Accompanying Materials for "On the Interpretation of Uninterpretable Interactions: A Survey of the Field 32 Years after Loftus." B. Literature Review

Eric-Jan Wagenmakers[1], Angelos-Miltiadis Krypotos[1], Amy Criss[2], and Geoff Iverson[3]

[1] University of Amsterdam

[2] Syracuse University

[3] University of California at Irvine

Correspondence concerning this article should be addressed to:

Eric-Jan Wagenmakers

University of Amsterdam, Department of Psychology

Roetersstraat 15

1018 WB Amsterdam, The Netherlands

Ph: (+31) 20–525–6420

Fax: (+31) 20-639-0279

E-mail may be sent to EJ.Wagenmakers@gmail.com.

Document's Description

This document includes the plots and the coresponding statistics of all 2×2 statistically significant ($p<0.05$) interactions reported in Psychology and Aging for the year 2008. If the 2×2 interactions were modified by a 2×3 interaction which was emphasized by the authors, then the 2×2 interactions were not plotted. We also did not include interactions that served as predictors in regression analyses. For all papers that did not satisfy our inclusion criteria, please see the last section of this document.

Each interaction was classified in one of the following categories: 1) Strongly Interpretable, 2) Strongly Uninterpretable, 3) Weakly Interpretable, 4) Strongly Interpretable–Vague, 5) Strongly Uninterpretable–Vague, 6) Weakly Interpretable–Vague.

The first category concerns qualitative interactions whereas the second one quantitative. The third one concerns both quantitative and qualitative interactions with at least one non–significant main effect. When post–hoc tests were reported, we classified interactions based on those categories. When post–hoc test were not reported, we classified interactions based on visual inspection of the data. For those cases, we used the same first three categories but included the term 'Vague' to reflect the absence of the corresponding tests.

The final classification of each interaction is included in the corresponding figures.

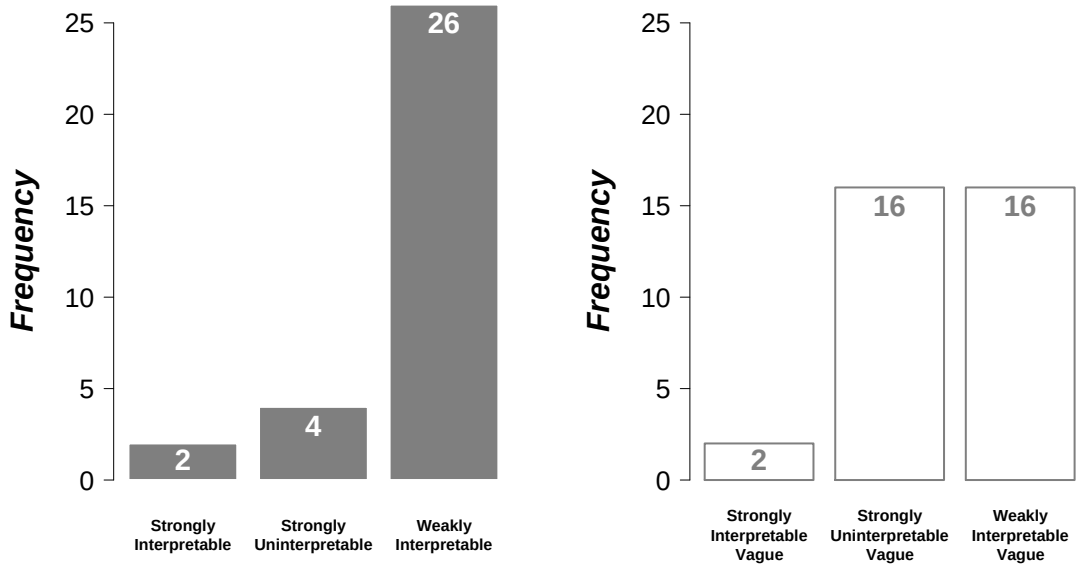

Figure 1. Number of interactions for each category.

Papers with statistically significant interactions.

Issue 1

Lisa Emery, Thomas M. Hess (p. 2–12)

Although the present paper reports many interactions, most of them are 2×3 because the variable of valence, which has three levels, is included.

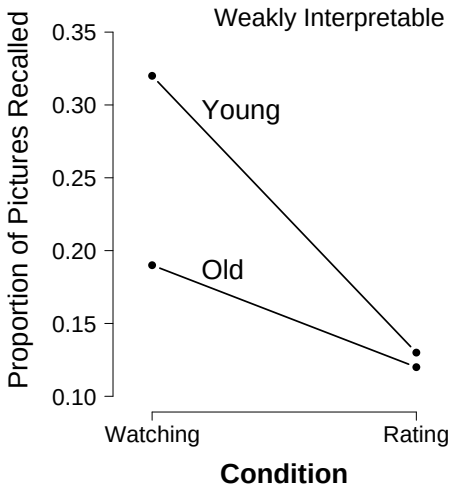

Figure 2. Age group × viewing condition.

Comments on the Interaction of Figure 2.

The values were extracted from the top section of Table 2 (Emery , & Hess, 2008, p. 7). In order to plot the values, we computed the mean of the three levels of valence for the corrected recognition scores.

Authors' Description of the Statistics.

"... Age Group × Viewing Condition interaction, F(1, 107) = 4.34, p = .04, $\eta_p^2 = .04$." (Emery & Hess, 2008, p. 6).

"... To clarify the two–way interactions, we conducted separate analyses of viewing condition and valence in each age group. These analyses indicated that the described effect of valence was only significant in the older adults, F(2, 106) = 13.22, p < .001, $\eta_p^2 = .20$, and the effect of viewing condition was only significant in the young adults, F(1, 54) = 15.45, p < .001, $\eta_p^2 = .22$." (Emery & Hess, 2008, p. 7).

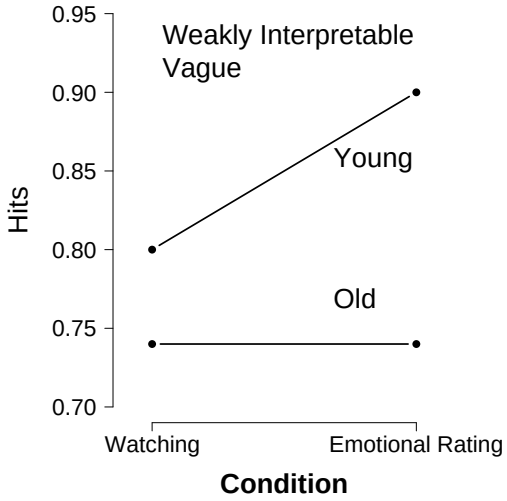

Figure 3. Age × condition interaction on hits

Comments on the Interaction of Figure 3.

The values were extracted from the top section of Table 2 (Emery & Hess, 2008, p. 7).

Authors' Description of the Statistics.

> " It should be noted, however, that the general effect of condition on the hit rates mirrors the pattern of the corrected recognition results, although the Age × Condition interaction on the hits is only significant in Experiment 1, F(1, 107)= 7.20, p = .008, η_{p^2} = .06 This effect does not further interact with valence, F(2, 214) = .01, p = .99, η_{p^2} = .00 . . . " (Emery & Hess, 2008, p. 9).

Sunghan Kim, M. Karl Healey, David Goldstein, Lynn Hasher, Ursula J. Wiprzycka (p. 33–38)

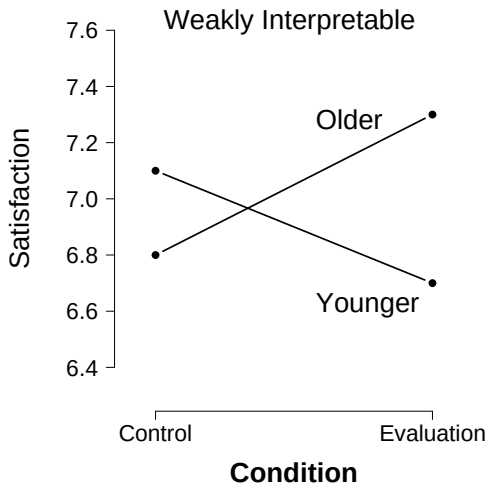

Figure 4. Age × condition interaction for satisfaction (liking) ratings–Initial.

Comments on the Interaction of Figure 4.

The values were extracted from Figure 1 (Kim et al., 2008, p. 36). Although the lines cross, the interaction is classified as weakly interpretable due to the corresponding post–hoc tests.

Authors' Description of the Statistics.

> "We conducted a 2 (age: younger, older) × 2 (condition: control, evaluation) between–subjects analysis of variance (ANOVA) on the initial liking ratings. The main effect of age, $F(1, 197) = 4.10$, $p < .05$, was qualified by a significant interaction between age and condition, $F(1, 197) = 6.56$, $p < .05$. There was no age difference in the control condition, $t(99) = 0.37$, whereas in the evaluation condition, older adults showed greater satisfaction than did younger adults, $t(98) = 3.30$, $p < .01$. Older adults in the evaluation condition gave higher ratings than did older control participants, $t(101) = 2.18$, $p < .05$, but evaluation did not influence satisfaction for younger adults, $t(96) = 1.42$, $p = .16$." (Kim et al., 2008, p. 35).

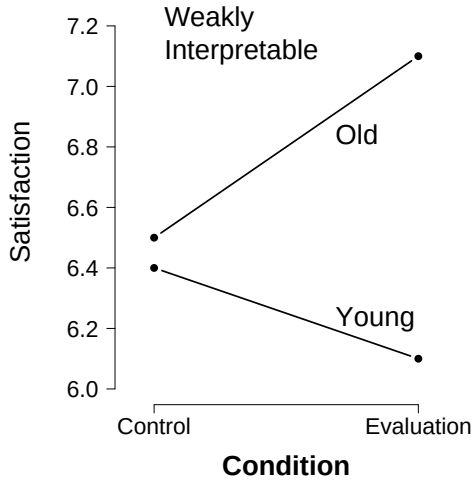

Figure 5. Age × condition interaction for satisfaction (liking) ratings–2 weeks later.

Comments on the Interaction of Figure 5.

The values were extracted from Figure 1 (Kim et al., 2008, p. 36).

Authors' Description of the Statistics.

"We conducted a parallel ANOVA with actual satisfaction after 2 weeks as the dependent measure. Again, there was the main effect of age, $F(1, 170) = 6.07$, $p < .05$, qualified by the Age × Condition interaction, $F(1, 170) = 4.23$, $p < .05$. Planned comparisons revealed that after 2 weeks, older adults in the evaluation condition remained more satisfied than did younger adults, $t(87) = 3.30$, $p < .01$, and there were still no age differences in the control condition, $t(83) = 0.28$." (Kim et al., 2008, p. 35).

Comments on the Interaction of Figure 6.

The values were extracted from Figure 1 (Kim et al., 2008, p. 36). In order to plot the interaction, we computed the mean of the control and evaluation condition for each rating time (Control and Evaluation).

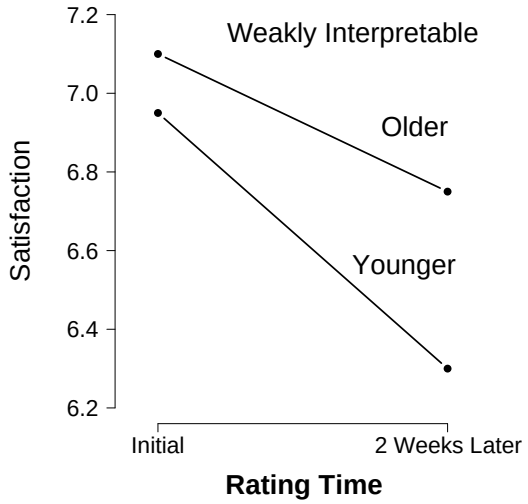

Figure 6. Age × rating time interaction for future satisfaction prediction.

Authors' Description of the Statistics.

"This analysis yielded the significant main effect of rating time, $F(1, 197) = 5.84$, $p < .05$, which was qualified by the significant interaction between rating time and age, $F(1, 197) = 7.79$, $p < .01$. Younger adults predicted that their satisfaction would decrease significantly, $F(1, 96) = 11.66$, $p < .001$, but older adults did not predict a significant change, $F(1, 101) = 0.08$." (Kim et al., 2008, p. 35).

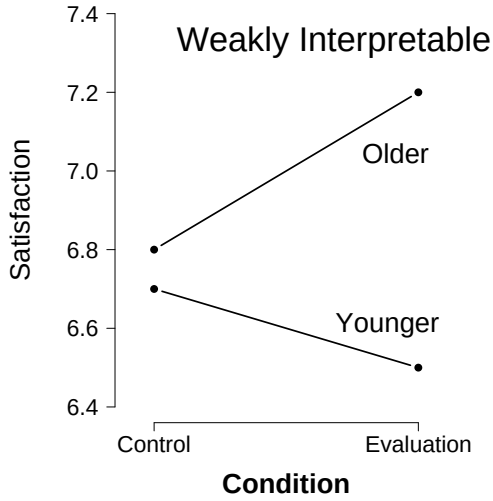

Figure 7. Age × condition interaction for future satisfaction prediction.

Comments on the Interaction of Figure 7.

The values were extracted from Figure 1 (Kim et al., 2008, p. 36). In order to plot the interaction, we computed the mean of each rating time (Initial, 2 Weeks Later) for each condition (Control, Evaluation). This interaction was reported in a footnote.

Authors' Description of the Statistics.

> "... qualified by the significant interaction between age and condition, $F(1, 197) = 4.90$, $p < .05$. Further analyses showed no age difference in the control condition (*initial condition on the graph*), $F(1, 99) = 0.58$, and a significant age difference in the evaluation condition (*two weeks later on the graph*), $F(1, 98) = 15.47$, $p < .001$." (Kim et al., 2008, p. 35).

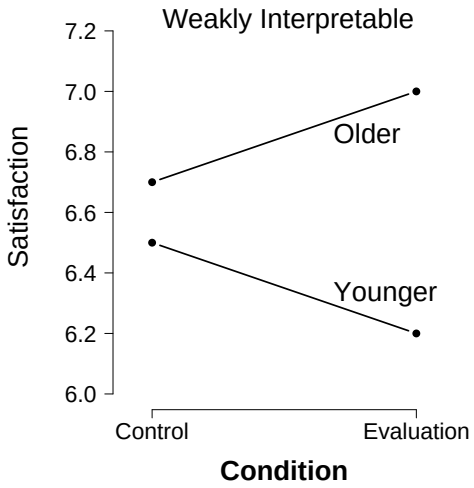

Figure 8. Age × condition interaction for the 'Predicted' versus 'Actual liking after two weeks', rating time.

