REPLY

All Data Collection and Analysis Methods Have Limitations: Reply to Rabbitt (2011) and Raz and Lindenberger (2011)

Timothy A. Salthouse
University of Virginia

The commentaries on my article contain a number of points with which I disagree but also several with which I agree. For example, I continue to believe that the existence of many cases in which between-person variability does not increase with age indicates that greater variance with increased age is not inevitable among healthy individuals up to about 80 years of age. I also do not believe that problems of causal inferences from correlational information are more severe in the cognitive neuroscience of aging than in other research areas; I contend instead that neglect of these problems has led to confusion about neurobiological underpinnings of cognitive aging. I agree that researchers need to be cautious in extrapolating from cross-sectional to longitudinal relations, but I also note that even longitudinal data are limited with respect to their ability to support causal inferences.

Keywords: cross-sectional, longitudinal, mediation, moderation

I appreciate the thoughtful remarks of Rabbitt (2011) and Raz and Lindenberger (2011), and I welcome the opportunity to respond to them. Because of space limitations, I am not able to address all of the issues raised in the commentaries, but in the following I have attempted to respond to what appear to be the most important substantive issues on which we may disagree. The initial section of my target article (Salthouse, 2011b) used results from two recent projects to illustrate major characteristics of cognitive aging. In both projects there were nearly linear age trends in the cross-sectional means and longitudinal changes in measures of cognitive functioning, but in each case the between-person variance was approximately constant across most of adulthood. On the basis of these findings and others cited in the article, I stated that "age-related differences in mean performance can occur without concomitant increases in between-person variability" (p. 754). The relation of age to variability was not a major focus in the article, but it is relevant to the interpretation of moderation analyses because, to the extent that variance systematically increases with increased age in healthy adults, stronger brain–cognition relations at older ages might be a statistical consequence of greater variance rather than a reflection of the emergence of new or stronger brain–cognition relations.

The commentators objected to my suggestion that variability does not inevitably increase with age (Rabbitt, 2011; Raz & Lindenberger, 2011). I agree that an expectation of greater variance at older ages is intuitively plausible and that this finding has been reported in a number of studies. However, because there are cases in which mean age-related declines occur without apprecia-

ble age-related increases in between-person variance, I suggested that the relation was not inevitable. In addition to the results from the two projects portrayed in the figures, several additional studies were cited, including cross-sectional data from the nationally representative samples used to establish the norms for various cognitive test batteries (described in Salthouse, 2010) and longitudinal data from several different studies. In all of the cited studies, negative age trends were reported in measures of level or change in cognitive functioning, with small to nonexistent age-related increase in variance. A similar pattern is evident in scatter plots of various brain structure variables, as several studies cited in Salthouse (2011b) contained figures indicating nearly linear negative age trends in the average value, with little or no increase in variance in samples of healthy adults up to about 80 years of age. Among these are studies of total or regional brain volume (e.g., Abe et al., 2008; Allen, Bruss, Brown, & Damasio, 2005; DeCarli et al., 2005; Fotenos, Mintun, Snyder, Morris, & Buckner, 2008; Fotenos, Snyder, Girton, Morris, & Buckner, 2005; Good et al., 2001; Kennedy et al., 2009; Lemaitre et al., 2010; Sowell et al., 2003; Walhovd et al., 2009; Zimmerman et al., 2006), studies of cortical thickness (e.g., Ecker et al., 2009; Lemaitre et al., 2010; Salat et al., 2004), and studies of white matter integrity based on diffusion tensor imaging (e.g., Abe et al., 2008; Charlton et al., 2006; Grieve, Williams, Paul, Clark, & Gordon, 2007; Hsu et al., 2008; Michielse et al., 2010; Rovaris et al., 2003; Salat et al., 2005; Stadlbauer, Salmonowitz, Strunk, Hammen, & Ganslandt, 2008; Sullivan & Pfefferbaum, 2006; Voineskos et al., 2010; Westlye et al., 2010). The research findings regarding age-related increases in variability are clearly mixed, perhaps in part because of the inclusion of individuals in early stages of dementia or in the period of terminal decline in some studies, but the existence of studies such as those cited above is consistent with the claim that age-related decreases in the means can occur without inevitable age-related

increases in variance in samples of relatively healthy adults between about 20 and 80 years of age.

Raz and Lindenberger (2011) noted that the discussion in Salthouse (2011b) emphasized linear age relations whereas age relations are sometimes nonlinear. I completely agree that nonlinear age relations can occur and that in some respects they are more interesting than linear trends. Unfortunately, nonlinear trends may not be detected when, as in many of the studies reviewed, the age range of the research participants is restricted or only extreme groups of young and old adults are compared. Nevertheless, nonlinear age relations could be accommodated in most analytical methods, and none of the substantive points in the article depend on the specific form of the age relations.

The commentators had markedly different views about the usefulness of mediation-like procedures with cross-sectional data. Rabbitt (2011) noted several examples from his own work in which they were informative, as when control of the variance in measures of brain volume reduced the relation between measures of balance and of cognition. In contrast, Raz and Lindenberger (2011) argued strongly against mediation analyses of cross-sectional data, particularly when they are used to make inferences about longitudinal changes. I agree that cross-sectional mediation analyses can be misleading if they are used as a proxy for longitudinal relations, and many researchers have clearly interpreted them in this fashion. However, it is worth considering whether the problem is the analytical methods or the inferences about developmental phenomena based upon results of the methods.

