Effect of Age on Variability in the Production of Text-Based Global Inferences

Lynne J. Williams[1]*, Joseph P. Dunlop[2], Hervé Abdi[2]*

1 Centre for Brain Fitness, Rotman Research Institute at Baycrest, Toronto, Ontario, Canada, **2** School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, Texas, United States of America

Abstract

As we age, our differences in cognitive skills become more visible, an effect especially true for memory and problem solving skills (i.e., fluid intelligence). However, by contrast with fluid intelligence, few studies have examined variability in measures that rely on one's world knowledge (i.e., crystallized intelligence). The current study investigated whether age increased the variability in text based global inference generation–a measure of crystallized intelligence. Global inference generation requires the integration of textual information and world knowledge and can be expressed as a gist or lesson. Variability in generating two global inferences for a single text was examined in young-old (62 to 69 years), middle-old (70 to 76 years) and old-old (77 to 94 years) adults. The older two groups showed greater variability, with the middle elderly group being most variable. These findings suggest that variability may be a characteristic of both fluid and crystallized intelligence in aging.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: lwilliams@research.baycrest.org (LJW); herve@utdallas.edu (HA)

Introduction

To understand spoken or written language, we need to integrate lexical, semantic, and contextual information and generate appropriate representations [1,2]. Obviously, this process is highly dependent upon knowledge and memory [2–10], which are both sensitive to aging. But what happens to language comprehension when we age? The generally accepted view suggests that memory for textual details declines as memory declines with age [2,4,6,11–13]. By contrast, however, older adults can access semantic information and understand complex linguistic representations as well as, or even better than, young adults in contexts where language comprehension is not dependent upon memory performance [2–10]. This apparent stability in older adults' language comprehension performance is intriguing because text comprehension is a very complex activity (see e.g., [14]) that typically involves remembering the gist of the text rather than the surface details [15–21].

Kintsch [15] and colleagues suggested, that in order to remember the information in a text we need to reduce the amount of information by transforming the verbatim information into an abstract version of the text (see also [22,23]). This abstract representation comes in the form of global inferences, which represent holistic concepts such as the theme or main point of a text [15,16,22–26]. These global inferences reduce the amount of information to be stored in memory because they integrate the text specific information with the individual's world knowledge and experience (i.e., extra-textual information). Moreover, because global inferences represent generalized information (i.e., the text information is extended to contexts beyond the text itself, see e.g., [9]), we generate global inferences in order to fill informational gaps within the text and this allows us to incorporate the information from the text into our own world knowledge [27,28].

Interestingly, the capacity to generate global inferences appears stable across age. For example, Ulatowska et al. [9] reported that there was no age difference in forming global representations of text in a longitudinal study of global inference generation in older adults. Similarly, Olness [29] found no differences between college-aged, middle-aged, and older adults in generating global inferences for didactic and non-didactic texts. Yet, there is growing evidence that knowledge structures thought to remain sable in aging–such as vocabulary and global inferences–may, in fact, be variable. For example, Christensen [30] found increased variability in older adults for measures of both memory, spatial, and reasoning skills (i.e., fluid intellectual abilities) as well as verbal abilities, including vocabulary (i.e., a crystallized intellectual ability). Similarly, Caskie, Schaie, and Willis [31] found considerable variability in verbal, spatial, and reasoning abilities in adults between 25 and 81 years of age. In addition, the patterns of variability were different for verbal abilities versus spatial and reasoning skills. In particular the changes in verbal abilities showed later onset, greater variability in the timing of onset, as well as greater variability in the overall rate of change. At the level of text comprehension, Hertzog, Dixon, and Hultsch [32] found significant variability in memory for textual information not accounted for by text-related factors in a group of seven elderly women. Likewise, Dixon and colleagues [33] reported an age increased variability in text recall for the stories used in the Logical Memory subtest of the Wechsler Memory Scale. Together all these findings suggest that an age-related increase in variability of the knowledge structures underlying linguistic ability and global inference generation may be a hallmark of cognitive aging, in the same

way as the age-related increase in variability in reaction time, memory, and other cognitive abilities [34–42]. Therefore, we decided to investigate if age increased the variability of global inferences of participants. In order to do so, we measured age-related variability in generating global inferences among three groups of older adults.

Results

Thirty-four participants between the ages of 62 and 94 years were divided into three age groups for the purpose of analysis. The young-old (Y-O) group consisted of 12 individuals (62 to 69 years of age), the middle-old (M-O) and old-old (O-O) elderly groups consisted of 11 participants each (70 to 76 and 77 to 94 years of age, respectively). Each participant gave 2 possible lessons for each of 12 Aesop fables ([43], see Supplementary Information S.1). Each lesson was scored categorically according to the criteria outlined in the Method section 4.3.1. Data were analyzed using discriminant correspondence analysis (DICA) [44–52].

DICA is a multivariate technique developed to classify observations described by qualitative and/or quantitative variables into a-priori defined groups and has been used to discriminate clinical populations, such as early versus middle stage Alzheimer's disease [51] and autistic paranoia from paranoid schizophrenia [52]. Based on correspondence analysis (CA), DICA is a type of principal component analysis (PCA)–specifically tailored for the analysis of categorical data–that represents the rows and columns as points in (a high dimensional) space [45,49–51,53–57]. Just like PCA, DICA finds the most important dimensions of variance of the data. These dimensions are uncorrelated to each other and ordered by the amount of the data variance that they explain. Rows and columns can be plotted as maps by using their coordinates on these dimensions. In order to reveal the pattern of variables associated with group differences, DICA analyzes a data table in which each row sums the behavior of the participants of a given group (see [51] for more information). DICA is then obtained from the CA of this summed table. This analysis reveals the similarities and differences in patterns of performance across the age groups. See Method section 4.3.3, File S2, Figure S1 and [44,51,58,59] for more information.

The DICA derived two factors accounting, respectively, for 85 percent and 15 percent of the data variance. The eigenvalues (λ), proportion of explained variance (τ), and the contributions of each variable and group to the total variance for Factors 1 and 2 are shown in Table 1. The higher the contribution, the more important that variable (or observation) is in defining a given factor.

2.1 Age-related Patterns of Global Inference Generation

The DICA uncovered age related patterns in lesson generation performance. Factor 1 separated the Y-O from the M-O and O-O groups (see Figure 1). Because DICA reliably separated the Y-O from the other groups, the effect size is quite large and is detectable with our current sample size. However, to ensure that we could detect a reasonable effect size, we computed an a-posteriori effect size analysis using G*Power 3 [60,61]. For the purpose of power analysis, multivariate discriminant analysis can be considered under the MANOVA framework [62,63]. With an α of .05 and achieved power $(1 - \beta)$ of .95, we had an effect size $(f^2(V))$ of .41. This effect size is equivalent to a critical Pillai's V of 0.6 across the 3 groups, meaning that the between group-variance is 60% of the total variance. This effect size and critical V were considered adequate to be able to discriminate between the Y-O, M-O, and O-O groups.

Table 1. Eigenvalues (λ), proportion of inertia (τ), contributions of the age groups and scoring categories for Factors 1 and 2.

	Factor 1	Factor 2
Eigenvalue (λ)[a]	0.0105	0.0018
Proportion of Explained variance (τ)	0.8539	0.1461
Contributions[b]		
Age Group		
Young Elderly	0.6449	0.0022
Middle Elderly	0.2118	0.4646
Old Elderly	0.1433	0.5332
Scoring Categories		
Switch Perspective		
Switch	0.1281	0.0004
Paraphrase	0.3344	0.0009
Linguistic Form[c]		
Proverb	0.0207	0.3624
Non-Proverb	0.0037	0.0653
Generalization Level (LESSON 1)		
Extratextual	0.0648	0.0061
Text Specific	0.0570	0.0054
Generalization Level (LESSON 2)		
Text Specific	0.0412	0.0017
Extratextual	0.0977	0.0041
Viewpoint Adopted (LESSON 1)		
Main Character	0.0523	0.0102
Supporting Character	0.1081	0.0085
Other	0.0290	0.1517
Viewpoint Adopted (LESSON 2)		
Main Character	0.0274	0.1221
Supporting Character	0.0347	0.2448
Other	0.0009	0.0162
Representation of Theme[d] (LESSON 1)		
Accurate	n/a	n/a
Inaccurate	n/a	n/a
Representation of Theme[d] (LESSON 2)		
Accurate	n/a	n/a
Inaccurate	n/a	n/a

[a]Note that in correspondence analysis, the eigenvalues (λ) are never greater than 1.
[b]Contributions are the proportion of variance of a given factor explained by the age group or scoring category.
[c]Proverbial form for lesson 1 and lesson 2 combined due to similar profiles in previous versions of the analysis.
[d]Representation of Theme was included as a supplementary element, therefore, it did not contribute to the explained inertia of the factors.
doi:10.1371/journal.pone.0036161.t001