Comments on the Interaction of Figure 8.

The values were extracted from Figure 1 (Kim et al. 2008, p. 36). In order to plot the interaction, we computed the mean values of the 'predicted' and the 'actual satisfaction after 2 weeks' values.

Authors' Description of the Statistics.

> "...and the interaction between age and condition, $F(1, 170) = 4.90$, $p < .05$, mirroring the patterns found when actual satisfaction after 2 weeks was analyzed in isolation. As with all other analyses of the various satisfaction ratings, there was the age effect in the evaluation condition, $F(1, 87) = 19.38$, $p < .0001$, but not in the control condition, $F(1, 83) = 1.71$, $p = .20$." (Kim et al., 2008, p. 35).

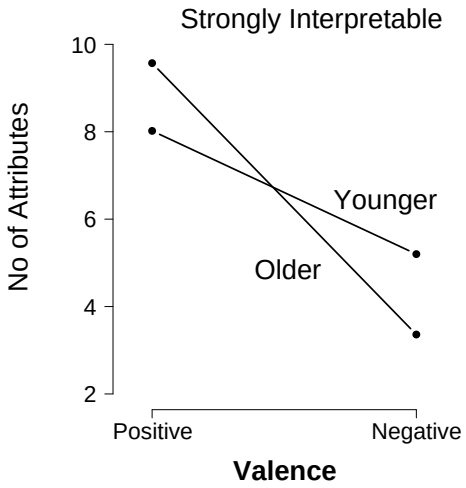

Figure 9. Age × valence interaction for number of attributes.

Comments on the Interaction of Figure 9.

The values were extracted from Table 1 (Kim et al. (2008), p. 36).
Authors' Description of the Statistics.

"There was, however, the significant effect of valence, $F(1, 100) = 91.00$, $p < .01$, indicating that both age groups listed more positive attributes than negative attributes. It is critical that there was the significant interaction, $F(1, 100) = 12.83$, $p < .01$, which can be attributed to the fact that whereas older adults produced significantly more positive attributes than did younger adults, $F(1, 100) = 4.33$, $p < .05$, older adults produced significantly fewer negative attributes, $F(1, 100) = 13.04$, $p < .01$." (Kim et al., 2008, p. 36).

Gillian Rowe, Lynn Hasher, Josée Turcotte (p. 79–84)

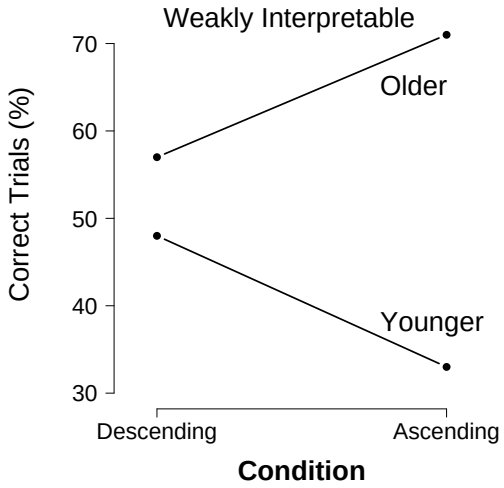

Figure 10. Age × condition interaction (experiment 1).

Comments on the Interaction of Figure 10.

The values were extracted from the text (please see below). This interaction was classified as 'Strongly Interpretable' although there was a marginal significant post–hoc test ($t(34) = 1.94$, $p = .06$).

Authors' Description of the Statistics.

> "A 2 (age) × 2 (condition) analysis of variance (ANOVA) on these scores showed young adults had higher span scores than older adults, $F(1, 65) = 38.31$, $MSE = 3.71$, $p < .01$, Ms = 65 and 40, respectively. Age interacted with condition, $F(1, 65) = 5.99$, $p = .02$, with older adults performing significantly better in the descending ($M = 48$, $SD = 22$) than in the ascending ($M = 33$, SD= 16) condition, $t(31) = 2.07$, $p < .05$, and, somewhat surprisingly, young adults showing the opposite pattern, that is, marginally poorer span scores in the descending ($M = 57$, $SD = 20$) than in the ascending ($M = 71$, $SD = 16$) condition, $t(34) = 1.94$, $p = .06$ (see Figure 2A). Note that despite the benefit that older adults received

from the descending version, their overall scores in that condition were still reliably lower than those of the comparable group of young adults, t(31) = 2.56, p = .02." (Rowe, Hasher, & Turcotte, 2008, p. 81).

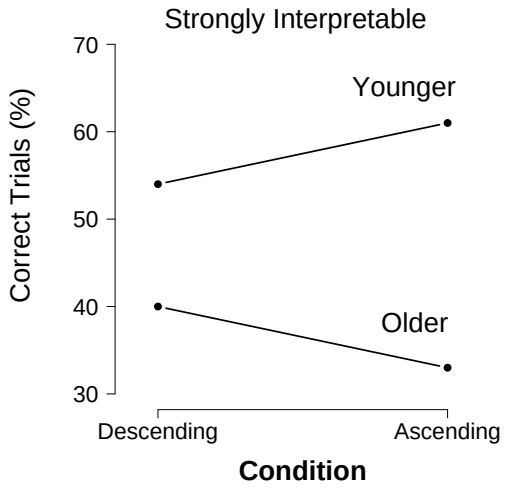

Figure 11. Age × condition interaction (experiment 2).

Comments on the Interaction of Figure 11.

The values were extracted from the text (please see below).
Authors' Description of the Statistics.

"Age again interacted with condition; F(1, 107) = 9.08, p < .01, with older adults performing significantly better in the descending (M = 40, SD = 8) than in the ascending (M = 33, SD = 10) condition, t(50) = 2.79, p < .01, and with young adults again showing the opposite pattern: reliably poorer span scores in the descending (M = 54, SD = 13) than in the ascending (M = 61, SD = 10) condition, t(57) = 2.03, p < .05 (see Figure 2B). Note that, as in Experiment 1, despite the improvement in the performance of older adults and the decline in the performance of young adults in the descending, as opposed to the ascending, condition,

the age difference in the descending condition remained significant, $t(54) = 4.87$, $p < .01$." (Rowe, Hasher, & Turcotte, 2008, p. 82).

Michael Ross, Steven J. Spencer, Craig W. Blatz,
Elaine Restorick (p. 85–92)

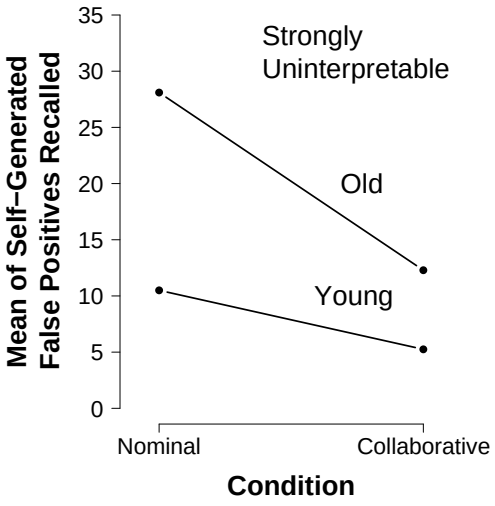

Figure 12. Age × condition (self–generated false positives rating).

Comments on the Interaction of Figure 12.

The values were extracted from Table 1 (Ross, Spencer, Blatz, & Restorick, 2008, p. 88).

Authors' Description of the Statistics.

> "The Age × Condition interaction was significant, F(1, 60) = 9.93, p < .01, partial η^2 = 0.14. (see Table 1 for means). Older couples listed significantly more self–generated false positives than younger couples in both the collaborative, F(1, 60) = 9.13, p < .01, partial η^2 = 0.47., and nominal, F(1, 60) = 53.54, p < .01, partial η^2 = 0.52 , group conditions. However, the decrease in self–generated false positives in the collaborative condition was greater for older, F(1, 60) = 44.44, p < .01, partial η^2 = 0.51., than for younger couples, F(1, 60) = 4.92, p < .03, partial η^2 = 0.22." (Ross, Spencer, Blatz, & Restorick, 2008, p. 88).

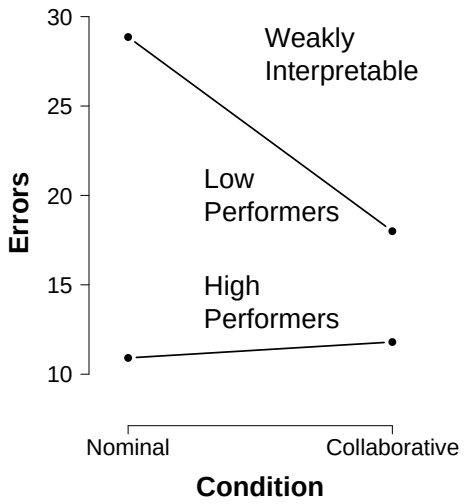

Figure 13. Age × condition (self–generated false positives rating).

Comments on the Interaction of Figure 13.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"High performers evidenced a similar number of false positives in the nominal (M = 10.91, SD = 6.41) and collaborative (M = 11.80, SD = 5.29) conditions (F < 1). In contrast, low performers made many more errors in the nominal (M = 28.86, SD = 10.69) than in the collaborative (M = 18.00, SD = 6.36) condition, F(1, 60) = 17.28, p < .01, partial $\eta^2 = 0.30$. When age was included as a factor in this analysis, it did not qualify the interaction (F < 1)." (Ross, Spencer, Blatz, & Restorick, 2008, p. 88).

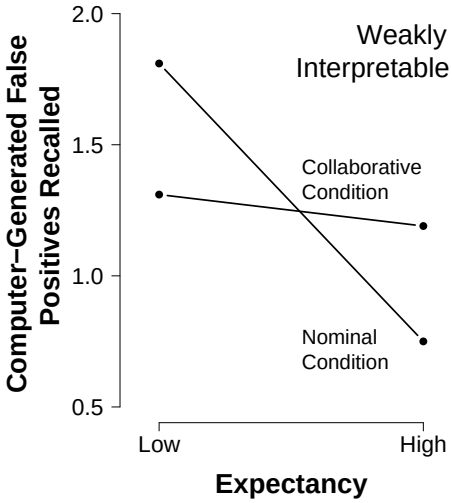

Figure 14. Condition × expectacy–younger group.

Comments on the Interaction of Figure 14.

The values were extracted from Table 2 (Ross, Spencer, Blatz, & Restorick, 2008, p. 89).

Authors' Description of the Statistics.

"The main effect of condition was nonsignificant ($F < 1$), but the Condition × Expectancy interaction was significant, $F(1, 60) = 5.97$, $p = .02$, partial $\eta^2 = 0.25$. (see Table 2 for means). In the nominal group condition, younger couples listed significantly more high–expectancy than low–expectancy computer–initiated false positives, $F(1, 60) = 15.33$, $p < .01$, partial $\eta^2 = 0.67$.; this effect of expectancy was eliminated in the collaborative condition ($F < 1$). The reduction in the expectancy effect in the collaborative condition was due to a slight increase in false positives in the low–expectancy condition and a slight decrease in false positives in the high–expectancy condition; neither of these trends approached significance ($ps > .10$)." (Ross, Spencer, Blatz, & Restorick, 2008, p. 89).

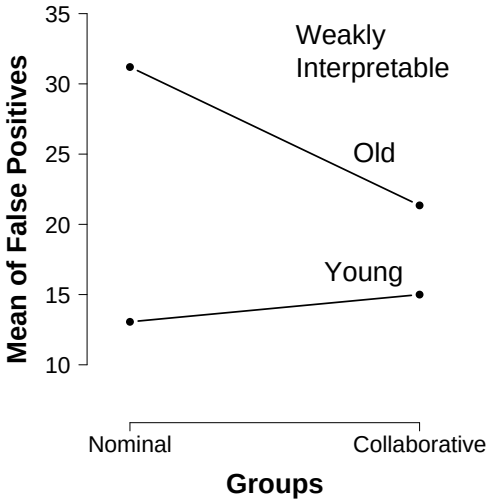

Figure 15. Age × condition (memory talk in collaborative groups).

Comments on the Interaction of Figure 15.

The values were extracted from table 3 (Ross, Spencer, Blatz, & Restorick, 2008. p. 90).

Authors' Description of the Statistics.

> "When we compared the number of erroneous suggestions during discussion in collaborative groups with the number of memory errors reported in nominal groups, the Condition × Age interaction was significant, F(1,60) = 9.18, p < .01, partial η^2 = 0.13. Only older couples inhibited errors. Older couples discussed fewer erroneous answers than older nominal groups reported, F(1, 60) = 12.79, p < .01, partial η^2 = 0.24. In contrast, younger couples discussed nonsignificantly more wrong answers (F < 1) than younger nominal groups reported." (Ross, Spencer, Blatz, & Restorick, 2008, p. 89).

Melanie Cohn, Stephen M. Emrich, Morris Moscovitch
(p. 93–103)

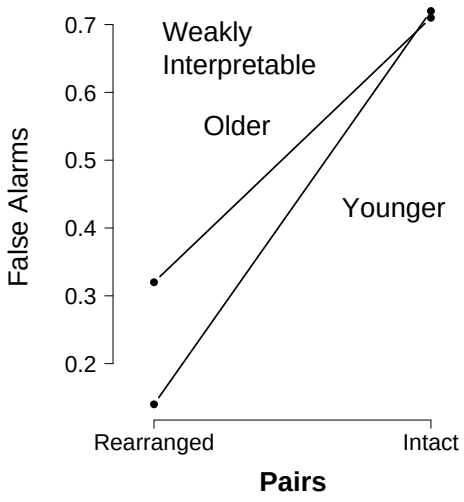

Figure 16. Groups × pairs.

Comments on the Interaction of Figure 16.

The values were extracted from Table 1–last two columns (Cohn, Emrich, & Moscovitch, 2008, p. 95).

Authors' Description of the Statistics.

> "As shown in Table 1, older adults false alarm rate to rearranged pairs was significantly greater than that of younger adults, $t(46) = 4.26$, $p < .001$, $d = 1.27$, whereas there was no age–related difference in hit rates to intact pairs, $t < 1$, $d = 0.10$. These data suggested that older adults difficulties in discriminating between studied and novel associations on an associative recognition task were due primarily to their poor recall–to–reject ability and less so to their recall–to–accept ability: interaction, $F(1, 46) = 7.23$, $p < .01$, partial $\eta^2 = 0.14$." (Cohn, Emrich, & Moscovitch, 2008, p. 97).

