement on effort mobilization for SBP and nHF-HRV. However, the evidence for nHF-HRV is twice as large as the evidence for SBP. Furthermore, the current SBP findings are subject to the same limitations as the SBP results described in the main text, i.e., the evidence for the moderating role of self-involvement should be described as weak. Finally the evidence for DBP and HR was weak: only $\lambda = 3.01$ for DBP and $\lambda = 0.39$ ($1/\lambda = 2.53$) for HR. Cell means and standard errors of cardiovascular change scores in the test period can be found in Table S5.

Table S5

Cell means and standard errors of SBP, nHF-HRV, DBP, and HR change scores in the full test period

Age	Self-involvement		SBP	nHF	DBP	HR
	Low	*M*	1.73	0.06	0.81	-0.34
Young		*SE*	1.41	0.04	0.65	0.99
	High	*M*	4.15	0.06	1.55	-0.49
		SE	1.33	0.03	0.60	0.62
Older	Low	*M*	12.58	-0.02	5.49	1.77
		SE	1.98	0.04	0.80	0.61
	High	*M*	12.06	-0.08	4.79	0.70
		SE	2.13	0.05	0.92	0.55

Note. Blood pressure values are in mmHg; nHF-HRV values are in n.u.; HR values are in bpm.

S6. Subjective ratings

Table S6. below presents cell means and standard errors for subjective ratings of tasks for the story and test periods that were not presented in the main analyses.

Table S6

Cell means and standard errors of the workload and task perception ratings in the story and test periods

			Young adults		Older adults	
			Low self-involvement	High self-involvement	Low self-involvement	High self-involvement
Physical demand	Story	*M*	0.23	0.13	0.35	0.31
		SE	0.09	0.06	0.11	0.08
	Test	*M*	0.27	0.29	0.19	0.45
		SE	0.10	0.12	0.08	0.20
Temporal demand	Story	*M*	2.68	3.36	3.50	2.71
		SE	0.40	0.36	0.44	0.42
	Test	*M*	2.73	1.88	3.30	2.78
		SE	0.38	0.28	0.45	0.40
Performance	Story	*M*	1.78	1.67	2.08	1.93
		SE	0.24	0.22	0.27	0.26
	Test	*M*	2.32	2.14	2.64	2.65
		SE	0.31	0.27	0.30	0.31
Frustration	Story	*M*	1.15	0.99	0.48	0.61
		SE	0.30	0.27	0.17	0.26
	Test	*M*	1.40	1.23	1.25	1.22
		SE	0.33	0.33	0.36	0.37
Tiredness	Story	*M*	8.77	7.83	5.83	5.05
		SE	1.01	1.20	0.87	0.92
	Test	*M*	10.05	7.83	9.58	9.45
		SE	0.89	1.00	1.28	1.14
Success probability	Test	*M*	13.55	14.54	13.33	12.50
		SE	0.96	0.87	1.22	1.18
Experimenter anxiety	Test	*M*	2.82	2.92	2.46	3.30
		SE	0.61	0.61	0.56	0.58
Test results anxiety	Test	*M*	11.41	7.79	11.04	12.15
		SE	1.19	1.06	1.54	1.30