The results of the DICA are shown in Figure 1. The scoring categories are shown in separate displays to help reading the map. The variable contributing the most to Factor 1 is "switching perspectives between lesson types." The young elderly group more frequently switched perspectives than the middle and old elderly groups. Success in switching between LESSON 1 and LESSON 2 is more strongly associated with a LESSON 1 that incorporated information from outside of the text (i.e., extra-textual) and
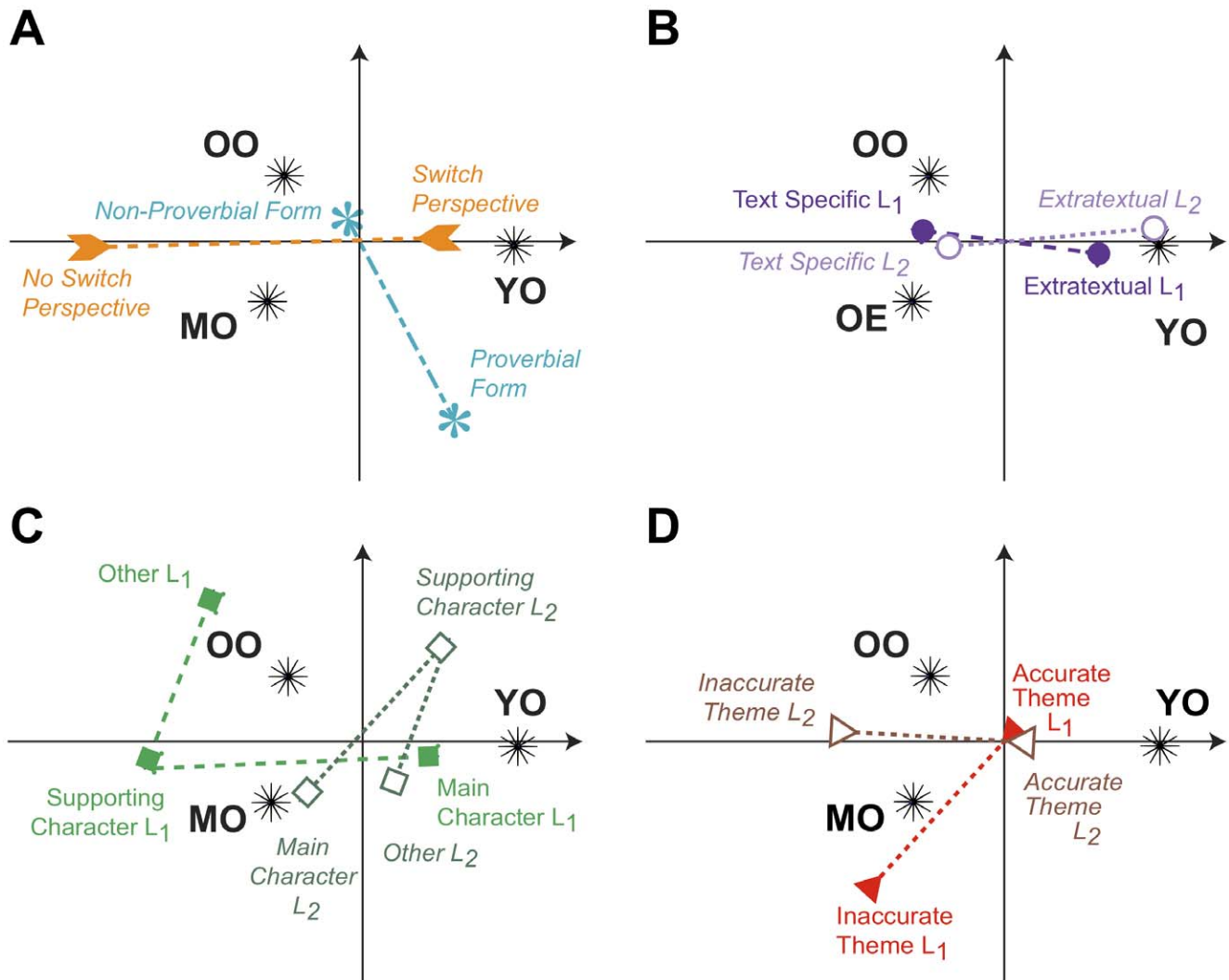
Figure 1. Discriminant correspondence analysis. Variables shown along Factors 1 and 2. Lambda (λ) and tau (τ) are the eigenvalues and the percentage of explained inertia (i.e., variance) for a given factor ($\lambda_1 = .0105$, $\tau_1 = .8539$; $\lambda_2 = .0018$, $\tau_2 = .1461$). All sub-figures are plotted on the same scale along each factor. (A) Switch Perspective and Linguistic Form collapsed across both lesson types. Note that young elderly group switched perspectives between lesson types more frequently than the middle or old elderly groups. (B) Generalization Level for each lesson type. Note that the young elderly group produced extra-textual lessons more frequently. Extra-textual lessons incorporate information from outside of the text. (C) Character Viewpoint for each lesson type. Note that they young elderly group more frequently adopted the viewpoint of the main character for the best lesson (LESSON 1) and the supporting character for the alternate lesson (LESSON 2). (D) Representation of Theme was included as a supplementary element. Supplementary elements are variables that were not included in the calculations, but were projected into the space to see their placement along the factors. They are used to aid with interpretation. Note that the young elderly group more frequently produced lessons reflecting accurate fable themes for both LESSON 1 and LESSON 2. Note that in correspondence analysis, the eigenvalues are never greater than 1.
doi:10.1371/journal.pone.0036161.g001

represented the main character's viewpoint. Successful switches in perspective also were more frequently stated as proverbs and showed themes consistent with the fable for both lesson types. By contrast, the middle and older elderly groups switched perspectives less frequently than the young elderly group. Failure to switch perspective between lesson types was associated with maintaining the main character's viewpoint for LESSON 2 and producing text specific lessons for both lesson types (i.e., the information content of each lesson did not go beyond information stated explicitly in the fable). Switch failures also were characterized by more frequent use of non-proverbial linguistic forms (i.e., a concrete interpretation) and inaccurate representation of the fable theme for both LESSONs 1 and 2.

Factor 2 distinguished the middle and old elderly groups. The middle elderly group had a slightly greater tendency to maintain the main character's perspective on LESSON 2. Furthermore, the middle elderly group produced LESSON 1 showing an inaccurate fable theme more frequently than those produced by the old elderly group. However, the old elderly group's LESSON 1 had a slightly greater tendency, on average, to adopt neither the main nor the supporting character's perspectives. The old elderly group also tended to state both lessons in a non-proverbial form.

The performance of the groups and the individual participants by age group are shown in Figure 2. The young elderly participants clustered more tightly together, indicating that they were predominantly successful in switching perspectives. The tight
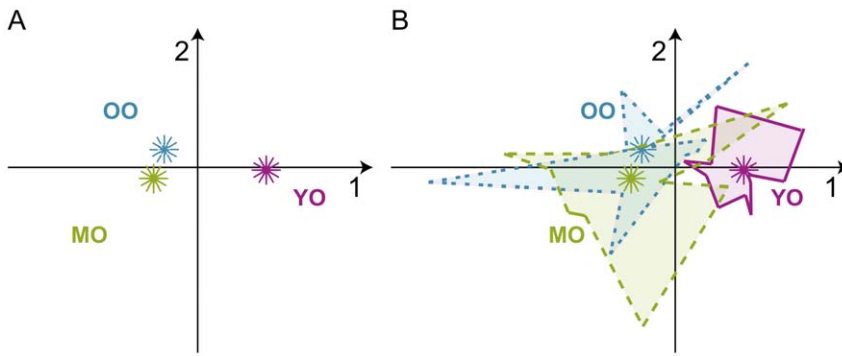
Figure 2. Discriminant correspondence analysis. Participants shown by age group along Factors 1 and 2. Lambda (λ) and tau (τ) are the eigenvalues and the percentage of explained inertia (i.e., variance) for a given factor ($\lambda_1 = .0105$, $\tau_1 = .8539$; $\lambda_2 = .0018$, $\tau_2 = .1461$). All sub-figures are plotted on the same scale along each factor. (A) Barycenters (weighted average) of the groups, (B) Convex hull. The convex hull represents the average performance of individual participants within each age group. Individual participants were projected into the DICA space as supplementary elements. Supplementary elements are variables or observations that were not included in the calculations, but were projected into the space to see their placement along the factors. Note that in correspondence analysis, the eigenvalues are never greater than 1.
doi:10.1371/journal.pone.0036161.g002

grouping also suggests that the young elderly group showed less between participant variability in generating lessons. The middle and old elderly participants, by contrast, were more dispersed. Some of the middle and old elderly participants showed a pattern of lesson generation similar to the young elderly participants, while others did not. This suggests greater between participant variability, especially in the ability to switch perspectives between LESSONS 1 and 2.

2.2 Variability in Global Inference Generation

The variability in generating global inferences within the age groups was evaluated using a bootstrap procedure [64–66]. The bootstrap produced 95% confidence interval ellipses for each age group (see Figure 3; a description of the bootstrap is presented in the File S2.6.2). The area of a confidence interval ellipse represents the variability within each group. When the confidence ellipses do not overlap there is a significant difference between the groups at the $p = .05$ level. Consequently, the confidence ellipses show that the young elderly group is reliably different from the middle and

old elderly groups because there is no overlap with the confidence ellipses of the other two groups. In addition, the size of the young elderly group's ellipse is smaller, indicating that there is less variability within this group. Although the middle and old elderly groups were not reliably distinguished, the middle elderly group, surprisingly, had the ellipse with the greatest area indicating that the middle elderly group showed the most variability (see also Figures 2A and 2c for actual dispersion in group performance).

2.3 Quality of the DICA Model

We evaluated the quality of our DICA model by computing the amount of variance explained by the DICA ($R^2 = .23$, $p < .01$; see Supplementary Information for details). We also evaluated how the model would generalize to new participants by using a jackknife procedure (also called "leave one out" procedure). The jackknife procedure [64,67,68] removes, in turn, each of the participants from the sample and performs a new DICA on the remaining participants. The distance between the removed participant (projected into the new DICA space as a supplementary element) and each of the groups is computed and the participant is assigned to the closest group (see [44] and [68] for more information about the jackknife in DICA). The results of the jackknife are summarized in Table 2. The columns represent the original group assignment and the rows represent the DICA assignment. From Table 2, we found that of the 34 possible assignments, only 13 were correct. The young elderly participants were more reliably assigned to their group (9 out of 12 correctly assigned) than participants from the middle and old elderly groups (2 correct assignments out of 11 participants for each group). This
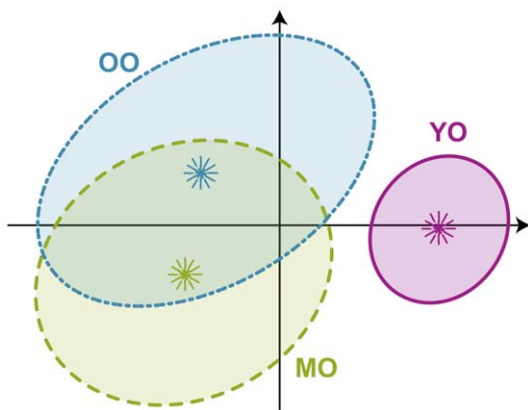

Figure 3. Discriminant correspondence analysis. 95% confidence intervals for age groups shown on factors 1 and 2. Lambda (λ) and tau (τ) are the eigenvalues and the percentage of explained inertia (i.e., variance) for a given factor ($\lambda_1 = .0105$, $\tau_1 = .8539$; $\lambda_2 = .0018$, $\tau_2 = .1461$). Confidence ellipses represent the variability within the group. Ellipses showing no overlap represent different populations. Note that in correspondence analysis, the eigenvalues are never greater than 1.
doi:10.1371/journal.pone.0036161.g003

Table 2. Actual versus dica participant classification into young elderly, middle elderly, and old elderly groups.

DICA Assigned Group	Actual Group		
	Young Elderly	Middle Elderly	Old Elderly
Young Elderly	9	3	4
Middle Elderly	1	2	5
Old Elderly	2	6	2

Note: Numbers in italicized print represent correctly classified participants.
doi:10.1371/journal.pone.0036161.t002

Table 3. Participant characteristics.