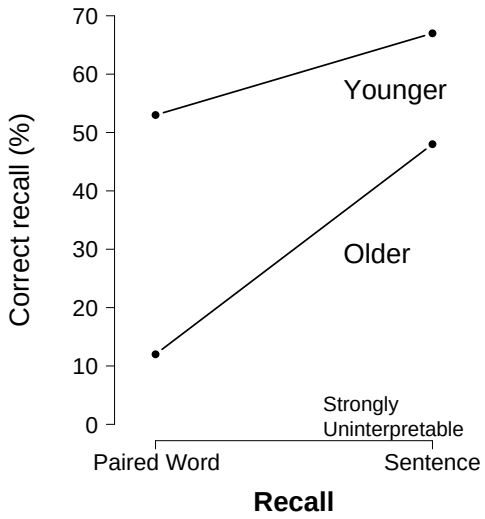

Figure 17. Age × recall.

Comments on the Interaction of Figure 17.

The values were extracted from Figure 4 (Cohn, Emrich, & Moscovitch, 2008, p. 99).

Authors' Description of the Statistics.

> "As shown in Figure 4, age–related impairments were found on both the paired–word recall, $t(46) = 6.71$, $p < .001$, $d = 2.00$; and sentence recall measures, $t(46) = 2.40$, $p < .05$, $d = 0.70$; but the paired–word recall was significantly more compromised than was the sentence recall: interaction, $F(1, 46) = 27.31$, $p < .001$, partial $\eta^2 = .37$." (Cohn, Emrich, & Moscovitch, 2008, p. 99).

Issue 2

Matthew G. Rhodes, Alan D. Castel, Larry L. Jacoby
(p. 239–249)

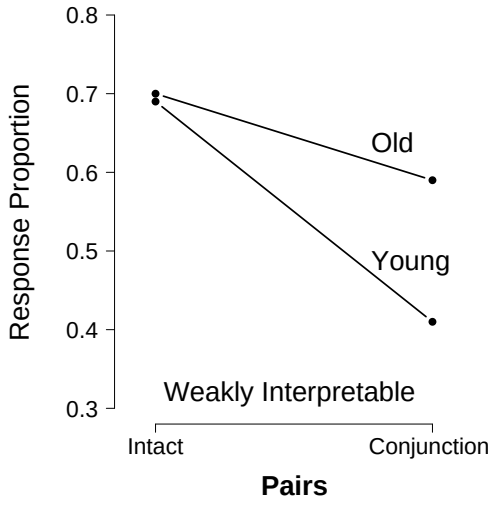

Figure 18. Age group × item type (experiment 1).

Comments on the Interaction of Figure 18.

The values were extracted from the text (please see below).
Authors' Description of the Statistics.

"Follow–up tests confirmed that older adults (M =.59) were significantly more likely to endorse conjunction pairs than were younger adults (M = .41), F(1, 57) = 13.32, η_p^2 =.19. In contrast, older (M = .70) and younger (M = .69) adults endorsed intact pairs approximately equally often (F < 1). No other reliable interactions were evident." (Rhodes, Castel, & Jacoby, 2008, p. 241).

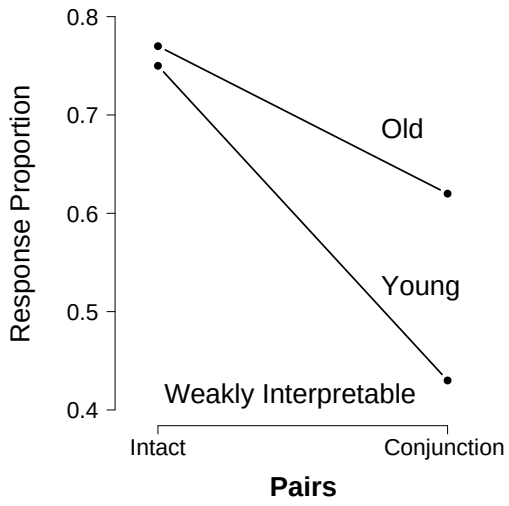

Figure 19. Age group × item type (experiment 2).

Comments on the Interaction of Figure 19.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"A significant Item Type × Age Group interaction was also present, F(1, 46) = 8.16, η_p^2 =.15. In particular, whereas older (M = .77) and younger (M = .75) adults endorsed intact pairs with relatively equal frequency (F < 1), older adults (M = .62) were far more likely to endorse conjunction pairs than were younger adults (M = .43), F(1, 46) =12.42, η_p^2 = .21." (Rhodes, Castel, & Jacoby, 2008, p. 243).

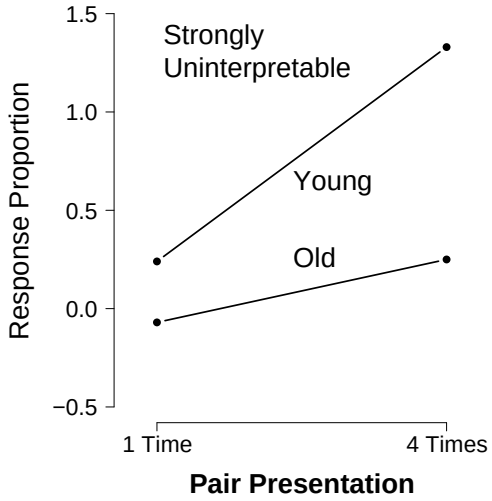

Figure 20. Age group × repetition (experiment 3).

Comments on the Interaction of Figure 20.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"In addition, a reliable Age Group × Repetition interaction was present. This reflects the fact that for younger adults, discriminability was reliably better for pairs composed of faces presented four times (M = 1.33) than for pairs composed of faces presented one time (M = 0.24), F(1, 29) = 27.18, η_p^2 = .48. Older adults likewise exhibited better discriminability for pairs composed of faces presented four times (M =0.25) than those presented one time (M = –0.07), F(1, 29) = 4.52, η_p^2 = .14, but the effect of repetition was of a smaller magnitude than that apparent for younger adults." (Rhodes, Castel, & Jacoby, 2008, p. 245).

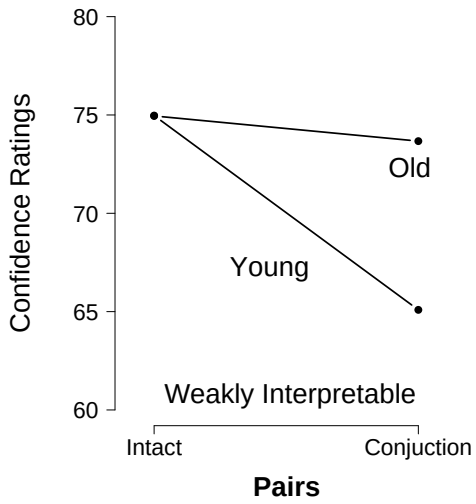

Figure 21. Age group × item type (experiment 3).

Comments on the Interaction of Figure 21.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"A reliable Age Group × Item Type interaction was also present, F(1, 49) = 8.59,η_p^2 = .15. Specifically, younger adults exhibited significantly higher confidence ratings for intact pairs (M = 74.97) than for conjunction (M = 65.09) pairs, F(1, 24) = 17.25, η_p^2 = .42. However, older adults confidence ratings did not distinguish between intact (M = 74.95) and conjunction (M = 73.67) pairs (F < 1)." (Rhodes, Castel, & Jacoby, 2008, p. 246).

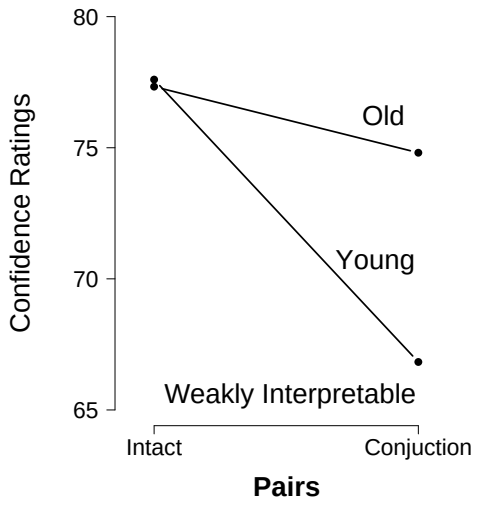

Figure 22. Age group × item type (footnote).

Comments on the Interaction of Figure 22.

Authors' Description of the Statistics.

"...but a reliable Age Group × Item Type interaction was present, F(1, 58) = 8.50, η_p^2 = .13. Follow–up analyses indicated that younger adults exhibited reliably higher confidence responses to intact (M = 77.60) pairs than those to conjunction (M = 66.83) pairs, F(1, 29) = 19.20, η_p^2 = .40. In contrast, older adults confidence in intact (M = 77.33) pairs was only marginally different from their confidence in conjunction (M = 74.81) pairs, F(1, 29) = 3.25, p = .08, η_p^2 = .10." (Rhodes, Castel, & Jacoby, 2008, p. 246).

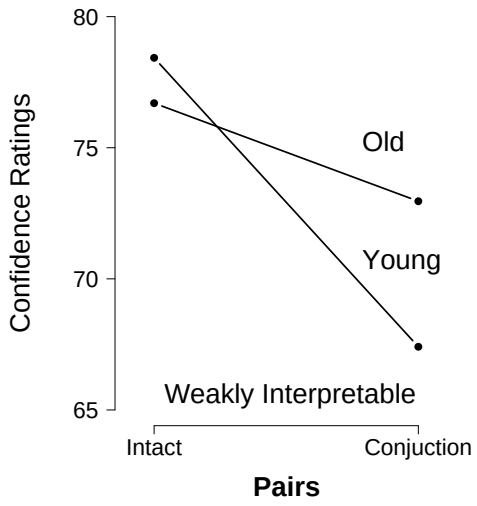

Figure 23. Face age × item type (footnote).

Comments on the Interaction of Figure 23.

The values were extracted from the text. This interaction was mentioned in a footnote.

Authors' Description of the Statistics.

> "... a reliable Face Age × Item Type interaction was present, F(1, 55) = 9.02, η_p^2 = .14. In particular, for intact pairs, confidence ratings did not differ for older (M = 76.70) faces and younger (M = 78.43) faces, F(1, 57) = 1.49, p = .23, η_p^2 = .03. However, confidence in intact responses to conjunction pairs was significantly higher for older (M = 72.96) faces than for younger (M = 67.41) faces, F(1, 56) = 7.26, η_p^2 = .12." (Rhodes, Castel, & Jacoby, 2008, p. 246).

Linda A. Henkel (p. 250–262)

Comments on the Interaction of Figure 24.

The values were extracted from Figure 1 (Henkel, 2008, p. 254). In order to plot the interaction we computed the means for each 'similarity to be seen item'.

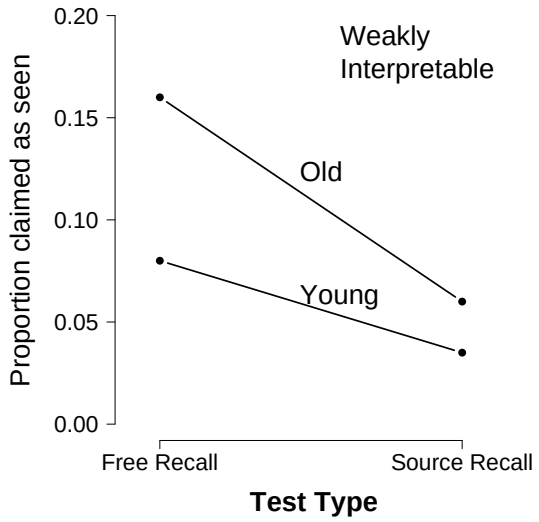

Figure 24. Test type × age interaction.

Author's Description of the Statistics.

"A significant Test Type × Age interaction was also found, $F(1, 82) = 4.70$, $MSE = 0.01$, $p < .05$. Both young and older adults benefited from the repeated source recall tests in that source misattributions were significantly lower for both age groups after the source recall tests than after the free recall tests. However, whereas older adults made significantly more errors following repeated free recall tests than did young adults, no significant age difference was found after repeated source recall tests." (Henkel, 2008, p. 253–254).

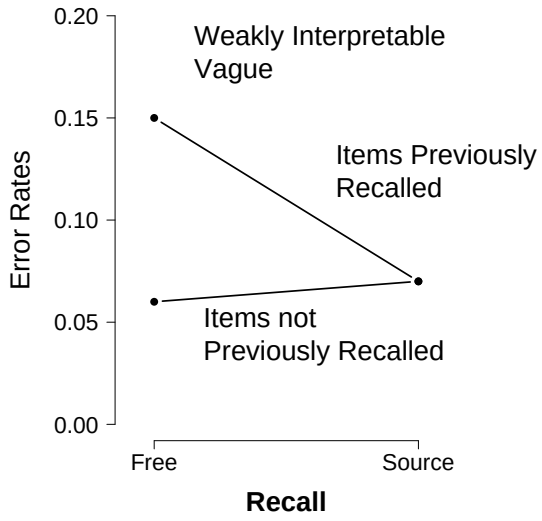

Figure 25. Prior recall status × specificity of earlier recall test.

Comments on the Interaction of Figure 25.

The values were extracted from the text.[1]

Author's Description of the Statistics.

"The Prior Recall Status (previously recalled, not previously recalled) × Specificity of Earlier Recall Tests interaction was significant, $F(1, 82) = 9.69$,MSE = 0.01, $p < .01$ as was the Prior Recall Status × Similarity interaction, $F(1, 82) = 25.01$, MSE = 0.01, $p < .001$. For items that had been previously recalled, more errors were made following free recall tests (.15) than following source recall tests (.07), and more errors were made for physically similar (.16) than for control items (.07). For items that had not been previously recalled, the specificity of the prior test did not influence error rates (free recall = .06, source recall = .05), and the relative difference in errors for physically similar (.07) and control items (.04) was less marked." (Henkel, 2008, p. 255).

[1]The values mentioned in the author's text concern the interaction of Figure 26 as well.

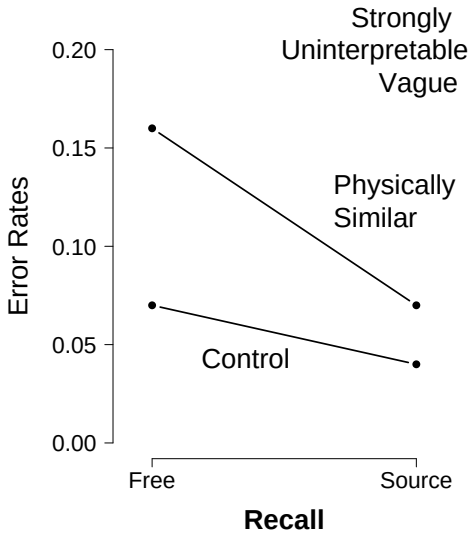

Figure 26. Prior recall status × similarity.

Comments on the Interaction of Figure 26.

Data were extracted from the text (please see below).

Author's Description of the Statistics.

"...as was the Prior Recall Status × Similarity interaction, $F(1, 82) = 25.01$, $MSE = 0.01$, $p < .001$. For items that had been previously recalled, more errors were made following free recall tests (.15) than following source recall tests (.07), and more errors were made for physically similar (.16) than for control items (.07). For items that had not been previously recalled, the specificity of the prior test did not influence error rates (free recall = .06, source recall = .05), and the relative difference in errors for physically similar (.07) and control items (.04) was less marked." (Henkel, 2008, p. 255).

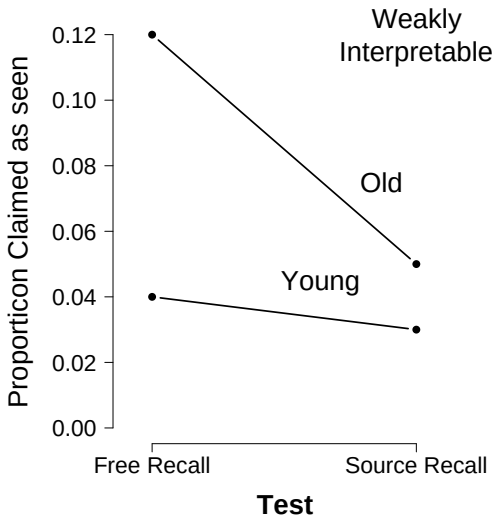

Figure 27. Test type × age group–experiment 2.

Comments on the Interaction of Figure 27.

The values were extracted from Figure 3 (Henkel, 2008, p. 257). We computed the means of the three item types in order to plot the interaction.

Author's Description of the Statistics.

> "Although there was a significant main effect for test type, this was qualified by a significant interaction with age group, $F(1, 110) = 9.32$, $MSE = 0.01$, $p < .01$. As in Experiment 1, older adults made significantly more source errors than young adults following the repeated free recall tests, and there was no significant difference in errors between the two age groups following the repeated source recall tests. " (Henkel, 2008, p. 256).

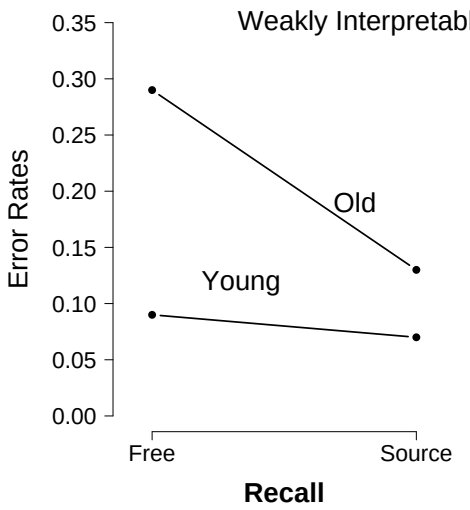

Figure 28. Age × test type.

Comments on the Interaction of Figure 28.

The values were extracted from Figure 4. We computed the values of the three 'similarity variables'.

Author's Description of the Statistics.

> "...follow-up comparisons on the significant Age × Test Type interaction, F(1, 110) = 8.75, MSE = 0.04, p < .01, show this was the case after repeated free recall tests but not after repeated source recall tests." (Henkel, 2008, p. 258).

Olivier Piguet, Emily Connally, Anne C. Krendl, Jessica R. Huot, Suzanne Corkin (p. 307–314)

Note: The present paper includes plenty 2×3 interactions which were not plotted because they did not meet our inclusion criteria. On Figure 29 different colours were used because the two lines were overplotted by each other.

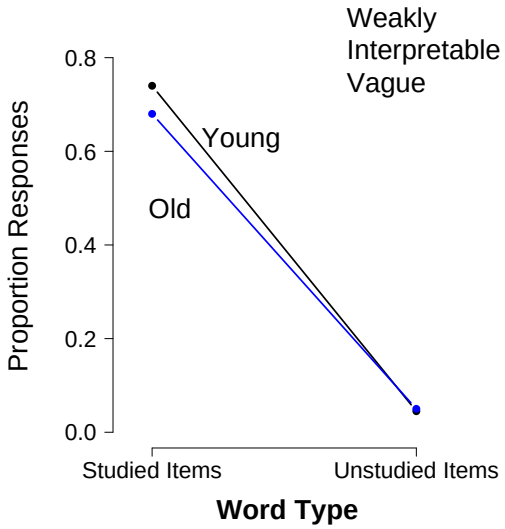

Figure 29. Word type × group (remembered words, list).

Comments on the Interaction of Figure 29.

The values were extracted from Figure 1 (Piguet et al., 2008, p. 310).

Authors' Description of the Statistics.

"A significant Word Type × Group interaction was also present, F(1, 70) = 11.28, p = .001, η^2 = .14. In other words, young adults had higher recognition scores than older adults on all studied words (all ps < .05)." (Piguet et al., 2008, p. 309).

Mei–Ching Lien, Eric Ruthruff, David Kuhns (p. 330–341)

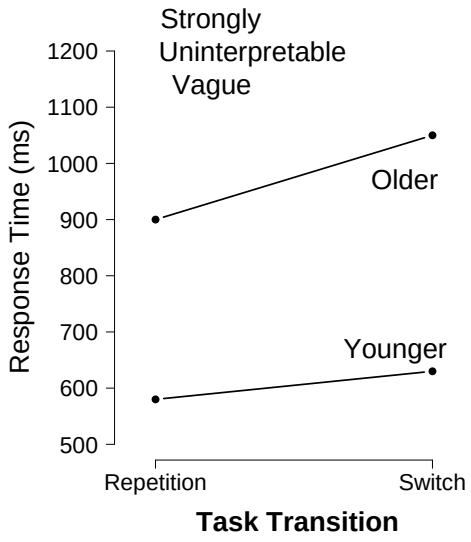

Figure 30. Expectacy cost × group–experiment 1.

Comments on the Interaction of Figure 30.

Authors' Description of the Statistics. The values were extracted from figure 2 (Lien, Ruthruff, & Kuhns, 2008, p. 334).

> "This interaction was stronger for older adults than for younger adults, $F(1, 74) = 17.72$, $p < .0001$, $MSE = 7.036$." (Lien, Ruthruff, & Kuhns, 2008, p. 334).

Raoul Bell, Axel Buchner, Iris Mund (p. 377–391)

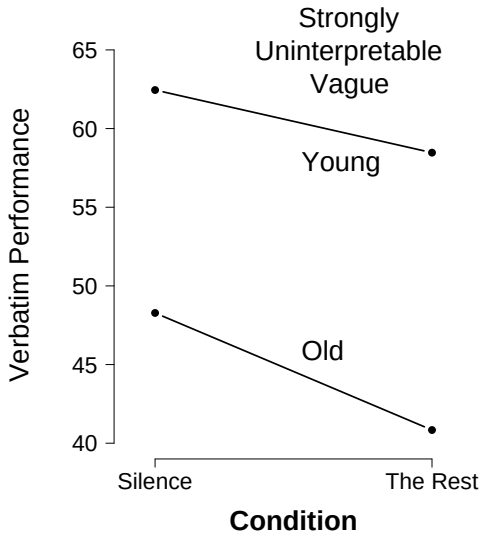

Figure 31. Age × condition (silence versus the rest)–Verbatim Performance

Comments on the Interaction of Figure 31.

"The values were extracted from Table 1 (Bell, Buchner, & Mund, 2008, p. 380). In order to plot the interaction we computed the mean of all conditions except the silent one."

Authors' Description of the Statistics.

> "...but the interaction between age and the variable contrasting the silence condition with all other conditions was significant, $F(1, 101) = 6.56$, $p < .01$, $\eta^2 = .06$." (Bell, Buchner, & Mund, 2008, p. 381).

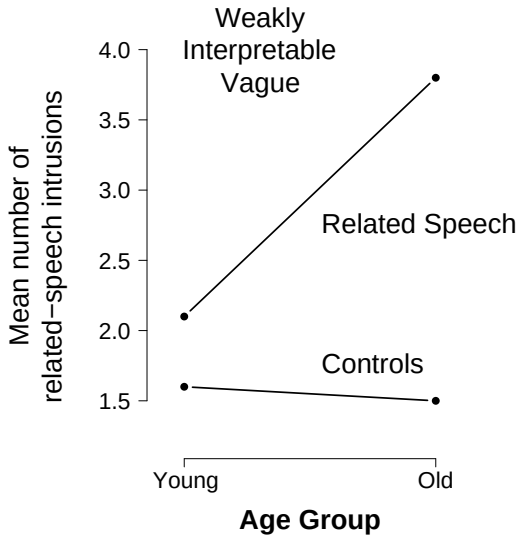

Figure 32. Age × related-speech conditions.

Comments on the Interaction of Figure 32.

The values were extracted from Figure 1 (Bell, Buchner, & Mund, 2008, p. 381). In order to plot the interaction, we computed the means of the control conditions, as those were shown on Figure 1 (Bell, Buchner, & Mund, 2008, p. 381).

Authors' Description of the Statistics.

> "The interaction between age and the variable contrasting the related–speech condition with the control conditions was significant, $F(1, 101) = 16.50$, $p < .01$, $\eta^2 = .14$, which is due to the large age difference in the number of related-speech intrusions in the related-speech condition, $t(101) = 4.22$, $p < .01$, $\eta^2 = .15$." (Bell, Buchner, & Mund, 2008, p. 381).

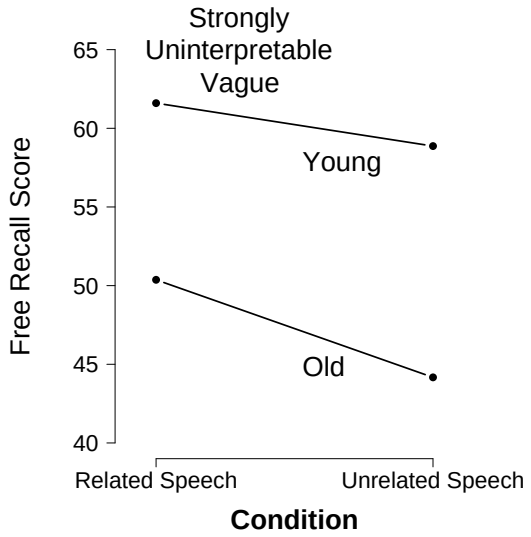

Figure 33. Age × condition – experiment 2.

Comments on the Interaction of Figure 33.

The values were extracted from Table 1 (Bell, Buchner, & Mund, 2008, p. 380).
Authors' Description of the Statistics.

"The interaction between age and the variable contrasting the unrelated condition with the related speech condition was significant, $F(1, 97) = 4.30$, $p = .04$, $\eta^2 = .04$."(Bell, Buchner, & Mund, 2008, p. 382).

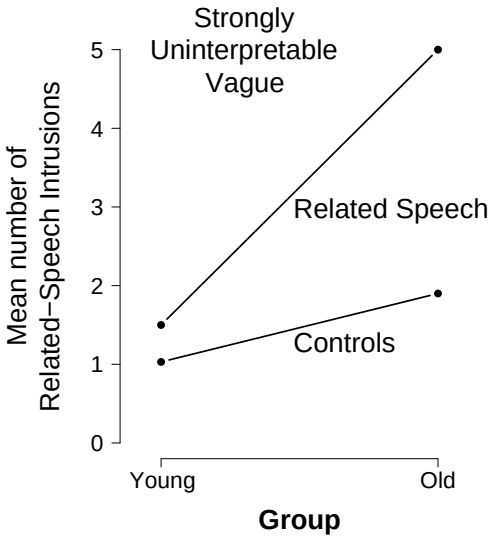

Figure 34. Age × condition (related versus controls).

Comments on the Interaction of Figure 34.

The values were extracted from Figure 2 (Bell, Buchner, & Mund, 2008, p. 383). In order to plot the interaction we computed the values across all control conditions.

Authors' Description of the Statistics.

> "The interaction between age and the variable contrasting the related-speech condition with the control conditions was significant, $F(1, 97) = 13.43$, $p < .01$, $\eta^2 = .12$." (Bell, Buchner, & Mund, 2008, p. 382).

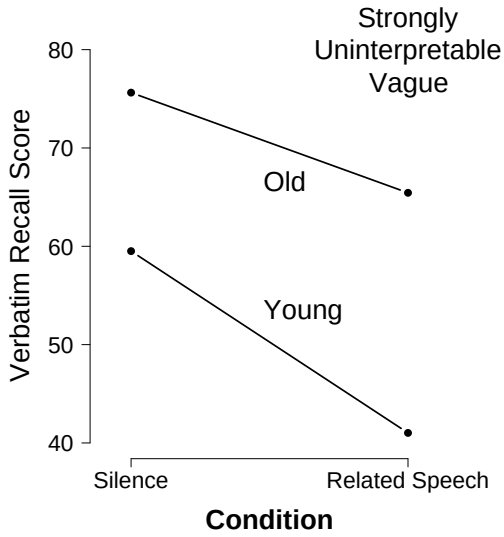

Figure 35. Age × distractor condition.

Comments on the Interaction of Figure 35.

The values were extracted from Table 1 (Bell, Buchner, & Mund, 2008, p. 380).
Authors' Description of the Statistics.

"A repeated measures MANOVA revealed main effects of age, $F(1, 89) = 34.11$, $p < .01$, $\eta^2 = .28$, and of distractor condition, $F(1, 89) = 159.09$, $p < .01$, $\eta^2 = .64$, and an interaction between both variables, $F(1, 89) = 13.30$, $p < .01$, $\eta^2 = .13$." (Bell, Buchner, & Mund, 2008, p. 385).

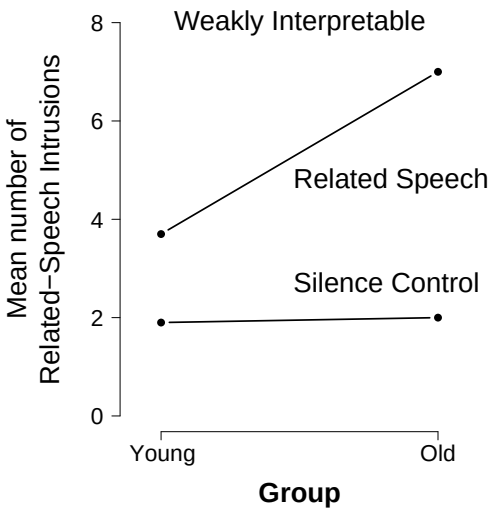

Figure 36. Age × distractor type.