	Age (yrs)		Education (yrs)		Similarities[a]		LM I[b]	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Young Elderly	65.58	(2.27)	16.25	(3.11)	27.50	(4.19)	18.67	(2.46)
Middle Elderly	72.18	(2.32)	16.91	(3.39)	27.27	(2.97)	18.45	(2.98)
Old Elderly	82.73	(5.41)	14.36	(3.26)	22.40	(8.46)	18.10	(2.64)
	LM II[b]		Trails (B — A)[c]		WCST (Total Correct)[d]			
	Mean	SD	Mean	SD	Mean	SD		
Young Elderly	26.00	(7.31)	29.13	(22.48)	50.58	(7.56)		
Middle Elderly	23.00	(8.98)	43.06	(35.06)	43.36	(8.37)		
Old Elderly	14.80	(8.98)	125.92	(98.87)	42.40	(12.21)		

Note: All scores from psychometric testing are represented as raw scores for the given subtest;
[a]Wechsler Adult Intelligence Scale III [110];
[b]Wechsler Memory Scale III [111];
[c]Trail Making Test [112];
[d]Wisconsin Card Sorting Test, 64 Card Version [113].
doi:10.1371/journal.pone.0036161.t003

difference in classification reflects the larger variability of the middle and old elderly groups.

Discussion

Although most studies examining cognitive performance variability in the elderly have examined skills that are known to decrease with age (e.g., fluid intelligence abilities, reaction time (RT) or memory [39,40,69–83]), skills that remain stable or improve across age (i.e., crystallized intelligence) also show inter-trial variability. However, this age-associated pattern of variability may differ between the two intelligence domains. For example, variability in RT for speeded tasks shows that older adults are consistently more variable than younger adults [81–83] and that this increased RT variability is associated with poorer cognitive performance in normally aging older adults [84–86]. The general increase in variability in the M-O and O-O groups relative to the Y-O group supports this view and may be associated with the older two groups' general difficulty in switching perspectives between LESSON 1 and LESSON 2.

By contrast, when older adults show increased variability in gist recall accuracy (rather than RT or detail recall), this increase in variability tends to be associated with poor health, rather than normal age-related change [32,87]. In normally aging adults, increased item-to-item variability in non-speeded tasks (such as gist recall) is associated with higher mean performance and may actually be an indicator of learning rather than decline [76,88]. The finding that the M-O group showed greater variability than the O-O group suggests that, at least in non-speeded tasks, increased variability may not be completely maladaptive. The strict view of increased variability indicating cognitive decline predicts a linear association between variability and age (see e.g., [81–86]), yet the current data do not show this pattern. Rather, they suggest the possibility varying patterns of variability at different life stages, especially given that the Y-O, M-O, and O-O individuals were cognitively normal and successfully performed the task.

If we consider that learning may also be a mechanism for increased variability in aging, then the M-O group would be expected to show the greatest amount of variability because this group has the largest proportion of recently retired individuals undergoing a major life change. For example, Adam and colleagues [89,90] have found sudden decreases in cognitive functioning immediately following retirement, a pattern which suggests that there may be an increase in variability in cognitive performance around this time. Such a change in variability would be similar to the recursive increases in variability and subsequent plateauing during periods of social and cognitive development during childhood and adolescence [88].

Although showing increased variability associated with age, the current results show a mixed pattern. This suggests that multiple mechanisms may underlie the increase in performance variability for crystallised intellectual abilities in older adults and that the relationship between age and variability may not be as straightforward as with fluid intellectual skills. Nevertheless, these findings show that variability with age may not be just an indicator of decline, but may also signal new learning. As Garrett et al. [91] so aptly said, "variability is more than just noise" (p. 4914).

Methods

4.1 Participants

Thirty-four participants between the ages of 62 and 94 years were divided into three age groups for the purpose of analysis. The young elderly (Y-O) group consisted of 12 individuals (62 to 69 years of age), the middle (M-O) and older (O-O) elderly groups consisted of 11 participants each (70 to 76 and 77 to 94 years of age, respectively). All participants were highly educated, with an average of 15 years of formal education. All participants were living in the community and were self-reported native English speakers. None exhibited clinical signs of impaired cognitive performance as tested by the 7 Minute Screen [92,93]. All participants scored within normal age limits on a hearing screening that included the Erber Sentences [94], CID Sentences [95], and a self-report of hearing loss. All subjects made no errors on a visual narrative screener where they read aloud an additional fable typeset in the same font as the stimulus fables. This study was approved by the Internal Review Board (IRB) of the University of Texas at Dallas. All participants gave written informed consent. Table 3 gives the participant characteristics.

Table 4. Scoring Criteria for LESSON 1 and LESSON 2 Lesson Responses.

	Definition	Example
Switching Perspective		
Switch	Lesson 2 represents a different idea, theme, or viewpoint than LESSON 1	LESSON 1: Be careful whom you trust; LESSON 2: Sometimes good intentions go astray
Fox & Goat [Px O-O 7]		
Paraphrase	LESSON 2 represents the same general idea, theme, or viewpoint as LESSON 1; Lessons may be stated in a different linguistic form.	LESSON 1: United we stand, divided we fall; LESSON 2: Strength in numbers
Father & Sticks [Px M-O 8]		
Generalization Level[a]		
Text Specific	Lesson remains tied to the characters, actions, or events of the fable.	Make sure that what you say is true, because you may need to prove it
Boasting Traveller [Px Y-O 7]		
Extratextual	Lesson extends beyond the actions and events of the fable. Contains inferred information not contained in the original fable.	Gold and rubies are not the only treasures in this world
Farmer & Sons [Px Y-O 9]		
Viewpoint Adopted		
Main Character	Lesson is stated from the viewpoint of the main fable character	When you cry, make sure its true. Or when you do anything or want help from anyone, make sure its true
Shepherd Boy [Px Y-O 3]		
Supporting		
Character	Lesson is stated from the viewpoint of the supporting fable character	People get tired of being made fools of and they eventually learn not to respond
Shepherd Boy [Px Y-O 10]		
Other	Lesson is stated from a perspective that does not distinctly adopt the viewpoint of either the main or supporting fable characters; Lesson may reflect a mixed or indeterminate viewpoint, or is not character specific	The moral of the story is that each of us has a gift
		Crane & Peacock [Px M-O 7]
		That there's not, um, a same fix for each person
		Father & Sticks [Px Y-O 5]
Linguistic Form		
Proverbial	Lesson is stated in the form of a proverb	United we stand, divided we fall
Father & Sticks [Px O-O 2]		
Non-Proverbial	Lesson is not proverbial	Try not to take on more than you can handle in detail. Try and break it down in smaller amounts and complete each effort singly
Father & Sticks [Px O-O 7]		
Representation of Theme[b]		
Accurate	The theme represented in the lesson is appropriate for the fable	The proof of the pudding is in the eating. If you have to if you have to boast, then you should be able to perform what you're saying
Boasting Traveller [Px O-O 2]		
Inaccurate	The theme represented in the lesson is inappropriate for the fable	If you pick at a sore, it will get worse
Miser [Px M-O 11]		

Y-O = Young elderly, M-O = Middle elderly, O-O = Old elderly;
[a]Modified from Olness [29];
[b]Scoring category included as supplementary variable in the dica analysis due to high frequency of accurate responses. A supplementary variable is one that was not included in the analysis, but was placed in the display to aid with interpretation of the factors.
doi:10.1371/journal.pone.0036161.t004

4.2 Stimuli and Task

We selected twelve short narratives from George Townsend's translations of Aesop's fables [43]. We used fables because cultural knowledge is transmitted via their didactic form. This transmission of cultural knowledge takes the form of a lesson or moral (i.e., types of global inferences). In addition, the role of the fables in transmitting knowledge or "general truths" gives to the fables a similar function to proverbs in discourse. However, unlike proverbs, fables require the theme, lesson, or moral to be inferred from the characters' actions and their consequences. Meaning in proverbs, by contrast, is derived from the text itself and not from its application to real-world contexts because proverbs are already stated in a global inference-like format [9,96,97]. Because fables are didactic, readers can interpret them at two levels: literally, at

the level of the text itself (i.e., a textual interpretation), or metaphorically, as a guide to culturally appropriate behavior in real-life contexts (i.e., an extratextual interpretation; [98–102]). Given that multiple interpretations of each fable is possible, fables can be, at least in part, interpreted as each reader chooses [103] and therefore interpretations of a given fable can vary with the reader, the information that is chosen as salient during comprehension (e.g., a given character's actions), and the overall level of generalization (i.e., textual versus extratextual).

All fables employed two characters, contained three episodes (i.e., setting, action, and resolution components), were between 10 and 21 propositions in length [15], and contained no mixture of anthropomorphized animal and human characters. We then modified the fables to exclude specific mention of character attributes (e.g., lazy, wise, etc.) and any specific mention of the moral or lesson. Fables are shown in File S1. We asked participants to generate two different lessons or morals for each of the 12 fables. We instructed participants to first give what they considered to be the "best" lesson for the fable (LESSON 1). We then asked participants to generate a second possible lesson for each fable that reflected a different interpretation or perspective (LESSON 2). The examiner read the fable to participants and a card with the printed fable was within view during generation of both lessons to minimize memory demands.

4.3 Analysis

4.3.1 Response coding. Lessons were scored categorically according to: (a) whether there was a switch in perspective between LESSON 1 and LESSON 2, (b) whether the lesson reflected text specific or extratextual content [9,29], (c) whether the lesson portrayed the viewpoint of the main or of the supporting character [98], and (d) whether the lesson was given in the form of a statement or proverb, that is, a literal or metaphorical interpretation [104]. The accuracy or semantic fit of each lesson theme was scored in reference to the original fable. Representation of theme was not included as an active variable in the analysis due to the high degree of accurate semantic representation produced by all three age groups (91% accurate). Table 4 shows further definitions of the scoring categories with examples.

4.3.2 Inter-rater reliability. Inter-rater reliability was analyzed on a random 20% of the data by comparing the first author's coding with the code ratings of a second trained rater. Point-by-point agreement was 79%. A Cohen's Kappa was calculated to correct for chance agreement ($\kappa = 0.621$), corresponding to a "substantial" rating of agreement [105].

4.3.3 Statistical analysis. We used discriminant correspondence analysis (DICA) to analyze the coded lesson responses. DICA combines the features of correspondence analysis (CA) and discriminant analysis ([44,106]; see also [51] for a tutorial on language oriented applications). Correspondence analysis (CA) is a type of principal component analysis (PCA)–specifically tailored for the analysis of categorical data–that represents the rows and columns as points in a (high dimensional) space [45,49,50,53–55]. In addition, CA (and consequently, DICA) can handle data sets with few observations described by many nominal variables [44,45,51,107].

Just like PCA, CA finds orthogonal factors or dimensions that reveal the patterns and the associations between the row and column profiles. The importance of the factors is determined by their inertia (i.e., a quality akin to variance), denoted by λ and the proportion of explained inertia, denoted by τ. CA converts contingency tables into visual displays (i.e., maps) in which the row profiles and column profiles represent points in the display. The proximity of the points within the display represents their

degree of associa'tion. Points distributed more closely in space are more strongly associated than those that are farther apart. In addition, CA places no constraints on the data; therefore, the pattern seen in the maps represents associations contained within the data and not those superimposed by an external model [47,48].