Comments on the Interaction of Figure 36.

The values were extracted from Figure 3 (Bell, Buchner, & Mund, 2008, p. 385).
Authors' Description of the Statistics.

"The effect of distractor condition was significant, $F(1, 89) = 67.54$, $p < .01$, $\eta^2 = .43$, as was the interaction between age and distractor type, $F(1, 89) = 16.32$, $p < .01$, $\eta^2 = .16$. This interaction can be attributed to the fact that older adults made more related-speech intrusions than younger adults in the related-speech condition, $t(89) = 4.78$, $p < .01$, 2 $\eta^2 = .20$, but not in the control condition, $t(89) = 0.48$, $p = .64$, $\eta^2 = .01$." (Bell, Buchner, & Mund, 2008, p. 385).

Michael J. Frank, Lauren Kong (p. 392–398)

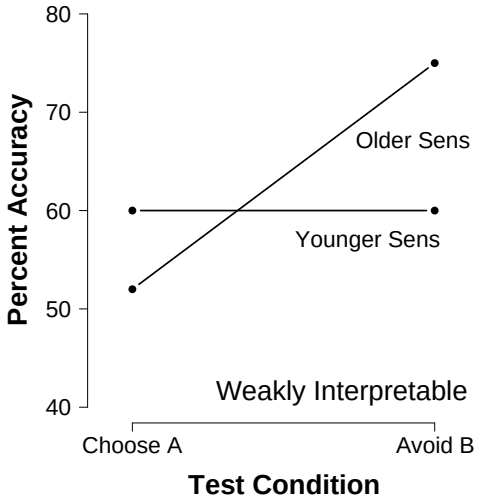

Figure 37. Novel test–pair performance in younger and older sens (seniors).

Comments on the Interaction of Figure 37.

Data were extracted from Figure 1B (Frank , & Kong, 2008, p. 394).
Authors' Description of the Statistics.

"Nevertheless, the interaction between age group and test–pair condition (choose–A and avoid–B) was significant, $F(1, 41) = 4.7$, $p = .036$. O–O seniors were better at negative than positive learning, $F(1, 21) = 16.1$, $p = .0006$, whereas Y–O seniors did not differ in these measures." (Frank , & Kong, 2008, p. 395).

Susan R. Old, Moshe Naveh–Benjamin (p. 467–472)

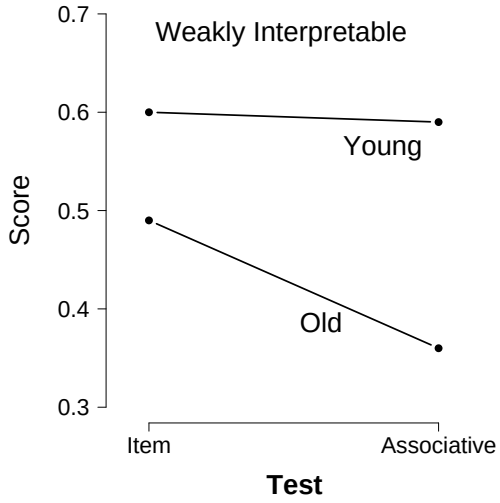

Figure 38. Age × test interaction.

Comments on the Interaction of Figure 38.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"It is important, however, that there was a significant 2 (age) × 2 (test: item, associative) interaction, $F(1, 54) = 4.95$, $MSE = .037$, $p < .05$. Whereas the younger adults item and associative scores did not differ, $t(27) = 0.60$, $p > .50$ ($M = 0.60$ and 0.59, $SD = 0.12$ and 0.16, for item and associative, respectively), older adults scored significantly higher on the item than on the associative measure, $t(27) = 3.11$, $p < .01$ ($M = 0.49$ and 0.36, $SD = 0.15$ and 0.18)." (Old, & Naveh–Benjamin, 2008 ,p. 469).

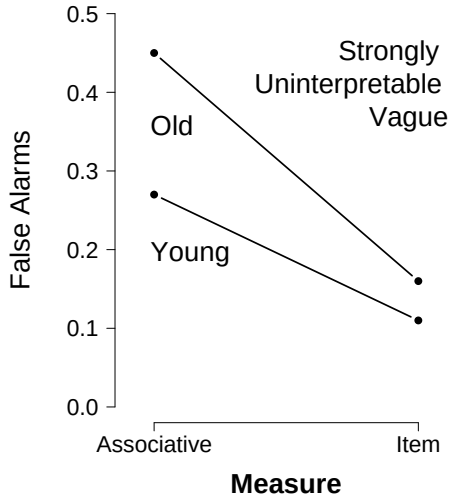

Figure 39. Age × test–false alarms.

Comments on the Interaction of Figure 39.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"Interestingly, whereas the three–way interaction was again not significant, this analysis yielded a significant 2 (age) × 2 (test) interaction, $F(1, 54) = 17.31$, $MSE = .013$, $p < .001$. Whereas older adults produced more false alarms than younger adults on both test measures, age differences were larger on the associative measure, $t(54) = 6.07$, $p < .001$ ($M = 0.27$ and 0.45, $SD = 0.09$ and 0.13 for younger and older adults, respectively), than on the item measure, $t(54) = 2.20$, $p < .05$ ($M = 0.11$ and 0.16, $SD = 0.07$ and 0.10 for younger and older adults, respectively)." (Old, & Naveh–Benjamin, 2008 ,p. 471).

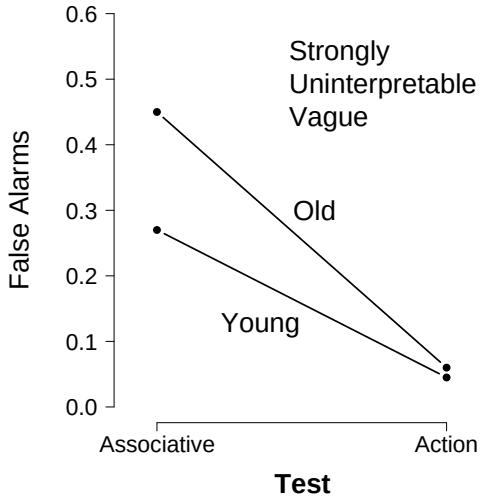

Figure 40. Age × test interaction (associative versus action test).

Comments on the Interaction of Figure 40.

The values were extracted from Table 1 (Old, & Naveh–Benjamin, 2008, p. 469). This interaction was classified as 'Strongly Uninterpretable', although it looks like 'Weakly Interpretable;. We based our desicion on the fact that authors mention that Age differences in terms of false alarms were larger on the associative than the action test (Old, & Naveh–Benjamin, 2008, p. 471).

Authors' Description of the Statistics.

> "Two three–way ANOVAs, each comparing one of the item test measures with the associative measure, revealed nonsignificant three–way interactions (both ps $> .4$), but both revealed significant Age × Test interactions. Age differences in terms of false alarms were larger on the associative than the action test, $F(1, 54) = 21.02$, $MSE = .017$, $p < .001$." (Old, & Naveh–Benjamin, 2008 ,p. 471).

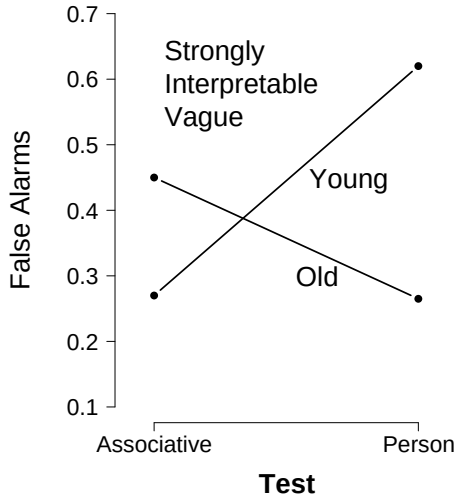

Figure 41. Age × test interaction (asociative versus person test).

Comments on the Interaction of Figure 41.

The values were extracted from Table 1 (Old, & Naveh–Benjamin, 2008, p. 469).

Authors' Description of the Statistics.

> "Two three–way ANOVAs, each comparing one of the item test measures with the associative measure, revealed nonsignificant three–way interactions (both ps $< .4$), but both revealed significant Age × Test interactions. Age differences in terms of false alarms were larger on the associative than the action test, $F(1, 54) = 21.02$, $MSE = .017$, $p < .001$, and on the associative than the person test, $F(1, 54) = 5.20$, $MSE = .025$, $p < .05$." (Old, & Naveh–Benjamin, 2008 ,p. 471).

Anjali Thapar, Sarah Malaya Sniezek (p. 473–477)

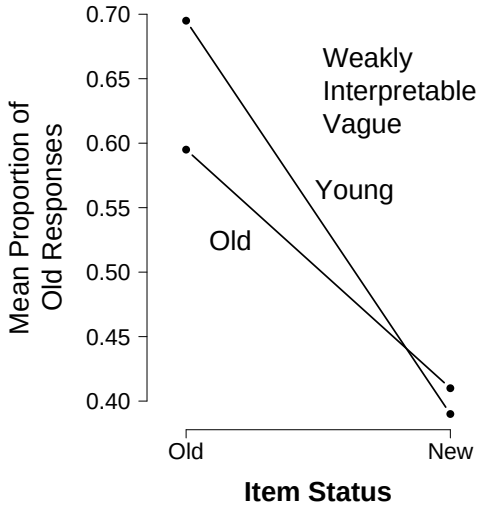

Figure 42. Age × item status–experiment 1.

Comments on the Interaction of Figure 42.

The values were extracted from Table 1 (Thapar, & Sniezek, 2008, p. 475). In order to plot the interaction we took the mean of the two test formats (Revealed, Intact) for each level of item status (Old, New).

It is noted that the values mentioned in the text explain the interaction but were not used in order to plot it.

Authors' Description of the Statistics.

> "However, there was a significant Age × Item Status interaction, F(1, 44) = 9.48, p = .004, η_p^2 = .18, and a significant Item Status × Test Format interaction, F(1, 44) = 9.39, p = .004, η_p^2 = .18. The Age × Item Status interaction reflected the fact that the younger adults had a higher proportion of hit rates compared to the older adults (younger: M = 0.70, SD = 0.14; older: M = 0.60, SD = 0.19), F(1, 44) = 4.96, p = .03, η_p^2 = .10, but the two groups had similar proportions of false alarms (younger: M = 0.39, SD = 0.12; older: M = 0.41, SD = 0.20), F(1, 44) = 0.20, p = .66, η_p^2 = .004." (Thapar, & Sniezek, 2008, p. 475).

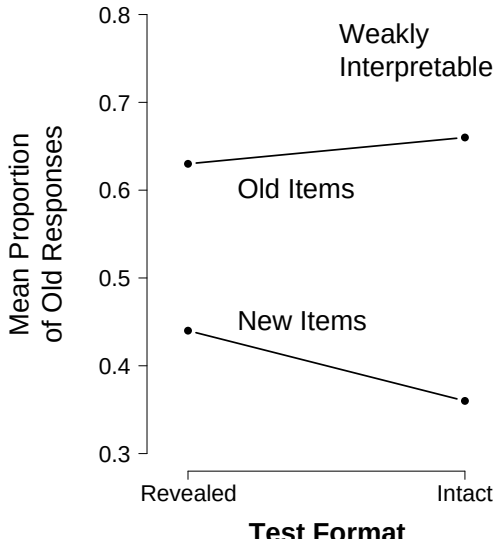

Figure 43. Item × test format–experiment 1.

Comments on the Interaction of Figure 43.

The values were extracted from the text although they are similar to the values of Table 1 (Thapar, & Sniezek, 2008, p. 475) if someone computes the mean between the two age groups (young and old).

Authors' Description of the Statistics.

> " However, there was a significant Age × Item Status interaction, F(1, 44) = 9.48, p = .004, η_p^2 = .18, and a significant Item Status × Test Format interaction, F(1, 44) = 9.39, p = .004, η_p^2 = .18. The Item Status × Test Format interaction reflected that . . . a revelation effect was observed in new items (revealed: M = 0.44, SD = 0.17; intact: M = 0.36, SD = 0.16), F(1, 45) = 9.78, p =.003, η_p^2 = .18, but was not observed in old items (revealed: M = 0.63, SD = 0.18; intact: M = 0.66, SD = 0.17), F(1, 45) = 1.42, p = .24, η_p^2 = .03." (Thapar, & Sniezek, 2008, p. 475).

Comments on the Interaction of Figure 44.

In similar fashion as for the interaction of figure 42, the values were extracted from Table 1 (Thapar, & Sniezek, 2008, p. 475) in order to plot the interaction we

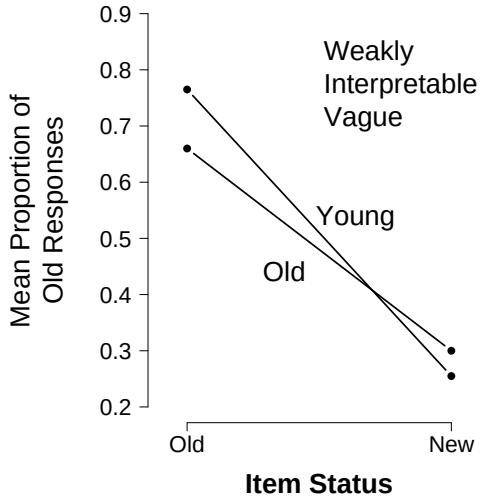

Figure 44. Age × item status–experiment 2.

took the mean of the two test formats (Revealed, Intact) for each level of item status (Old, New).

Authors' Description of the Statistics.

"The Age × Item Status interaction reflected the fact that younger adults had a higher proportion of hit rates compared to the older adults (younger: M = 0.76, SD = 0.14; older: M = 0.66, SD = 0.20), F(1, 66) = 6.68, p < .01, $\eta_p^2 = .09$, but the two groups had a similar proportion of false alarms (younger: M = 0.26, SD = 0.20; older: M = 0.30, SD = 0.20), F(1, 66) = 1.01, p = .32, $\eta_p^2 = .02$." (Thapar, & Sniezek, 2008, p. 476).

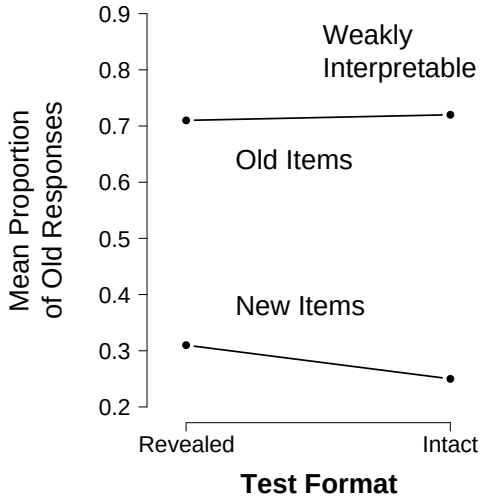

Figure 45. Age × test format–experiment 2.

Comments on the Interaction of Figure 45.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"The Item Status × Test Format interaction reflected the fact that a revelation effect was observed in new items (revealed: M = 0.31, SD = 0.22; intact: M = 0.25, SD = 0.17), F(1, 67) = 10.50, p = .002, η_p^2 = .14, but not old items (revealed: M = 0.71, SD = 0.18; intact: M = 0.72, SD = 0.18), F(1, 67) = 0.30, p = .59, η_p^2 = .004." (Thapar, & Sniezek, 2008, p. 476).

Issue 3

Lili Sahakyan, Peter F. Delaney, Leilani B. Goodmon
(p. 621–633)

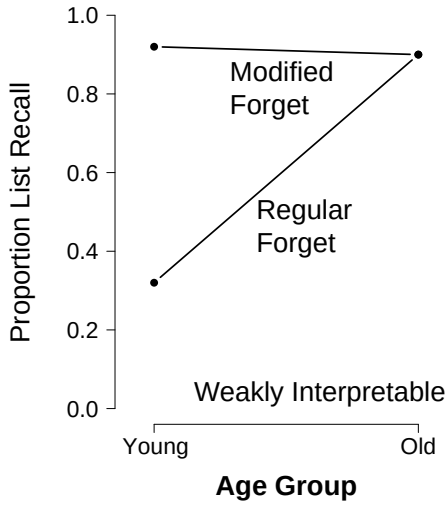

Figure 46. Cue × group–experiment 2–(directed forgetting costs).

Comments on the Interaction of Figure 46.

The values were extracted from Figure 3 (Sahakyan, Delaney, & Goodmon, 2008, p. 627).

Authors' Description of the Statistics.

> "However, these effects were moderated by a significant two–way interaction, $F(1, 89) = 24.51$, $MSE = .080$, $p < .001$, $\eta^2 = .216$, indicating that younger adults were equally likely to employ forgetting strategies with the standard forget cue and the modified forget cue ($t < 1$), whereas older adults employed forgetting strategies less often with the standard forget cue than with the modified forget cue, $t(27) = 4.03$, $p < .01$. In sum, for young participants, virtually everyone tried to do something to forget, but for older participants, apparently more prompting was needed in order to engage them in deliberate forgetting strategies." (Sahakyan, Delaney, & Goodmon, 2008, p. 627).

Mark A. McDaniel, Keith B. Lyle, Karin M. Butler,
Courtney C. Dornburg (p. 646–656)

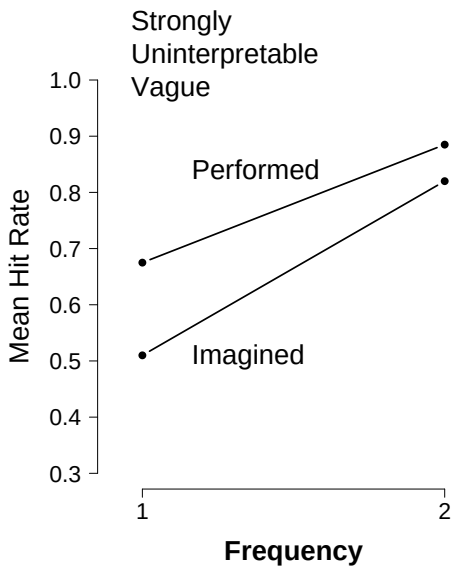

Figure 47. Frequency (1 versus 2) × source (imagined and performed only) interaction.

Comments on the Interaction of Figure 47.

The values were extracted from Table 2 (Emery, Hale, & Myerson, 2008, p. 650). In order to plot the interaction we computed the mean of both groups (Younger and Older) for each source (Imagined only and Performed only).

Authors' Description of the Statistics.

> "...and this resulted in a significant interaction of frequency and source, $F(1, 49) = 5.62$, $MSE = 0.02$, $p = .02$." (McDaniel, Lyle, Butler, & Dornburg 2008, p. 650).

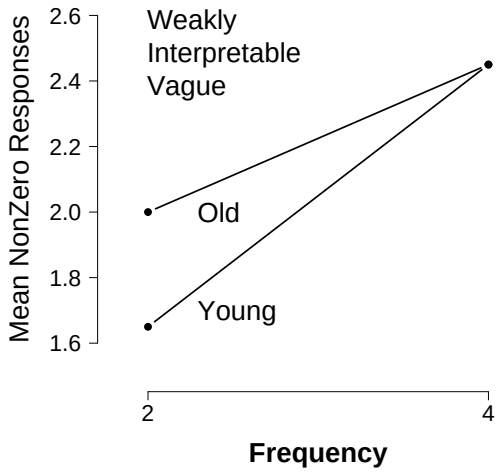

Figure 48. Frequency (2 versus 4) × age (young versus old) interaction.

Comments on the Interaction of Figure 48.

The values were extracted from Table 3 (Emery, Hale, & Myerson, 2008, p. 651). In order to plot the interaction, we computed the mean of both source (Imagination and Performance) for each group (Younger and Old).

Authors' Description of the Statistics.

> "...There was, however, a significant interaction between age and frequency, F(2, 80) = 4.54, MSE = 0.36, p = .01. As seen in the top half of Table 3, the increase in estimates between actions imagined two times versus four times was larger for younger adults (Mdifference = 1.20) than older adults (Mdifference = 0.50). To confirm that this difference drove the Age × Frequency interaction, we submitted estimates from the 2× and 4× conditions to a 2 (age) × 2 (frequency) ANOVA. The interaction remained significant, F(1,40) = 9.75, MSE = 0.29, p = .003. Despite this interaction, mean frequency estimates did not differ significantly as a function of age at any level of frequency (smallest p = .14), and overall, older and younger adults were generally similar in tracking the number of

prior acts of imagination." (McDaniel, Lyle, Butler, & Dornburg 2008, p. 651).

Lori E. James, Kethera A. Fogler, Sarah K. Tauber (p. 657–664)

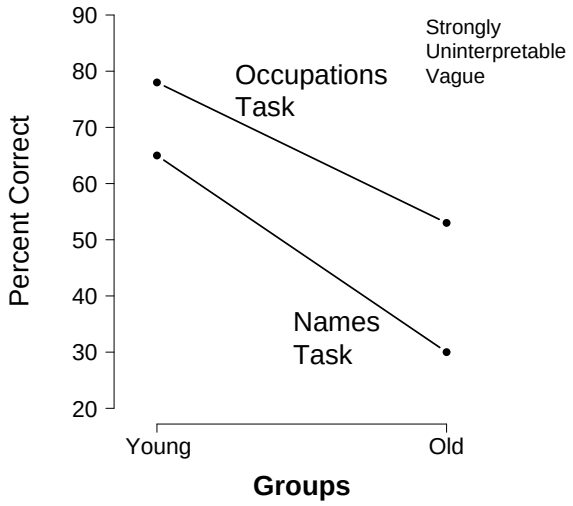

Figure 49. Age × information type (experiment 1).

Comments on the Interaction of Figure 49.

The values were extracted from Figure 1 (James, Fogler, & Tauber, 2008, p. 660).

Authors' Description of the Statistics.

> "Age interacted with information type, $F(1, 46) = 7.96$, partial $\eta^2 = .15$, $p < .01$, because older adults did disproportionately worse than young adults on facename matching compared to faceoccupation matching, although age differences were significant for both facename matching, $t(46) = 6.47$, $p < .01$, and faceoccupation matching, $t(46) = 4.58$, $p < .01$." (James, Fogler, & Tauber, 2008, p. 659).

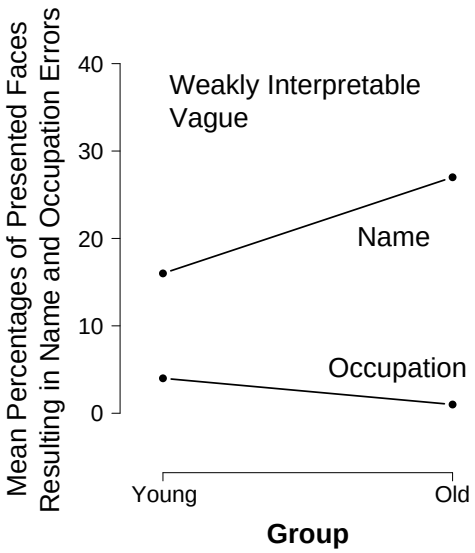

Figure 50. Age × error type (experiment 1).

Comments on the Interaction of Figure 50.

The values were extracted from Table 1 (Top section. James, Fogler, & Tauber, 2008, p. 660).

Authors' Description of the Statistics.

> "This ANOVA yielded the same pattern of effects, including the critical interaction of age and error type that demonstrates a disproportionate deficit in name matching for older adults, $F(1, 46) = 9.00$, partial $\eta^2 = .16$, $p < .01$." (James, Fogler, & Tauber, 2008, p. 660).

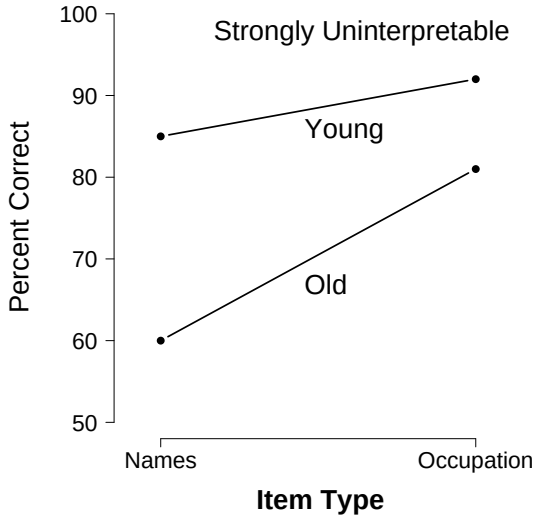

Figure 51. Age × information type (experiment 2).

Comments on the Interaction of Figure 51.

The values were extracted from Figure 1 (Middle Panel. James, Fogler, & Tauber, 2008, p. 660).

Authors' Description of the Statistics.

> " . . . and age interacted with information type, F(1, 46) = 24.15, partial η^2 = .34 , p < .01. Although young adults significantly outperformed older adults for both facename and faceoccupation recognition, older adults did disproportionately worse than young adults on the multiple–choice test for facename associations, t(46) = 6.43, p < .01, than for faceoccupation associations, t(46) = 3.89, p < .01." (James, Fogler, & Tauber, 2008, p. 661).

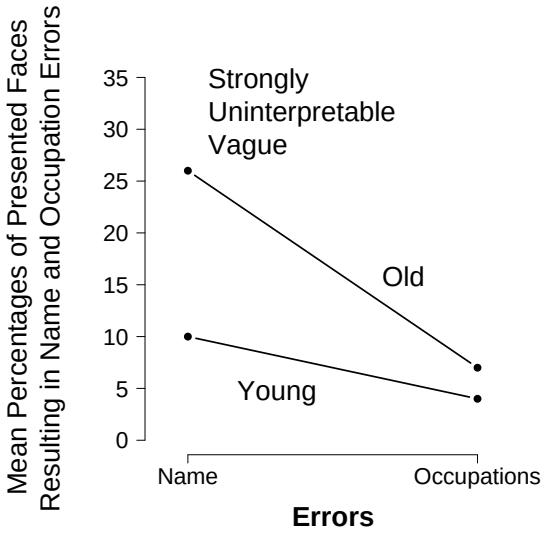

Figure 52. Age × error rate (experiment 2).

Comments on the Interaction of Figure 52.

The values were extracted from Table 1 (Middle section. James, Fogler, & Tauber, 2008, p. 660).

Authors' Description of the Statistics.

> "This ANOVA yielded the same pattern of effects, including the critical interaction of age and error type that demonstrates a disproportionate deficit in face-name association recognition for older adults, $F(1, 46) = 24.14$, partial $\eta^2 = .34$, $p < .01$ (see Table 1)." (James, Fogler, & Tauber, 2008, p. 661).

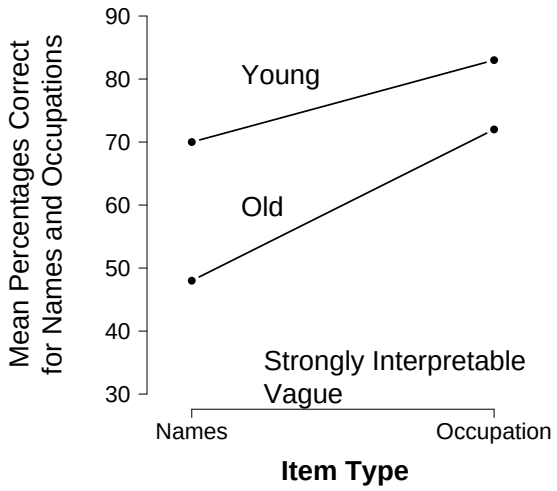

Figure 53. Age × item type interaction for the correct trials (experiment 2).

Comments on the Interaction of Figure 53.

The values were extracted from Table 2 (James, Fogler, & Tauber, 2008, p. 660).

Authors' Description of the Statistics.

> " Also as in Experiment 1, we analyzed the percentage correct data from only the first testing round. For these data, the critical interaction of age and item type was significant, $F(1, 46) = 4.25$, partial $\eta^2 = .09$, $p < .05$." (James, Fogler, & Tauber, 2008, p. 661).

Andrew E. Reed, Joseph A. Mikels, Kosali I. Simon (p. 671–675)

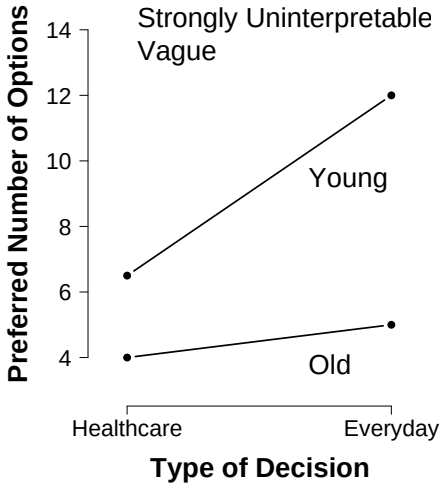

Figure 54. Age × domain interaction.