DICA is a multivariate technique developed to classify observations described by qualitative and/or quantitative variables into a-priori defined groups and therefore adds a discriminative component to CA. Here, we used DICA to analyze LESSON 1 and LESSON 2 responses and to classify participants into pre-defined age categories: young-old (Y-O), middle-old (M-O) and old-old (O-O) groups.

For the DICA, participants were grouped into the three age groups. Then, the pattern of performance of the participants in each group was combined into its common pattern of performance (see [51] for more information on how the common pattern is developed). Table 5 shows the age-group by lesson response contingency table, the common pattern of performance used for the DICA in the current study.

We then ran a CA on the common performances, which allowed us to examine the similarities and differences in patterns of performance across the age groups. CA and DICA also can be used

Table 5. Frequency of occurrence of scoring categories by lesson type for the young elderly, middle elderly, and old elderly groups (contingency table input into dica).

	Switch Perspective		Linguistic Form	
	Switch	Paraphrase	Proverbial	Non-Proverbial
Young Elderly	119	25	25	119
Middle Elderly	87	45	22.5	109.5
Old Elderly	89	43	14.75	117.25
	Generalization Level			
	LESSON 1		LESSON 2	
	Extratextual	Text Specific	Extratextual	Text Specific
Young Elderly	76	68	51	93
Middle Elderly	58	74	34	98
Old Elderly	57	75	36	96
	Character Viewpoint			
	LESSON 1			LESSON 2
	Main	Supporting	Other	Main
Young Elderly	113	19	12	72
Middle Elderly	91	28	13	78.67
Old Elderly	89	26	17	69.34
	Character Viewpoint			
	LESSON 2			
	Supporting	Other		
1−3 Young Elderly	57	15		
Middle Elderly	39.66	13.67		
Old Elderly	50.32	12.34		

to estimate the amount of variability within and between each category. To do this 95% confidence ellipses are constructed using a bootstrap resampling technique ([108,109]; see also File S2.6.2). A detailed mathematical appendix is included in the Supplementary Information.

Supporting Information

Figure S1
(TIFF)

File S1 Supporting Information PDF file.
(PDF)

File S2
(PDF)

Author Contributions

References

1. Burke DM (2006) Representation and aging. In: Bialystok E, Craik FIM, editors, Lifespan Cognition: Mechanisms of Change, Oxford University Press. pp 193–206.
2. Chesneau S, Jbabdi S, Champagne-Lavau M, Giroux F, Ska B (2007) Comprehension de textes, ressources cognitives et vieillssement [Text comprehension, cognitive resources and aging]. Psychologie et NeuroPsychiatrie du Viellissement 5: 47–64.
3. Hamm V, Hasher L (1992) Age and the availability of inferences. Psychology and Aging 7: 56–64.
4. Hosokawa A, Hosokawa T (2006) Cross-cultural study on adult age-group differences in the recall of the literal and interpretive meanings of narrative text. Japanese Psychological Research 2: 77–90.
5. Laver GD, Burke DM (1993) Why do semantic priming effects increase in old age? A metaanalysis. Psychology and Aging 8: 34–43.
6. Meyer BI (1987) Reading comprehension and aging. In: Schaie KW, editor, Annual Review of Gerontology and Geriatrics (Volume 7), New York: Springer.
7. Radvansky GA, Zwaan RA, Curiel JM, Copeland DE (2001) Situation models and aging. Psychology and Aging 16: 145–160.
8. North A, Ulatowska HK, Macaluso-Haynes S, Bell H (1986) Discourse performance in older adults. International Journal of Aging and Human Development 23: 267–283.
9. Ulatowska HK, Chapman SB, Highley Amy P PJ (1998) Discourse in healthy old-elderly adults: A longitudinal study. Aphasiology 12: 619–663.
10. Zacks RT, Hasher L (1988) Integrating information from discourse: do older adults show deficits? In: Light LL, editor, Language, Memory, and Aging, Cambridge: Cambridge University Press. pp 117–132.
11. Adams C, Labouvie-Vief G, Hobart CJ, Dorosz M (1990) Adult age group differences in story recall style. Journals of Gerontology: Psychological Sciences 45: 17–27.
12. Adams C (1991) Qualitative age differences in memory for text: a life-span developmental perspective. Psychology and Aging 6: 323–336.
13. Park D, Schwarz N (1999) Cognitive aging: A primer. Philadelphia, PA: Psychology Press.
14. Chastaing M, Abdi H (1980) La psychologie des injures [The psychology of insults]. Journal de Psychologie Normale et Pathologique 77: 31–62.
15. Kintsch W (1998) Comprehension: A Paradigm for Cognition. Cambridge: Cambridge University Press.
16. van Dijk TA, Kintsch W (1983) Strategies of Discourse Comprehension. New York: Academic Press.
17. Stahl C, Klauer KC (2008) A simplified conjoint recognition paradigm for the measurement of gist and verbatim memory. Journal of Experimental Psychology Learning, Memory, and Cognition 34: 570–586.
18. Reyna VF, Kiernan B (1994) Development of gist versus verbatim memory in sentence recognition: effects of lexical familiarity, semantic content, encoding instructions, and retention interval. Developmental Psychology 30: 178–191.
19. Gutchess AH, Schacter DL (2012) The neural correlates of gist-based true and false recognition. NeuroImage 59: 3418–3426.
20. Abdi H (1990) Additive-tree representation of verbatim memory. CUMFID 16: 99–124.
21. Abdi H (1985) Représentations arborées de l'information verbatim [Additive tree representations of verbatim information]. Bulletin de Psychologie 38: 633–643.
22. Graesser AC, Singer M, Trabasso T (1994) Constructing inferences during narrative text comprehension. Psychological Review 101: 371–395.
23. McKoon G, Ratcliff R (1992) Inference during reading. Psychological Review 99: 440–446.
24. Schank R, Abelson R (1977) Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures. Hillsdale, NJ: Lawrence Erlbaum Associates.
25. van Dijk TA (1977) Semantic macro-structures and knowledge frames in discourse comprehension. In: Just MA, Carpenter PA, editors, Cognitive Processes in Comprehension, Hillsdale, NJ: Lawrence Erlbaum Associates.
26. van Dijk TA (1980) Macrostructures. Hillsdale, NJ: Lawrence Erlbaum Associates.
27. Magliano J, Graesser AC (1991) A three-pronged method for studying inference generation in literary text. Poetics 20: 193–232.
28. Seifert CM, Robertson SP, Black JB (1985) Types of inferences generated during reading. Journal of Memory and Language 24: 405–422.
29. Olness G (2000) Expression of Narrative Main-Point Inferences in Adults: A Developmental Perspective. Doctoral dissertation, University of Texas at Dallas, Richardson, TX.
30. Christensen H (2001) What cognitive changes can be expected with normal ageing? Australian and New Zealand Journal of Psychiatry 35: 768–775.
31. Caskie GIL, Schaie KW, Willis SL (1999) Individual differences in the rate of change in cognitive abilities during adulthood. Paper presented at the Gerontological Society of America Conference.
32. Hertzog C, Dixon RA, Hultsch DF (1992) Intraindividual change in text recall of the elderly. Brain and Language 42: 248–269.
33. Dixon RA, Hertzog C, Friesen I, Hultsch DF (1993) Assessment of intraindividual change in text recall of elderly adults. In: Brownell HH, Joanette Y, editors, Narrative Discourse in Neurologically Impaired and Normally Aging Adults, San Diego: Singular. pp 77–101.
34. Hultsch DF, Hunter M, MacDonald SWS, Strauss E (2005) Inconsistency in response time as an indicator of cognitive aging. In: Duncan J, Phillips L, McLeod P, editors, Measuring the Mind: Speed, Control, and Age (2nd ed.), Oxford: Oxford University Press.
35. Luszcz MA (2004) Whats it all about: variation and aging. Gerontology 50: 5–6.
36. Martin M, Hofer SM (2004) Intraindividual variability, change, and aging: conceptual and analytical issues. Gerontology 50: 7–11.
37. Miller MB, Van Horn JD (2007) Individual variability in brain activations associated with episodic retrieval: a role for large-scale databases. International Journal of Psychophysiology 63: 205–13.
38. Nesselroade JR, Ram N (2004) Studying intraindividual variability: what we have learned that will help us understand lives in context. Research in Human Development 1: 9–29.
39. Ram N, Rabbitt P, Stoller B, Nesselroade JR (2005) Cognitive performance inconsistency: intraindividual change and variability. Psychology and Aging 20: 623–633.
40. Shammi P, Bosman E, Stuss DT (1998) Aging and variability in performance. Aging, Neuropsychology, and Cognition 5: 1–13.
41. Fiske DW (1957) The constraints on intra-individual variability in test responses. Educational and Psychological Measurement 17: 317–337.
42. Daly DL, Bath KE, Nesselroade JR (1963) On the confounding of inter- and intraindividual variability in examining change patterns. Intelligence. pp 33–36.
43. Aesop/Townsend GF Aesops fables by Aesop. URL Retrieved September 20, 2008 from http://www.gutenberg.org/dirs/etext91/aesop11h.htm and http://www.gutenberg.org/etext/21: Project Gutenberg.
44. Abdi H (2007) Discriminant correspondence analysis. In: Salkind NJ, editor, Encyclopedia of Measurement and Statistics, Thousand Oaks, CA: Sage. pp 270–275.
45. Abdi H, Williams LJ (2010) Correspondence analysis. In: Salkind NJ, editor, Encyclopedia of Research Design, Thousand Oaks, CA: Sage. pp 267–278.
46. Abdi H, Valentin D (2007) Multiple correspondence analysis. In: Salkind NJ, editor, Encyclopedia of Measurement and Statistics, Thousand Oaks, CA: Sage, volume 95. pp 651–657.
47. Benzécri JP (1973) L'Analyse des Données. Tome 1: La Taxonomie [Data Analysis Volume 1: Taxonomy]. Paris: Dunod.
48. Benzécri JP (1973) L'Analyse des Données. Tome 2: L'Analyse des Correspondances [Data Analysis, Volume 2: Correspondence Analysis]. Paris: Dunod.
49. Greenacre M (1984) Theory and Applications of Correspondence Analysis. London: Academic Press.
50. Greenacre M (2007) Correspondence Analysis in Practice (Second ed.). Boca Raton, FL: Chapman & Hall/CRC.
51. Williams LJ, Abdi H, French R, Orange JB (2010) A tutorial on multi-block discriminant correspondence analysis (MUDICA): A new method for analyzing discourse data from clinical populations. Journal of Speech Language and Hearing Research 53: 1372–1393.
52. Pinkham AE, Sasson NJ, Beaton D, Abdi H, Kohler CG, et al. Qualitatively distinct factors contribute to elevated rates of paranoia in autism and schizophrenia. Journal of Abnormal Psychology, (in press), 121.

53. Abdi H, Williams LJ (2010) Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics 2: 433–459.

54. Blasius J, Greenacre M (2006) Correspondence analysis and related methods in practice. In: Greenacre M, Blasius J, editors, Multiple Correspondence Analysis and Related Methods, Boca Raton, FL: Chapman & hall. pp 3–40.

55. Clausen SE (1998) Applied Correspondence Analysis: An introduction. Thousand Oaks, CA: Sage.

56. Abdi H (2010) Partial least squares regression and projection on latent structure regression (PLS Regression). Wiley Interdisciplinary Reviews Computational Statistics 2: 97–106.

57. Krishnan A, Williams LJ, McIntosh AR, Abdi H (2011) Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. NeuroImage 56: 455–75.

58. Abdi H, Williams LJ, Connolly AC, Gobbini MI, Dunlop JP, et al. (2012) Multiple Subject Barycentric Discriminant Analysis (MUSUBADA): How to assign scans to categories without using spatial normalization. Computational and Mathematical Methods in Medicine.

59. Abdi H, Williams LJ, Valentin D, Bennani-Dosse M (2012) STATIS and DISTATIS: optimum multitable principal component analysis and three way metric multidimensional scaling. Wiley Interdisciplinary Reviews: Computational Statistics 4: 124–167.

60. Faul F, Erdfelder E, Lang AG, Buchner A (2007) G*Power 3: a exible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods 39: 175–191.

61. Faul F, Erdfelder E, Buchner A, Lang AG (2009) Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behavior Research Methods 41: 1149–60.

62. Hwang D, Schmitt WA, Stephanopoulos G, Stephanopoulos G (2002) Determination of minimum sample size and discriminatory expression patterns in microarray data. Bioinfomatics 18: 1184–1193.

63. Chi YY (2011) Multivariate methods. Wiley Interdisciplinary Reviews Computational Statistis 4: 35–47.

64. Efron B, Gong G (1983) A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician 37: 36–48.

65. Efron B, Tibshirani R (1993) An Introduction to the Bootstrap. Boca Raton, FL: Chapman & Hall/CRC.

66. Lebart L (2006) Explorer l'espace des mots: du linéaire au non-linéaire [Exploring word-space: from the linear to the non-linear]. Journées internationales d'analyse statistique des données textuelles 8.

67. Efron B (1979) Bootstrap methods: Another look at the Jackknife. The Annals of Statistics 7: 1–26.

68. Abdi H, Williams LJ (2010) Jackknife. In: Salkind NJ, editor, Encyclopedia of Research Design, Thousand Oaks, CA: Sage. pp 655–660.

69. Anstey KJ, Low LF (2004) Normal cognitive changes in aging. Australian Family Physician 33: 783–7.

70. Tucker-Drob EM, Salthouse TA (2008) Adult age trends in the relations among cognitive abilities. Psychology and Aging 23: 453–60.

71. Salthouse TA, Ferrer-Caja E (2003) What needs to be explained to account for age-related effects on multiple cognitive variables? Psychology and Aging 18: 91–110.

72. Salthouse TA (2001) Attempted decomposition of age-related inuences on two tests of reasoning. Psychology and Aging 16: 251–263.

73. Salthouse TA, Fristoe N, McGuthry KE, Hambrick DZ (1998) Relation of task switching to speed, age, and uid intelligence. Psychology and Aging 13: 445–461.

74. Lindenberger U, Baltes PB (1997) Intellectual functioning in old and very old age: cross-sectional results from the Berlin Aging Study. Psychology and Aging 12: 410–432.

75. Lindenberger U, Mayr U, Kliegl R (1993) Speed and intelligence in old age. Psychology and Aging 8: 207–220.

76. Allaire JC, Marsiske M (2005) Intraindividual variability may not always indicate vulnerability in elders' cognitive performance. Psychology and Aging 20: 390–401.

77. Bielak AAM, Hultsch DF, Strauss E, Macdonald SWS, Hunter MA (2010) Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. Neuropsychology 24: 731–41.

78. Bielak AAM, Hultsch DF, Strauss E, Macdonald SWS, Hunter MA (2010) Intraindividual Variability Is Related to Cognitive Change in Older Adults : Evidence for Within-Person Coupling. Psychology and Aging 25: 575–586.

79. Brose A, Schmiedek F, Lövdén M, Molenaar PCM, Lindenberger U (2010) Adult age differences in covariation of motivation and working memory performance: contrasting between-person and within-person findings. Research in Human Development 7: 61–78.

80. Christensen H, Dear K, Anstey KJ, Parslow RA, Sachdev P, et al. (2005) Within-occasion intraindividual variability and preclinical diagnostic status: is intraindividual variability an indicator of mild cognitive impairment? Neuropsychology 19: 309–317.

81. Dixon RA, Garrett DD, Lentz TL, MacDonald SWS, Strauss E, et al. (2007) Neurocognitive markers of cognitive impairment: exploring the roles of speed and inconsistency. Neuropsychology 21: 381–399.

82. Hilborn JV, Strauss E, Hultsch DF, Hunter MA (2009) Intraindividual variability across cognitive domains: investigation of dispersion levels and performance profiles in older adults. Journal of Clinical and Experimental Neuropsychology 31: 412–424.

83. Hultsch DF, MacDonald SWS, Dixon RA (2002) Variability in reaction time performance of younger and older adults. The Journals of Gerontology Series B, Psychological Sciences and Social Sciences 57: P101–P115.

84. Papenberg G, Bäckman L, Chicherio C, Nagel IE, Heekeren HR, et al. (2011) Higher intraindividual variability is associated with more forgetting and dedifferentiated memory functions in old age. Neuropsychologia 49: 1879–1888.

85. MacDonald SWS, Hultsch DF, Dixon RA (2003) Performance variability is related to change in cognition: evidence from the Victoria Longitudinal Study. Psychology and Aging 18: 510–523.

86. Jackson JD, Balota DA, Duchek JM, Head D (2012) White matter integrity and reaction time intraindividual variability in healthy aging and early-stage Alzheimer disease. Neuropsychologia 50: 357–366.

87. Li SC, Aggen SH, Nesselroade JR, Baltes PB (2001) Short-term uctuation in elderly people's sensorimotor functioning predict text and spatial memory performance: the MacArthur Successful Aging Studies. Gerontology 47: 100–116.

88. Siegler RS (1994) Cognitive variability: a key to understanding cognitive development. Current Directions in Psychological Science 3: 1–5.

89. Adam S, Bonsang E, Germain S, Perelman S (2007) Retraite, activités non professionnelles et vieillissement cognitif: une exploration à partir des données de SHARE [Retirement, non-professional activities and cognitive aging: an investigation based on SHARE data]. Économie et Statistique 403: 83–96.

90. Bonsang E, Adam S, Perelman S Does retirement affect cognitive functioning? Technical report, Network for Studies on Pensions, Aging and Retirement. URL http://arno.uvt.nl/show.cgi?fid = 113197.

91. Garrett DD, Kovacevic N, McIntosh AR, Grady CL (2011) The importance of being variable. The Journal of Neuroscience 31: 4496–503.

92. Solomon PR, Pendlebury WW (1998) Recognition of Alzheimer's disease: the 7 Minute Screen. Family Medicine 4: 265–271.

93. Solomon PR, Hirschoff A, Kelly B, Relin M, Brush M, et al. (1998) A 7 minute neurocognitive screening battery highly sensitive to Alzheimers disease. Archives of Neurology 55: 349–355.

94. Erber NP (1982) Washington, DC: Alexander Graham Bell Association for the Deaf.

95. Sims DG (1975) The validation of the cid everyday sentence test for use wih the severely hearing impaired. Journal of the Academy of Rehabilitative Audiology 8: 70–79.

96. Carnes P (1988) Proverbia in Fabula: Essays on the Relationship of the Fable and the Proverb. Bern: Verlag Peter Lang.

97. Ulatowska HK, Sadowska M, Kadzielawa D, Kordys J, Rymarczyk K (2000) Linguistic and cognitive aspects of proverb processing in aphasia. Aphasiology 14: 227–250.

98. Dorfman M, Brewer W (1994) Understanding the point of fables. Discourse Processes 17: 105–129.

99. Gaudreault R (1994) Textes narratifs, descriptifs et autres: Une approche sémiotique [Narrative, descriptive, and other texts: a semiotic approach]. Semiotic 99: 297–317.

100. Hanauer DI, Waksman S (2000) The role of explicit moral points in fable reading. Discourse Processes 30: 107–132.

101. Jamet MC (1988) Linguistique textuelle et analyse littéraire [Text linguistics and literary analysis]. Lingue del Mondo 53: 79–82.

102. Smith ME (1915) The fable and kindred forms. Journal of English and Germanic Philology 14: 519–529.

103. Dolby-Stahl SK (1980) Sour grapes: fable, proverb, unripe fruit. In: Burlakoff N, Lindahl C, editors, Folklore on Two Continents: Essays in Honor of Linda Daugh, Bloomington, IN: Trickster Press. pp 160–168.

104. Ulatowska HK, Olness GS, Williams-Hubbard LJ (2005) Macrostructure revisited: An examination of gist responses in aphasia. Brain and Language 35: 109–110.

105. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159–174.

106. Saporta G, Niang N (2006) Correspondence analysis and classification. In: Greenacre M, Blasius J, editors, Multiple Correspondence Analysis and Related Methods, Boca Raton, FL: Chapman & Hall/CRC. pp 371–392.

107. Le Roux B, Rouanet H (2010) Multiple Correspondence Analysis. Thousand Oaks, CA: SAGE.

108. Abdi H, Dunlop JP, Williams LJ (2009) How to compute reliability estimates and display confidence and tolerance intervals for pattern classifiers using the Bootstrap and 3-way multidimensional scaling (DISTATIS). NeuroImage 45: 89–95.

109. Efron B (1987) Better Bootstrap Confidence Intervals. Journal of the American Statistical Association 82: 171–185.

110. Wechsler D (1998) Wechsler Adult Intelligence Scale (3rd ed.). New York: Psychological Press.

111. Wechsler D (1998) Wechsler Memory Scale (3rd ed.). New York: Psychological Press.

112. Reitan R (1958) Validity of the trail making test as an indicator of organic brain damage. Perception and Motor Skills 8: 271–276.

113. Hart RP, Kwentus JA, Wade JB, Taylor JR (1988) Modified Wisconsin Card Sorting Test in elderly normal, depressed and demented patients. Clinical Neuropsychologist 2: 49–56.