Comments on the Interaction of Figure 54.

The values were extracted from Figure 1 (Reed, Mikels, & Simon, 2008, p. 674).
Authors' Description of the Statistics.

" . . . an Age × Domain interaction such that the difference was even greater for young adults relative to older adults(graph's legend). In addition to the predicted main effect of age on choice preference, we also observed an interaction between age and domain type (everyday versus health care) such that young adults showed a greater difference in preferred choice between health care and everyday decisions, relative to older adults." (Reed, Mikels, & Simon, 2008, p. 673).

Issue 4

Erika Dahlin, Lars Nyberg, Lars Bóckman, Anna Stigsdotter Neely (p. 720–730)

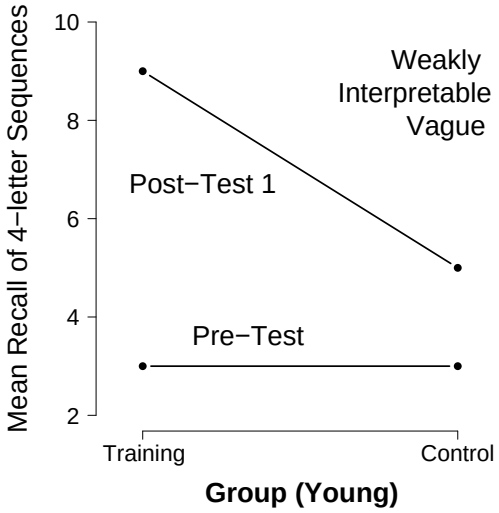

Figure 55. Group × session interaction (younger group).

Comments on the Interaction of Figure 55.

The values were extracted from Figure 1 (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 724). The interaction concerns only the young group.

Authors' Description of the Statistics.

> "More importantly, the Group × Session interaction was significant, $F(1, 24) = 34.51$, $MSE = 1.73$, $p < .001$ ..." (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 724).

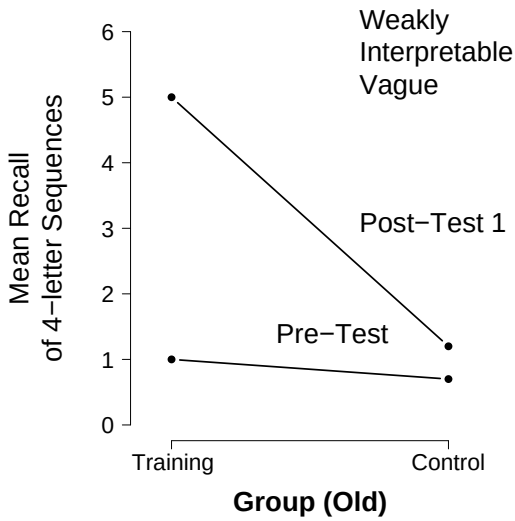

Figure 56. Group × session interaction (older group).

Comments on the Interaction of Figure 56.

The values were extracted from Figure 1 (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 724). The interaction concerns only the old group.

Authors' Description of the Statistics.

> "As with the young adult group, the Group × Session interaction was significant for the older participants, $F(1, 25) = 25.80$, $MSE = 1.71$, $p < .001$ . . . " (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 724).

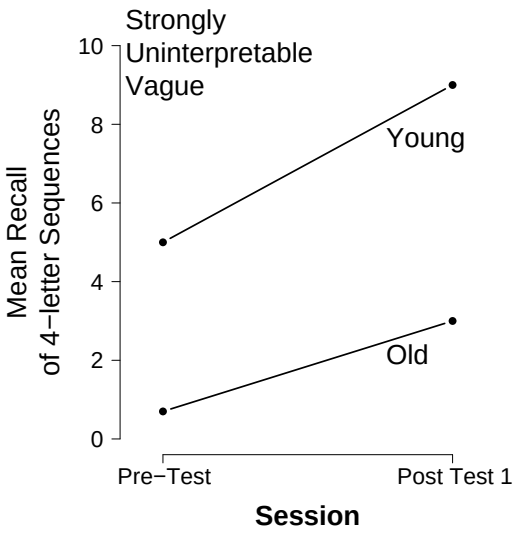

Figure 57. Group × session (differences during training).

Comments on the Interaction of Figure 57.

The values were extracted from Figure 1 (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 724). We considered only the bars referring to the training results.

Authors' Description of the Statistics.

> "More importantly, the Group × Session interaction was significant, F(1, 24) = 5.49, MSE = 11.27, p < .05, indicating that the improvement was more pronounced for the young trained than the old trained." (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 724).

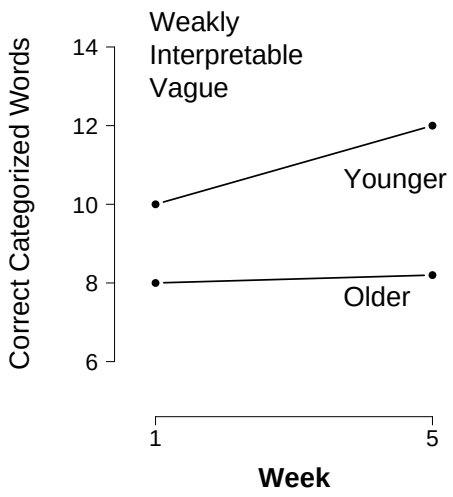

Figure 58. Group × session (keep track performance).

Comments on the Interaction of Figure 58.

Authors' Description of the Statistics.

"Finally, a reliable Group × Session interaction, $F(1, 26) = 6.47$, $MSE = 1.12$, $p < .05$, showed that the young participants, but not the older participants, improved keep–track performance across the 5 weeks." (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 725).

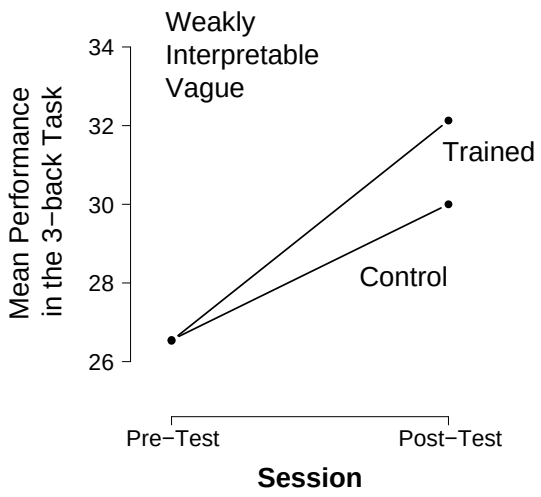

Figure 59. Group × session (3rd back task)–young people.

Comments on the Interaction of Figure 59.

The values were extracted from Table 2 (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 726).

Authors' Description of the Statistics.

> "A significant Group × Session interaction, $F(1, 24) = 5.93$, $MSE = 2.47$, $p < .05$, emerged for the young adults in the 3–back task, which indicated that the young trained improved more from pretest to Posttest 1 compared with the young controls." (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 725).

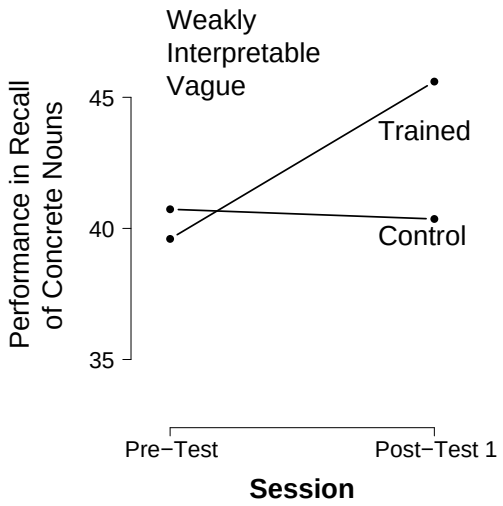

Figure 60. Group × session interaction for recall of concrete words. The interaction concerns only the younger group.

Comments on the Interaction of Figure 60.

The values were extracted from Table 2 (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 726).

Authors' Description of the Statistics.

> "For recall of concrete nouns, a significant Group × Session interaction was obtained for the young groups, $F(1, 24) = 9.48$, $MSE = 13.55$, $p < .005$, which indicated a more pronounced improvement between pretest to Posttest 1 for the young trained compared with young controls." (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 726).

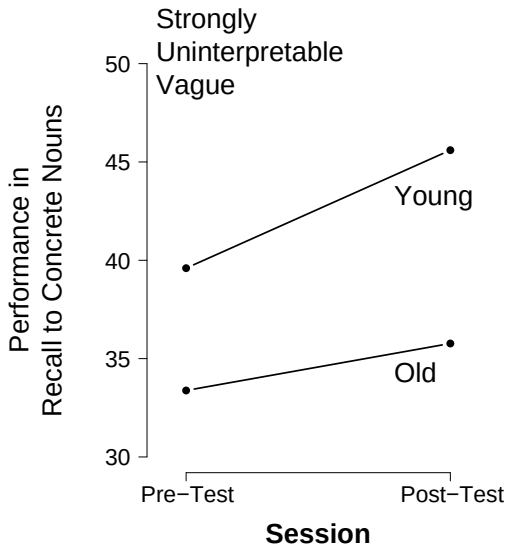

Figure 61. Group × session (episodic memory).

Comments on the Interaction of Figure 61.

The values were extracted from Table 2 (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 726).

Authors' Description of the Statistics.

> "To compare age differences in transfer effects for recall of concrete nouns, we performed a 2 (Group: young trained, old trained) × 2 (Session: pretest, Posttest 1) ANOVA, which revealed a significant Group × Session interaction, $F(1, 26) = 4.61$, $MSE = 9.87$, $p < .05$, showing that young trained improved more from pretest to Posttest 1 compared with old trained." (Dahlin, Nyberg, Bóckman, & Neely, 2008, p. 726).

Cindy Lustig, Kristin E. Flegal (p. 754–764)

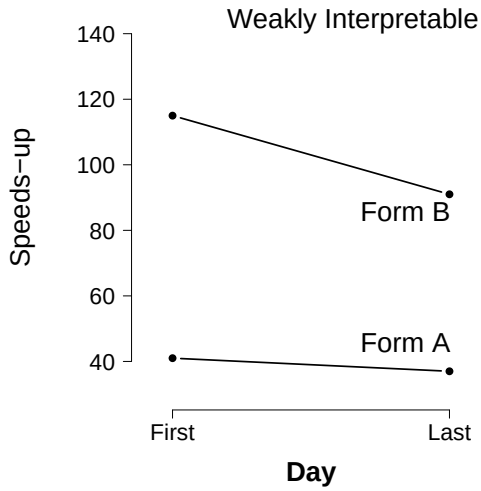

Figure 62. Day (first, last) × form (A, B).

Comments on the Interaction of Figure 62.

The values were extracted from the text (please see below).
Authors' Description of the Statistics.

" . . . The Training Day (first, last) × Form (A, B) interaction was significant, $F(1, 30) = 5.02$, $p < .05$ . . . " (Lustig, & Flegal, 2008, p. 759-760). ". . . Post hoc t–tests showed that speed–ups on Trail–Making, Version A were only marginal, from 41s to 37s, $t(31) = 1.61$, $p < .10$, corresponding with the lack of change on Pattern Comparison Test, another test of perceptual speed (9.9 items completed on both days, $t < 1$). In contrast, performance on Trail–Making, Version B, improved from pre– to posttest-115 s vs. 91 s, $t(31) = 2.53$, $p < .05$." (Lustig, & Flegal, 2008, p. 760).

Charlotte Mickler, Ursula M. Staudinger (p. 787–799)

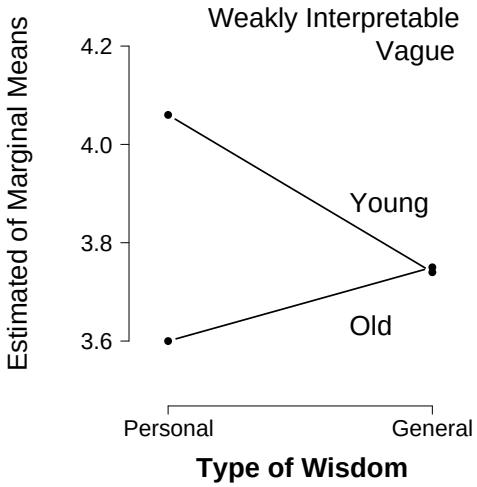

Figure 63. Age × wisdom.

Comments on the Interaction of Figure 63.

The values were extracted from Figure 5 (Mickler, & Staudinger, 2008, p. 796).
Authors' Description of the Statistics.

"The interaction of wisdom type with age was significant, $F(1, 149) = 6.31$, $p < .05$ (see Figure 5 for illustration), indicating that age plays a stronger role in PW (*Personal Wisdom*) than in GW (*General Wisdom*)." (Mickler, & Staudinger, 2008, p. 794).

Gillian Slessor, Louise H. Phillips, Rebecca Bull (p. 812–822)

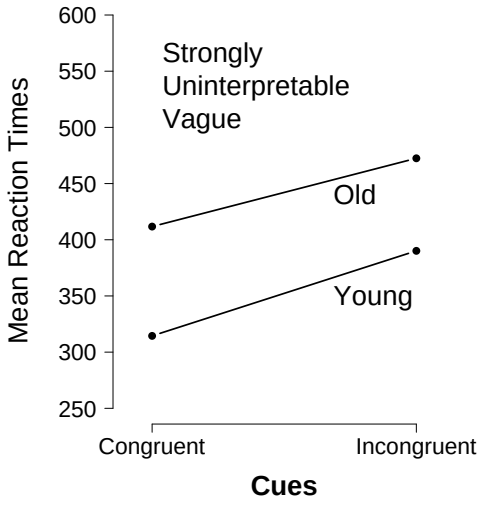

Figure 64. Age × cue congruity (arrow cueing task).

Comments on the Interaction of Figure 64.

The values were extracted from Table 3 (Slessor, Phillips, & Bull, 2008, p. 816).
Authors' Description of the Statistics.

"As for the emotion gaze cueing task, there was also an Age × Cue Congruity interaction, $F(1, 79) = 20.35$, $p < .001$, $\eta^2 = .21$. Paired samples t tests comparing RT in the congruent trials with RT on the incongruent trials showed that both younger, $t(44) = 10.37$, $p < .001$, $d = 1.55$, and older adults, $t(35) = 8.84$, $p < .001$, $d = 1.47$, responded significantly more quickly to congruent (vs. incongruent) trials." (Slessor, Phillips, & Bull, 2008, p. 819).