Supplementary Information
Effect of Age on the Variability in the Production of
Text-Based Global Inferences

Lynne J. Williams[1,*], Joseph P. Dunlop[2], Hervé Abdi[3,*]

1 Lynne J. Williams Centre for Brain Fitness, Rotman Research Institute, Baycrest, Toronto, ON, CANADA lwilliams@research.baycrest.org

2 Joseph P. Dunlop School of Behavioral and Brain Sciences, The University of Texas at Dallas, MS: GR4.1, 800 West Campbell Road, Richardson, TX USA joseph.dunlop@utdallas.edu

3 Hervé Abdi School of Behavioral and Brain Sciences The University of Texas at Dallas MS: GR4.1 800 West Campbell Road Richardson, TX 75080-3021 USA herve@utdallas.edu

∗ E-mail: lwilliams@research.baycrest.org, herve@utdallas.edu

Fables

S1.1 Boasting Traveller

Once upon a time, a world traveller boasted of performing superhuman feats in the places he had visited. The more he boasted, the more people gathered around to hear his stories. Once a small group of listeners had gathered, he told them, "While in Caracas, in front of a group of people like you, I leaped farther than any man in history. I bet no one can match my jump." Doubtful of the traveller's claims, one of the bystanders interrupted him, saying, "If your feat is true, you don't need any witnesses, just leap for us."

S1.2 Crane and Peacock

Once upon a time, a peacock was admiring his gorgeous tail. Preening, he marvelled at how the colors flashed in the sunlight. Seeing a crane, he made fun of the drabness of the crane's coloring. "I am colored like the robes of kings in gold and purple and all the colors of the rainbow. You do not have any color, so

you are no where near as beautiful or good as I am." "True, I may not be beautiful," replied the crane, "but I can fly among the heavens, while you are limited to the ground."

S1.3   Farmer and Sons

A farmer found out that he was dying. He wanted his children to treat his farm with the same care that he did. He loved his farm and wanted it to thrive. He called his sons to his bedside and said, "My children, you must know that there is a treasure hidden in my vineyards." After the man died, the children dug the ground trying to find this great wealth. They found no treasure, but amazingly enough, the vineyards produced a wonderful amount of fruit.

S1.4   Father and Sticks

A father once had sons who always argued with each other. Frustrated with their fighting, he pleaded with the boys to stop. When his pleading did not work, he was determined to show his sons the value of his advice. One day, the father asked his sons to bring him a bundle of sticks. Once they brought the bundle to him, the father asked his sons to break it. One by one, each son tried to break it, but even using all their strength, no one could. After the sons had failed, the father opened the bundle of sticks and gave one stick to each boy. With only one each, the sons had no trouble breaking the sticks.

S1.5   Fox and Goat

Once upon a time, a fox was walking through a field and slipped into a well. Frightened, the fox tried to climb out but the sides of the well were too steep. Later that day, a goat wandered by looking for some water to drink. Finding the fox, the goat asked if he knew where there was some water. With a sly thought, the fox told the goat of the magical properties of the well's water. The fox told the goat to jump into the well in order to get his drink. Believing the fox, the goat jumped in. As the goat was enjoying his drink, the fox told of his difficulty getting out of the well. Panicked, the goat cried in frustration at being stuck in the well. The fox said, "Do not worry, I have a plan. If you lean against the wall, I will run up your back to escape. And don't worry, I will help you get out after me." Reassured, the goat placed his front hooves against the wall. The fox quickly ran up his back, however, instead of helping the goat he kept on running.

S1.6   Lion and Mouse

Once upon a time, a mouse ran across a sleeping lion's face. Startled, the lion woke up and in his anger intended to kill the mouse. The mouse begged the lion not to harm him and said, "If you do not hurt me, I am sure I can find a way to return your favor." The lion laughed because he did not believe the mouse could ever help him, but he was intrigued and let the mouse go. Sometime later, a group of hunters caught the lion and tied him up with ropes. The lion roared and roared, but no one responded to his cries. When the mouse finally heard the lion's roar, he came to see what was the trouble. Seeing the lion tied up, the mouse chewed through the ropes.

S1.7   The Miser

Once upon a time, a man sold everything that he had and bought a lump of gold. Not wanting to lose his treasure, he buried the gold beside an old wall. He visited the spot every day to make sure that his gold was safe. One day, one of his workmen noticed the man's many visits to the same spot. Curious, the workman went to see what his employer was looking at. Discovering the man's secret, the workman dug up his employer's treasure. Not long after, the man discovered that his gold was gone.

S1.8   Old Woman and Physician

Once upon a time, an old woman became blind and she asked a doctor to restore her vision. They agreed that the doctor would only be paid if the woman regained her sight. Every day, the doctor came and applied an ointment to her eyes. On every visit, the doctor took some of the woman's valuables with him. After many treatments with the ointment, the doctor gave the woman a different medicine that restored her sight. When the woman opened her eyes, she could not see any of her valuable belongings, so she refused to pay. The doctor took the woman to court where she told the judge, "I did promise to pay the doctor when my sight was restored. He claims that I am healed, but I say I am still blind. When I lost my sight, I could see my valuables. Now, although he swears he has cured me, I am not able to see a single thing."

S1.9   Shepherd Boy

Once upon a time there was a young shepherd who was watching a flock of sheep near a village. Being young, he got bored easily so, several times the boy cried, "Wolf! Wolf!" The villagers came running to his aid. When they arrived, the boy would laugh because there was no real wolf and he called them just to pass the time. One day, a wolf attacked. The shepherd boy cried, "Wolf! Wolf! A wolf is eating my sheep!" but none of the villagers came to help him.

S1.10   Sick Lion

Once upon a time, an old lion discovered he could not catch his food by hunting anymore. In order to not go hungry, he decided to get his food by trickery. He pretended that he was very sick. The other animals, feeling terrible for the old lion, went to his cave in order to help him in his recovery. The old lion, hungry, quickly ate his visitors and waited for the next one to come so that he could have another meal. The fox, noticing that many animals were missing, went to visit the lion. She saw all the footprints going into the cave, but saw none coming out. Suspicious, the fox called out to the lion and asked how he was doing. "I am not well," responded the lion, "please come in so I can talk to you." "No, thank-you," said the fox, "I see many footprints going into your cave, but I do not see any coming out."

S1.11   Turtle and Rabbit

Once upon a time, a rabbit made fun of a turtle's short legs and slow pace. In good humor, the turtle teased, "These short legs will beat you in a race any day." Thinking that he was unbeatable, the rabbit accepted the challenge. Together the rabbit and the turtle decided that the fox should design the course. Once the course was ready, the rabbit and the turtle started the race at the same time. The rabbit put on a quick burst of speed, leaving the turtle behind. Not seeing the turtle behind him, the rabbit decided to take a nap. Meanwhile, the turtle pushed ahead with his persistent slow pace. After some time, the rabbit woke up and realized how late it was, so he nervously decided to use his full speed in order to win the race. Just as he was about to cross the finish line, he noticed that the turtle had already crossed the line and was ready to greet the rabbit.

S1.12 Two Roosters

Once upon a time there were two roosters who always fought over the command of the roost. One day, they had a really big fight and one of the roosters finally beat the other into submission. The rooster who lost hid in a quiet corner of the roost. The winner went to the top of a fence and started crowing and flapping his wings in order to boast of his victory. All of a sudden, an eagle swooped down and snatched away the winner. Later, the other rooster came out from his corner. Now, being the only rooster, he ruled the roost.

DICA: Formal Presentation

The goal of discriminant correspondence analysis (DICA) is to predict group membership of observations which are described by nominal variables (or by variables which represent the *amount* of some quantity). DICA is shown schematically in Figure S1.

S1.13 Notations

We have I observations each described by J variables. The values of the variables for the observations are stored in an I by J data matrix denoted $\mathbf{X}$. The observations of $\mathbf{X}$ are partitioned into N *a-priori* groups of interest with I_n being the number of observations of the nth group (and so $\sum_n^N I_n = I$). The elements of $\mathbf{X}$ are assumed to be positive or zeros and we assume that there are no empty rows or columns (i.e., rows or columns with only zero values).

S1.13.1 Notations for the Groups (Rows)

We denote by $\mathbf{Y}$ the I by N design matrix for the groups describing the rows of $\mathbf{X}$: $y_{i,n} = 1$ if row i belongs to group n, $y_{i,n} = 0$ otherwise.

S1.14 Step 1: Compute the group × variable matrix

The first step of DICA is to compute the N by J matrix of the total of each group. This matrix is called $\mathbf{S}$ and it is computed as

$$\mathbf{S} = \mathbf{Y}^\mathsf{T}\mathbf{X} \ . \tag{1}$$

The grand total of $\mathbf{S}$ is denoted s_{++} (i.e., $s_{++} = \mathbf{1}^{\mathsf{T}}\mathbf{S1}$). From matrix $\mathbf{S}$ we compute a matrix of barycentric row *profiles* denoted $\mathbf{R}^*$ and computed as

$$\mathbf{R}^* = \mathsf{diag}\{\mathbf{S1}\}^{-1}\,\mathbf{S} \tag{2}$$

where the diag operator transforms a vector into a diagonal matrix when applied to a vector and extract the vector of the diagonal elements when applied to a matrix. A row of $\mathbf{R}^*$ is a profile because it is made of non-negative numbers whose sum is equal to one. When transformed into profiles, two rows can be compared independently of their overall level. The *masses* of the barycenters are proportional to the sum of the corresponding groups. Specifically, the N by 1 group mass vector denoted $\mathbf{b}$ is computed as

$$\mathbf{b} = \mathbf{S1} \times s_{++}^{-1} \tag{3}$$

The diagonal barycenter mass matrix is obtained from the barycenter mass vector as

$$\mathbf{B} = \mathsf{diag}\{\mathbf{b}\}\ . \tag{4}$$

The "grand barycenter," denoted $\mathbf{c}$, is the overall barycenter of matrix $\mathbf{R}^*$, it is computed as

$$\mathbf{c} = \mathbf{S}^{\mathsf{T}}\mathbf{1} \times s_{++}^{-1}\ . \tag{5}$$

The *weights* of the columns are inversely proportional to their frequency. The weights are stored in a J by 1 vector denoted $\mathbf{w}$ and the corresponding J by J diagonal matrix is denoted $\mathbf{W}$. Specifically, $\mathbf{W}$ and $\mathbf{w}$ are computed as:

$$\mathbf{W} = \mathsf{diag}\{\mathbf{c}\}^{-1} \quad \text{and} \quad \mathbf{w} = \mathsf{diag}\{\mathbf{W}\}\ . \tag{6}$$

S1.15 Step 2: Correspondence Analysis of the $\mathbf{R}^*$ Barycenter Matrix

The $\mathbf{R}^*$ matrix is then analyzed using CA. Specifically, the first step of the analysis is to *center* $\mathbf{R}^*$ in order to create a matrix of centered profiles. This matrix, denoted $\mathbf{R}$, is computed as

$$\mathbf{R} = \mathbf{R}^* - \mathbf{1c}\ , \tag{7}$$

(with $\mathbf{1}$ being a N by 1 vector of 1s). Then, the matrix $\mathbf{R}$ is analyzed with a generalized singular value decomposition under the constraints provided by the matrices $\mathbf{B}$ (masses for the N groups) and $\mathbf{W}$ (weights for the columns) as [1–4]:

$$\mathbf{R} = \mathbf{P}\mathbf{\Delta}\mathbf{Q}^{\mathsf{T}} \quad \text{with} \quad \mathbf{P}^{\mathsf{T}}\mathbf{B}\mathbf{P} = \mathbf{Q}^{\mathsf{T}}\mathbf{W}\mathbf{Q} = \mathbf{I} \ , \tag{8}$$

where $\mathbf{\Delta}$ is the L by L diagonal matrix of the singular values (with L being the number of non-zero singular values), and $\mathbf{P}$ (respectively $\mathbf{Q}$) being the N by L (respectively J by L) matrix of the left (respectively right) generalized singular vectors of $\mathbf{R}$ (the singular vectors are also called *eigenvectors* and the the squared singular values are also called *eigenvalues*; see for details, [5]).

S1.15.1 Row Factor Scores

The N by L matrix of factor scores for the groups is obtained as

$$\mathbf{F} = \mathbf{P}\mathbf{\Delta} = \mathbf{R}\mathbf{W}\mathbf{Q} \ . \tag{9}$$

The variance of the columns of $\mathbf{F}$ is given by the square of the corresponding singular values (i.e., the "eigen-value" denoted λ, these are stored in the diagonal matrix $\mathbf{\Lambda}$). This can be shown by combining Equations 8 and 9 to give:

$$\mathbf{F}^{\mathsf{T}}\mathbf{B}\mathbf{F} = \mathbf{\Delta}\mathbf{P}^{\mathsf{T}}\mathbf{B}\mathbf{P}\mathbf{\Delta} = \mathbf{\Delta}^2 = \mathbf{\Lambda} \ . \tag{10}$$

S1.15.2 Column Factor Scores (Loadings)

In correspondence analysis, the roles of the row and the columns are symmetrical: They can be represented in the same map because they have the same variance. Therefore, the columns are described by *factor scores* which can also be interpreted as loadings. Column factor scores are used to identify the variables important for the separation between the groups. In DICA, the column factor scores are computed as (*cf.* Equations 8 and and 10):

$$\mathbf{G} = \mathbf{W}\mathbf{Q}\mathbf{\Delta} \ . \tag{11}$$

S1.16   Projection of the Observations in the Discriminant Space

The I rows of matrix $\mathbf{X}$ can be projected (as "supplementary" or "illustrative" elements) onto the space defined by the factor scores of the barycenters. The first step is to transform $\mathbf{X}$ into a matrix of centered row profiles called $\mathbf{L}$ and computed as:

$$\mathbf{L} = \left(\mathsf{diag}\{\mathbf{X1}\}^{-1}\mathbf{X} \right) - \mathbf{1c}, \tag{12}$$

(with $\mathbf{1}$ being an I by 1 vector of 1s). Then from Equations 8 and 9, we find that matrix $\mathbf{WQ}$ is a projection matrix. Therefore, the I by L matrix $\mathbf{H}$ of the factor scores for the rows of $\mathbf{X}$ can be computed as

$$\mathbf{H} = \mathbf{LWQ} = \mathbf{LG\Delta}^{-1}. \tag{13}$$

These projections are barycentric, because the weighted average of the factor scores of the rows of a group gives the factor scores of the group. Specifically, if we define $\mathbf{M}$ as the mass matrix for the observations as

$$\mathbf{M} = \mathsf{diag}\{\mathbf{m}\} = \mathsf{diag}\big\{\mathbf{X1} \times s_{++}^{-1}\big\} \ . \tag{14}$$

Note that the factor scores of the barycenters are the barycenter of the factor scores of the projections of the observations. This is shown by first computing the barycenters of the row factor scores as ($cf.$ Equation 2) as

$$\overline{\mathbf{H}} = \mathsf{diag}\{\mathbf{YM1}\}^{-1}\mathbf{YMH} \ , \tag{15}$$

then plugging in Equation 13 and developing. Taking this into account, Equation 9 gives

$$\overline{\mathbf{H}} = \mathsf{diag}\{\mathbf{YM1}\}^{-1}\mathbf{YMXWQ} = \mathbf{RWQ} = \mathbf{F} \ . \tag{16}$$

S1.17   Quality of the Prediction

The performance, or quality, of the prediction of a discriminant analysis is assessed by predicting the group membership of the observations and by comparing the predicted with the actual group membership. The pattern of correct and incorrect classifications can be stored in a *confusion* matrix in which the columns represent the actual groups and the row the predicted groups. At the intersection of a row and a column

is the number of observations from the column group assigned to the row group.

The performance of the model can be assessed for the observations used to compute the groups: this is the *fixed effect* model. In addition, the performance of the model can be estimated for *new* observations (i.e., observations not used to compute the model): this is the *random effect* model).

S1.17.1 Fixed Effect: Old Observations

The *fixed effect* model predicts the group assignment for the observations used to compute the barycenters of the groups. In order to assign an observation to a group, the first step is to compute the distance between this observation and all N groups. Then, the observation is assigned to the closest group. Several possible distances can be chosen, but a natural choice is the Euclidean distance computed in the factor space [6]. If we denote by $\mathbf{h}_i$ the vector of factor scores for the ith observation, and by $\mathbf{f}_n$ the vector of factor scores for the nth group, then the squared Euclidean distance (in the factor space) between the ith observation and the nth group is computed as

$$d^2\left(\mathbf{h}_i, \mathbf{f}_n\right) = \left(\mathbf{h}_i - \mathbf{f}_n\right)^{\mathsf{T}}\left(\mathbf{h}_i - \mathbf{f}_n\right) \ . \tag{17}$$

(Note that the Euclidean distance in the factor space is equivalent to the so called "chi-squared" distance in the original space). Obviously, other distances are possible (e.g., Mahalanobis distance, see for more details [6]), but the Euclidean distance has the advantage of being "directly read" on the map.

S1.17.1.1 Tolerance intervals The quality of the group assignment of the actual observations can be displayed using *tolerance* intervals. A tolerance interval encompasses a given proportion of a sample or a population. When displayed in two dimensions, these intervals have the shape of an ellipse and are called *tolerance ellipsoids*. For DICA, a group tolerance ellipsoid is plotted on the group factor score map. This ellipsoid is obtained by fitting an ellipse which includes a given percentage (e.g., 95%) of the observations. Tolerance ellipsoids are centered on their groups and the overlap of the tolerance ellipsoids of two groups reflects the proportion of misclassifications between these two groups.

S1.17.2 Random Effect: New Observations

The *random effect* model evaluates the quality of the assignment of *new* observations to groups. This estimation is obtained, in general, by using cross validation techniques that partition the data into a

learning set (used to create the model) and a *testing set* (used to evaluate the model). For DICA we use a variation of this approach called the jackknife (*a.k.a.* "leave one out") approach: Each observation is taken out from the data set, in turn, and then is projected onto the barycenter factor space computed from the remaining observations. This projection is then used to predict its group membership from the distances between the projected observation and the barycenters. In DICA the only pre-processing needed to project an observation consists into the transformation of this observation into a profile. This transformation does not require estimating parameters from the learning set, and this guarantees that the prediction of the left-out observation is random.

The assignment of an observation to a group follows the same procedure as for a fixed effect model: the observation is projected onto the group factor scores, and the observation is assigned to the closest group. Specifically, we denote $\boldsymbol{\ell}_i$ the profile vector for the ith observation, and the following matrices obtained without the ith observation are denoted (1) $\mathbf{X}_{-i}$, (2) $\mathbf{R}_{-i}$, (3) $\mathbf{B}_{-i}$ and (4) $\mathbf{W}_{-i}$ and refer to (1) the $I-1$ by J data matrix (2) the N by J barycenter matrix, (3) the N by N mass matrix, and (4) the J by J weight matrix. All these matrices are obtained using $I-1$ instead of I observations. Then the generalized eigendecomposition of $\mathbf{R}_{-i}$ is obtained as (*cf.* Equation 8):

$$\mathbf{R}_{-i} = \mathbf{P}_{-i}\boldsymbol{\Delta}_{-i}\mathbf{Q}_{-i}^{\mathsf{T}} \quad \text{with} \quad \mathbf{P}_{-i}^{\mathsf{T}}\mathbf{W}_{-i}\mathbf{P}_{-i} = \mathbf{Q}_{-i}^{\mathsf{T}}\mathbf{B}_{-i}\mathbf{Q}_{-i} = \mathbf{I} \tag{18}$$

The matrices of row and column factor scores denoted $\mathbf{F}_{-i}$ and $\mathbf{G}_{-i}$ are obtained as (*cf.* Equations 9 and 18)

$$\mathbf{F}_{-i} = \mathbf{P}_{-i}\boldsymbol{\Delta}_{-i} = \mathbf{R}_{-i}\mathbf{W}_{-i}\mathbf{Q}_{-i} \quad \text{and} \quad \mathbf{G}_{-i} = \mathbf{W}_{-i}\mathbf{Q}_{-i}\boldsymbol{\Delta}_{-i} . \tag{19}$$

The jackknifed projection of the ith observation, denoted $\widetilde{\mathbf{h}}_i$ is obtained (*cf.* Equation 13) as

$$\widetilde{\mathbf{h}}_i = \boldsymbol{\ell}_i\mathbf{W}_{-i}\mathbf{Q}_{-i} = \boldsymbol{\ell}_i\mathbf{G}_{-i}\boldsymbol{\Delta}_{-i}^{-1} . \tag{20}$$

Distances between the ith observation and the N groups can be computed (*cf.* Equation 17) with the factor scores. The observation is then assigned to the closest group. Note that the jackknife procedure assumes that there are no columns with only one non-zero entry. If there is such a column, we would create a "division by zero error" when the non-zero observation is jackknifed.