Linda K. Langley, Paul D. Rokke, Atiana C. Stark, Alyson L. Saville, Jaryn L. Allen, Angela G. Bagne (p. 873–855)

Note: In order to plot most of the interactions, we had to compute the values from the tables that were provided in the paper. However, for some interactions there were not enough information to compute the values.

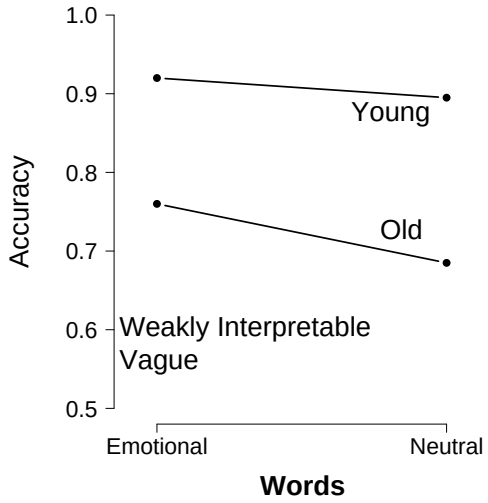

Figure 65. Age × valence (experiment 1a).

Comments on the Interaction of Figure 65.

The values were extracted from Table 3 (Langley et al., 2008, p. 878). In order to plot the interaction, we used the means across tasks for each stimulus (Emotional and Neutral) for each task (Single and Dual).

Authors' Description of the Statistics.

> "In addition to the main effect of valence, there was an Age × Valence interaction, $F(1, 58) = 4.55$, $p < .05$. Emotion effects, reported in Table 3, were calculated by subtracting neutral T2 accuracy from positive T2 accuracy. Both age groups identified positive words more accurately than neutral words, $ts(30) = 3.90$, $ps < .001$. However, emotion effects were

significantly greater for older adults (.07) than for young adults (.03), $F(1, 58) = 4.70$, $p < .05$. Note that the absence of other interactions involving valence suggests that the age difference in emotion effects generalized across tasks and lags and was not specific to the attentional blink." (Langley et al., 2008, p. 878).

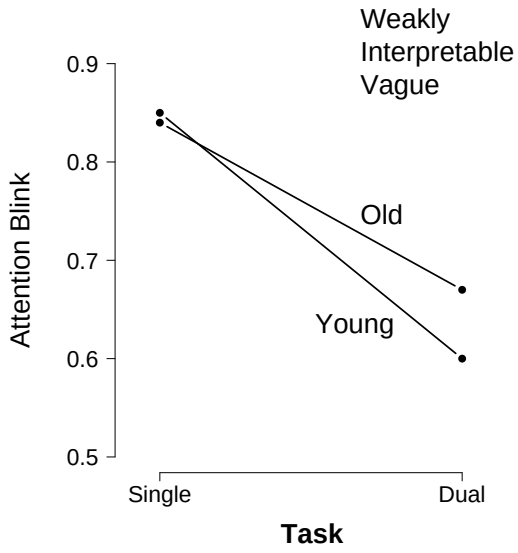

Figure 66. Age × task (experiment 1b)–attention blink.

Comments on the Interaction of Figure 66.

The values were extracted from the text (please see below).

Authors' Description of the Statistics.

"**Attentional blink.** There were two–way interactions of Age × Task, $F(1, 58) = 6.55$, $p < .05$, and Task × Lag, $F(7, 406) = 24.67$, $p < .0001$, reflecting attention effects. With similar accuracy on the single task (.84 for young adults, .85 for older adults), the Age × Task interaction was explained by greater task effects for older adults (a .25 decline in accuracy from the single task to the dual task) than for young adults (a .17 decline)." (Langley et al., 2008, p. 879).

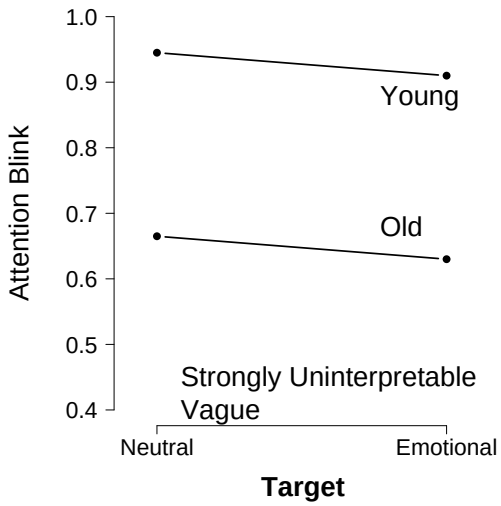

Figure 67. Age × valence (experiment 2a).

Comments on the Interaction of Figure 67.

To plot the described interaction, we used the mean of the Emotional and Neutral scores for both tasks (Dual & Single). Our choice was driven by the fact that there were not explicit scores on the task. Also, the emotionality rating (see Table 4) was not chosen, as this is the value that was computed posthoc by the authors to estimate emotionality.

Authors' Description of the Statistics.

> " **Emotion effects.** Although there was no main effect of valence, there was an Age × Valence interaction, $F(1, 58) = 7.69$, $p < .01$. As shown in Table 4, difference scores reflecting emotion effects (negative T2 accuracy × neutral T2 accuracy) were significantly greater for older adults (.03) than for young adults (–.01), $F(1, 58) = 7.56$, $p < .01$. Older adults identified negative targets more accurately than neutral targets, $t(30) = 2.38$, $p < .05$, but young adults did not, $p = .10$." (Langley et al., 2008, p. 880).

Papers without statistically significant interactions.

Issue 1

Michael Poulin, Roxane Cohen Silver (p. 13–23)

No statistically 2×2 significant interactions were reported.

Jennifer Tehan Stanley, Fredda Blanchard–Fields (p. 24–32)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Abby Heckman Coats, Fredda Blanchard–Fields (p. 39–51)

No interactions were reported.

Robert S. Stawski, Martin J. Sliwinski, David M. Almeida, Joshua M. Smyth, (p. 52–61)

No statistically 2×2 significant interactions were reported.

Steven David, Bob G. Knight (p. 62–69)

No statistically significant 2×2 interactions were reported.

Gitit Kavé, Nitza Eyal, Aviva Shorek, Jiska Cohen–Mansfield (p. 70–78)

No statistically significant 2×2 interactions were reported.

Susan R. Old, Moshe Naveh–Benjamin (p. 104–118)

No statistically significant 2×2 interactions were reported.

Koene R.A. Van Dijk, Pascal W.M. Van Gerven, Martin P.J. Van Boxtel, Wim Van der Elst, Jelle Jolles (p. 119–130)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Elizabeth A.L. Stine–Morrow, Lisa M. Soederberg Miller, Danielle D. Gagne, Christopher Hertzog (p. 131–153)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Denis Gerstorf, Nilam Ram, Christina Røcke, Ulman Lindenberger, Jacqui Smith (p. 154–168)

No interactions were reported

Cynthia M. Torges, Abigail J. Stewart, Susan Nolen–Hoeksema (p. 169–180)

No statistically significant 2×2 interactions were reported.

Randolph Blake, Matthew Rizzo, Sean McEvoy (p. 181–189)

No interactions were reported.

Herbert Heuer, Mathias Hegele (p. 190–202)

No statistically significant 2×2 interaction reported.

Matthias Kliegel, Theodor Jóger, Louise H. Phillips (p. 203–208)

No interactions were reported.

Christina M. Leclerc, Elizabeth A. Kensinger (p. 209–215)

No statistically significant 2×2 interactions were reported.

Sarah R. Weatherbee, Jason C. Allaire (p. 216–221)

No interactions were reported.

Dannii Y. Yeung, Helene H. Fung, Frieder R. Lang (222–226)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Issue 2

Yvonne Brehmer, Shu–Chen Li, Benjamin Straube, Gundula Stoll, Timo von Oertzen, Viktor Mller, Ulman Lindenberger (p. 227–238)

No statistically significant 2×2 interactions were reported.

Nora A. Murphy, Derek M. Isaacowitz (p. 263–286)

No interactions were reported.

Shevaun D. Neupert, Daniel K. Mroczek, Avron Spiro III (p. 287–296)

No statistically significant interactions were reported.

Myra Fernandes, Michael Ross, Melanie Wiegand, Emily Schryer (p. 297–306)

No statistically significant 2×2 interactions were reported.

Arjun Kumar, Brian C. Rakitin, Rohit Nambisan, Christian Habeck, Yaakov Stern (p. 315–329)

There were not enough data in order to plot the available interactions.

Zheng Bian, George J. Andersen (p. 342–352)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Arndt Bröder, Andrea Herwig, Stefan Teipel, Kristina Fast (p. 353–365)

There were more than three levels in each one of the reported interactions.

Ryan P. Bowles, Timothy A. Salthouse (p. 366–376)

No interactions were reported.

Karen L. Fingerman, Laura Miller, Susan Charles (p. 399–409)

There were not enough data in order to plot the reported interactions.

Phebe Cramer (p. 410–421)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Hanna van Solinge, Kéne Henkens (p. 422–434)

No significant interactions were reported.

William von Hippel, Julie D. Henry, Diana Matovic (p. 435–439)

No interactions were reported.

Helene H. Fung, Derek M. Isaacowitz, Alice Y. Lu, Heather A. Wadlinger, Deborah Goren, Hugh R. Wilson (p. 440–446)

No statistically significant 2×2 interactions were reported.

Lynn M. Martire, Richard Schulz, Charles F. Reynolds III, Jennifer Q. Morse, Meryl A. Butters, Gregory A. Hinrichsen (p. 447–452)

No interactions were reported.

Elliot M. Tucker–Drob, Timothy A. Salthouse (p. 453–460)

No interactions were reported.

Ensar Becic, Walter R. Boot, Arthur F. Kramer (p. 461–466)

No statistically significant 2×2 interactions were reported.

Issue 3

Michaela Riediger, Alexandra M. Freund (p. 479–494)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Susan Turk Charles, Laura L. Carstensen (p. 495–504)

No statistically significant 2×2 interactions were reported.

Cynthia A. Berg, Deborah J. Wiebe, Jonathan Butner, Lindsey Bloor, Chester Bradstreet, Renn Upchurch, John Hayes, Robert Stephenson, Lillian Nail, Gregory Patton (p. 505–516)

No interactions were reported.

JoNell Strough, Joseph P. McFall, Jennifer A. Flinn, Kelly L. Schuller (p. 517–530)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Patrik Hansson, Michael Rönnlund, Peter Juslin, Lars–Goran Nilsson (p. 531–544)

No statistically significant 2×2 interactions were reported.

Mathias Allemand, Daniel Zimprich, Mike Martin (p. 545–557)

No statistically significant interactions were reported.

M. Brent Donnellan, Richard E. Lucas (p. 558–566)

No statistically significant 2×2 interactions were reported.

David L. Roth, Michelle L. Ackerman, Ozioma C. Okonkwo, Louis D. Burgio (p. 567–576)

No interactions were reported.

Lia Nower, Alex Blaszczynski (p. 577–584)

No interactions were reported.

Raluca Petrican, Morris Moscovitch, Ulrich Schimmack (p. 585–594)

No statistically significant 2×2 interactions were reported.

Stuart W.S. MacDonald, David F. Hultsch, Roger A. Dixon (p. 595–607)

No statistically significant 2×2 interactions were reported.

Cécile Proust–Lima, Héléne Amieva, Luc Letenneur, Jean–Marc Orgogozo, Héléne Jacqmin–Gadda, Jean–Francois Dartigues (p. 608–620)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Lisa Emery, Sandra Hale, Joel Myerson (p. 634–645)

There were either no statistically significant 2×2 interactions or not enough information were provided in order to plot the ones available.

Dary D. Fiorentino (p. 665–670)

No interaction reported.

Sheung–Tak Cheng, Ying–Kit Yim (p. 676–680)

No statistically significant 2×2 interactions were reported.

Issue 4

Sebastian Jessberger, Fred H. Gage (p. 684–691)

No interactions were reported.

C.S. Green, D. Bavelier (p. 692–701)

No interactions were reported.

John J. McArdle, John J. Prindle (p. 702–719)

No interactions were reported.

Shu–Chen Li, Florian Schmiedek, Oliver Huxhold, Christina Rcke, Jacqui Smith, Ulman Lindenberger (p. 731–742)

There were either no statistically significant 2×2 interactions or not enough information was provided in order to plot the ones available.

Martin Buschkuehl, Susanne M. Jaeggi, Sara Hutchison, Pasqualina Perrig–Chiello, Christoph Dpp, Matthias Mller, Fabio Breil, Hans Hoppeler, Walter J. Perrig (p. 743–753)

No interactions were reported.

Chandramallika Basak, Walter R. Boot, Michelle W. Voss, Arthur F. Kramer (p. 765–777)

No statistically significant 2×2 interactions were reported.

Elizabeth A.L. Stine–Morrow, Jeanine M. Parisi, Daniel G. Morrow, Denise C. Park (p. 778–786)

No interactions were reported.

Tracie Harrison, Shelley Blozis, Alexa Stuifbergen (p. 723–732)

No interactions were reported.

Isabelle Bauer, Carsten Wrosch, Joelle Jobin (p. 800–811)

The interactions reported concerned regression coefficients.

Christina Röcke, Margie E. Lachman (p. 833–847)

No statistically significant 2×2 interactions were reported.

David Bunce, Rowena Handley, Stanley O. Gaines Jr. (p. 848–858)

No statistically significant interactions were reported.

Julia Spaniol, Andreas Voss, Cheryl L. Grady (p. 859–872)

No statistically significant 2×2 interactions were reported.

Simon Forstmeier, Andreas Maercker (p. 886–899)

The interactions concerned regression coefficients.

Roger Ratcliff (900–916)

No interactions were reported.

Melissa Lunsman, Jerri D. Edwards, Ross Andel, Brent J. Small, Karlene K. Ball, Daniel L. Roenker (p. 917–927)

The interactions concerned regression coefficients.

Rebecca E. Ready, Janessa O. Carvalho, Mark I. Weinberger (p. 928–933)

No interactions were reported.

Dominic Abrams, Richard J. Crisp, Sibila Marques, Emily Fagg, Lauren Bedford, Dimitri Provias (p. 934–939)

The interactions concerned regression coefficients.

Elizabeth A. McDade–Montez, David Watson, Michael W. O'Hara, Natalie L. Denburg (p. 940–9467)

No interactions were reported.