S1.17.2.1 Prediction intervals In order to display the quality of the prediction for *new* observations, we use *prediction* intervals. In order to compute these intervals, the first step is to project the jackknifed observations onto the original complete factor space. There are several ways to project a jackknifed observation onto the factor score space. Here we proposed a two-step procedure. First, the observation is projected onto the jackknifed space and is reconstructed from its projections. Then, the reconstituted observation is projected onto the full factor score solution. Specifically, a jackknifed observation is reconstituted from its factor scores as (*cf.* Equations 8 and 20):

$$\widetilde{\boldsymbol{\ell}}_i = \widetilde{\mathbf{h}}_i \mathbf{Q}_{-i}^{\mathsf{T}} \ . \tag{21}$$

The projection of the jackknifed observation is denoted $\widehat{\mathbf{h}}_i$ and is obtained by projecting $\widetilde{\boldsymbol{\ell}}_i$ as a supplementary element in the original solution. Specifically, $\widehat{\mathbf{h}}_i$ is computed as:

$$
\begin{aligned}
\widehat{\mathbf{h}}_i &= \widetilde{\boldsymbol{\ell}}_i \mathbf{W} \mathbf{Q} && (\textit{cf. } \text{Equation 9}) \\
&= \widetilde{\mathbf{h}}_i \mathbf{Q}_{-i}^{\mathsf{T}} \mathbf{W} \mathbf{Q} && (\textit{cf. } \text{Equation 21}) \\
&= \boldsymbol{\ell}_i \mathbf{W}_{-i} \mathbf{Q}_{-i} \mathbf{Q}_{-i}^{\mathsf{T}} \mathbf{W} \mathbf{Q} && (\textit{cf. } \text{Equation 20}) \ .
\end{aligned} \tag{22}
$$

Note that $\widehat{\mathbf{h}}_i$ can also be computed from the column factor scores as

$$\widehat{\mathbf{h}}_i = \boldsymbol{\ell}_i \mathbf{G}_{-i} \boldsymbol{\Delta}_{-i}^{-2} \mathbf{G}_{-i}^{\mathsf{T}} \mathbf{W}_{-i} \mathbf{G} \boldsymbol{\Delta}^{-1} \ . \tag{23}$$

The quality of the predicted group assignment of the observations as a random model can be displayed using *prediction* intervals. A prediction interval encompasses a given proportion of the predicted elements of a sample or a population. When displayed in two dimensions, these intervals have the shape of an ellipse and are called *prediction ellipsoids*. For DICA, a group prediction ellipsoid is plotted on the group factor score map. This ellipsoid is obtained by fitting an ellipse which includes a given percentage (e.g., 95%) of the predicted observations. Prediction ellipsoids are not necessarily centered on their groups, in fact the distance between the center of the ellipse and the group represents the estimation *bias*. Overlap of two predictions intervals directly reflects the proportion of misclassifications for the "new" observations.

S1.18   Quality of the Group Separation

S1.18.1 R^2 and Permutation Test

In order to evaluate the quality of the discriminant model, we use a coefficient inspired by the coefficient of correlation. Because DICA is a barycentric technique, the total variance (i.e., the *inertia*) of the observations to the grand barycenter (i.e., the barycenter of all groups) can be decomposed into two additive quantities: (1) the inertia of the observations relative to the barycenter of their own category, and (2) the inertia of the group barycenters to the grand barycenter.

Specifically, if we denote by $\bar{\mathbf{f}}$ the vector of the coordinates of the grand barycenter (i.e., each component of this vector is the average of the corresponding components of the barycenters), the total inertia, denoted $\mathcal{I}_{\mathsf{Total}}$, is computed as the sum of the squared distances of the observations to the grand barycenter (*cf.* Equation 17):

$$\mathcal{I}_{\mathsf{Total}} = \sum_i^I m_i d^2 \left(\mathbf{h}_i, \bar{\mathbf{f}}\right) = \sum_i^I m_i \left(\mathbf{h}_i - \bar{\mathbf{f}}\right)^\top \left(\mathbf{h}_i - \bar{\mathbf{f}}\right) \ . \tag{24}$$

In CA, the grand barycenter is the center of the space, and therefore $\bar{\mathbf{f}} = \mathbf{0}$ and Equation 24 reduces to

$$\mathcal{I}_{\mathsf{Total}} = \sum_i^I m_i \mathbf{h}_i^\top \mathbf{h}_i \ . \tag{25}$$

The inertia of the observations relative to the barycenter of their own category is abbreviated as the "inertia within." It is denoted $\mathcal{I}_{\mathsf{Within}}$ and computed as

$$\mathcal{I}_{\mathsf{Within}} = \sum_n^N \sum_{i \text{ in group } n} m_i d^2 \left(\mathbf{h}_i, \mathbf{f}_n\right) = \sum_n^I \sum_{i \text{ in group } n} m_i \left(\mathbf{h}_i - \mathbf{f}_n\right)^\top \left(\mathbf{h}_i - \mathbf{f}_n\right) \ . \tag{26}$$

The inertia of the barycenters to the grand barycenter is abbreviated as the "inertia between." It is denoted $\mathcal{I}_{\mathsf{Between}}$ and computed as

$$\mathcal{I}_{\mathsf{Between}} = \sum_i^I b_n \times d^2 \left(\mathbf{f}_n, \bar{\mathbf{f}}\right) = \sum_n^N b_n \times d^2 \left(\mathbf{f}_n, \bar{\mathbf{f}}\right) = \sum_n^N b_n \times \left(\mathbf{f}_n - \bar{\mathbf{f}}\right)^\top \left(\mathbf{f}_n - \bar{\mathbf{f}}\right) = \sum_n^N b_n \times \mathbf{f}_n^\top \mathbf{f}_n \ . \tag{27}$$

So the additive decomposition of the inertia can be expressed as

$$\mathcal{I}_{\mathsf{Total}} = \mathcal{I}_{\mathsf{Within}} + \mathcal{I}_{\mathsf{Between}} \ . \tag{28}$$

This decomposition is similar to the familiar decomposition of the sum of squares in the analysis of variance. This suggest that the intensity of the discriminant model can be tested by the ratio of between inertia by the total inertia, as is done in analysis of variance and regression. This ratio is denoted R^2 and it is computed as:

$$R^2 = \frac{\mathcal{I}_{\mathsf{Between}}}{\mathcal{I}_{\mathsf{Total}}} = \frac{\mathcal{I}_{\mathsf{Between}}}{\mathcal{I}_{\mathsf{Between}} + \mathcal{I}_{\mathsf{Within}}} \ . \tag{29}$$

The R^2 ratio takes values between 0 and 1, the closer to one the better the model. The significance of R^2 can be assessed by permutation tests, and confidence intervals can be computed using cross-validation techniques such as the jackknife (see [7]) or the bootstrap (see [8]).

S1.18.2 Confidence Intervals

The stability of the position of the groups can be displayed using *confidence* intervals. A confidence interval reflects the variability of a population *parameter* or its estimate. In two dimensions, this interval becomes a confidence ellipsoid. The problem of estimating the variability of the position of the groups cannot, in general, be solved analytically and cross-validation techniques need to be used. Specifically, the variability of the position of the groups is estimated by generating *bootstraped* samples from the sample of observations. A bootstraped sample is obtained by sampling *with replacement* from the observations (recall that when sampling with replacement some observations may be absent and some other maybe repeated). The "bootstraped barycenters" obtained from these samples are then projected onto the discriminant factor space and, finally, an ellipse is plotted such that it comprises a given percentage (e.g., 95%) of these bootstraped barycenters. When the confidence intervals of two groups do not overlap, these two groups are "significantly different" at the corresponding alpha level (e.g., $\alpha = .05$). In DICA, the bootstrap can be performed directly in the factor space by sampling the elements of matrix **H** and projecting their weighted means onto the factor space.

S1.19 Where to find more information about DICA

More detailed information about DICA and related methods can be downloaded from the third author's website at www.utdallas.edu/∼herve.

References

1. Abdi H (2007) Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In: Salkind NJ, editor, Encyclopedia of Measurement and Statistics, Thousand Oaks, CA: Sage. pp. 907–912.

2. Abdi H, Williams LJ (2010) Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics 2: 433–459.

3. Greenacre M (1984) Theory and Applications of Correspondence Analysis. London: Academic Press.

4. Williams LJ, Abdi H, French R, Orange JB (2010) A tutorial on multi-block discriminant correspondence analysis (MUDICA): A new method for analyzing discourse data from clinical populations. Journal of Speech Language and Hearing Research 53: 1372–1393.

5. Abdi H (2007) The eigen-decomposition: eigenvalues and eigenvectors. In: Salkind NJ, editor, Encyclopedia of Measurement and Statistics, Thousand Oaks, CA: Sage. pp. 304–308.

6. Abdi H (2007) Distance. In: Salkind NJ, editor, Encyclopedia of Measurement and Statistics, Thousand Oaks, CA: Sage. pp. 280–284.

7. Abdi H, Williams LJ (2010) Jackknife. In: Salkind NJ, editor, Encyclopedia of Research Design, Thousand Oaks, CA: Sage. pp. 655–660.

8. Efron B, Tibshirani R (1993) An Introduction to the Bootstrap. Boca Raton, FL: Chapman & Hall/CRC.
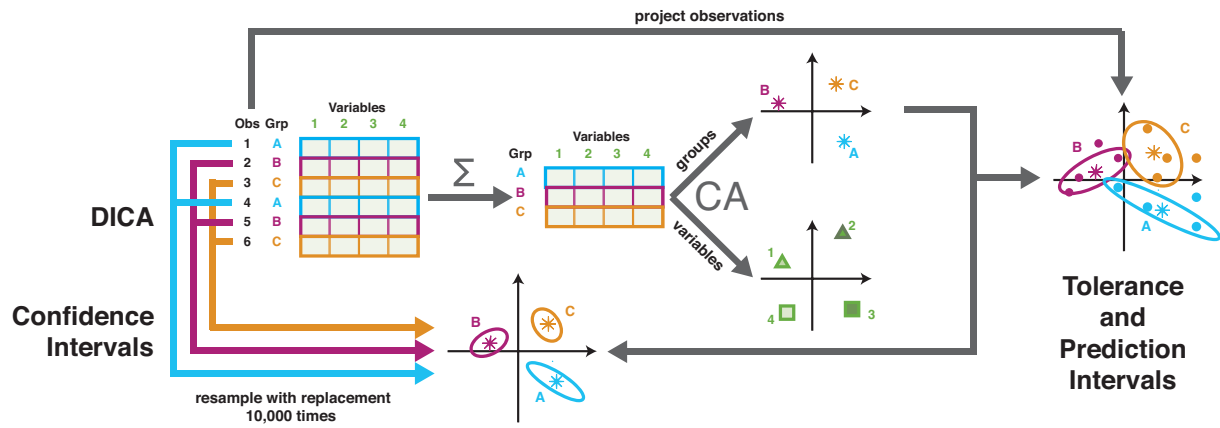
S2   Figure Captions

Figure S1. Schematic diagram of discriminant correspondence analysis (DICA)