Because the term *mediation* seems to imply change over time, in that if *X* mediates the *Y–Z* relation then *X* is often assumed to occur intermediate in the temporal sequence between *Y* and *Z* (Kraemer, Kiernan, Exxex, & Kupfer, 2008), mediation analyses can invite inferences about temporal relations that may not be justified. However, mediation techniques can also be viewed as a type of variance partitioning, similar to other statistical control methods such as various forms of multiple regression and partial correlation, and these procedures can be useful even if none of the variables involve a temporal dimension. That is, a discovery that the relation between *A* and *C* is reduced when variance in *B* is controlled but that there is little change in the *A–B* relation when variance in *C* is controlled or in the *B–C* relation when variance in *A* is controlled is likely to be informative even if all variables are static and have no temporal connotation. As an illustration, consider the interpretations of the preceding pattern if *A* referred to attention-deficit/hyperactivity disorder (ADHD), *B* referred to behavior problems, and *C* referred to cooperation with peers. Rather than merely indicating that ADHD was associated with low levels of peer cooperation, a pattern such as that outlined above would be consistent with an interpretation that behavior problems may be largely responsible for the relation between ADHD and peer cooperation. Whether information of this type is relevant to the ultimate causes of the phenomenon could be debated, but it can often serve to distinguish among alternative interpretations of relations observed at a particular point in time. Just as we should resist inappropriate extrapolation of results beyond the situations to which they apply, we should also resist universal rejection of analytical procedures that can be informative when their limitations are recognized.

I very much agree with the general point that researchers need to be cautious when making inferences about longitudinal changes

from cross-sectional differences. The distinction between difference and change is often ignored, even in the titles of articles in which the term *change* appears and all of the data are based on cross-sectional differences. It is therefore worth emphasizing that whether comparable patterns are evident in cross-sectional and longitudinal data is an empirical question and not a logical necessity. In fact, two recent projects from my laboratory revealed one situation (with the Connections variable) in which speed and fluid cognitive abilities had similar patterns of correlations with both cross-sectional differences and longitudinal changes (Salthouse, 2011a), and another situation (with the Mini-Mental State Examination variable) in which reasoning and memory abilities had somewhat different patterns of correlations with cross-sectional differences and with longitudinal changes (Soubelet & Salthouse, 2011).

Raz and Lindenberger (2011) suggested that multivariate longitudinal studies and experimental interventions are better alternatives than cross-sectional studies for reaching causal inferences. Although longitudinal data clearly contain more information than cross-sectional data because temporal information about change is included in addition to the cross-sectional information about differences, it is important to consider exactly how longitudinal data could be informative about causes. As I noted in the article, relations between the early change in one variable and the later change in another variable (i.e., lead-lag relations) can be more informative than simple correlations among the changes. However, lead-lag relations can be difficult to interpret in terms of causality without considerable knowledge (or strong assumptions) about the timing of critical events, including the onset of the change in the leading variable, the time until the leading change reaches a critical value, and the interval between the critical value in the leading variable and the onset of change in the lagged variable. I am not suggesting that these problems are insurmountable but rather that at the current time we have very little information relevant to the timing of critical events in either brain or cognitive variables. Another limitation of much of the existing research is that in many of the studies all of the participants have been adults over about 60 years of age, and correlations among changes occurring during a period when both variables have been changing for years or decades are likely to be less informative about causal sequences than are correlations obtained during a period when one or both variables are just beginning to change. I agree that longitudinal data are generally preferred over cross-sectional data, but I also believe that it is important to recognize that the mere existence of correlations among changes provides a limited basis for inferring causality. This is so because the relevant data are still observational, relations among changes after a prolonged period of change may not be informative about relations at earlier periods in time when the causal sequences were initiated, and even lagged relations may reflect preexisting characteristics of the individuals rather than direct causal influences.

Although I suspect, for the reasons mentioned above, that there are limits on how much can be learned about causes of aging from only studying older adults, I was puzzled by the implication that I believe that studying young adults would necessarily be informative about brain–cognition relations in aging or that it is "a sufficient basis for understanding the brain basis of cognitive aging" (Raz & Lindenberger, 2011, p. 793). My view is that brain–cognition relations might be fruitfully studied in young

adults only if powerful tests of moderation failed to find different *B–C* relations at different ages. However, as I noted in the article, the existing results on moderation are weak and inconsistent, and therefore very little is currently known about the relation between age and *B–C* relations. It is unfortunate that so few studies with either cross-sectional or longitudinal data have included large samples of adults across a wide age range, because moderation analyses could be very informative about the timing and age-specificity of brain–cognition relations.

In conclusion, I appreciate the time and efforts of the commentators because their remarks have provided an opportunity to clarify my position and highlight differences in perspectives. The major points of the article were that despite much research examining interrelations of aging, measures of brain structure, and measures of cognitive functioning, only tentative conclusions are currently possible regarding the neuroanatomical substrates of age-related cognitive decline. As I stated in the article, all data collection and analytical methods have limitations, and strong causal inferences are exceedingly difficult when the critical variable of age cannot be experimentally manipulated. Because no method is without limitations, I suggested that confidence in conclusions will likely be greatest when the results are found to converge across multiple methods involving different sets of assumptions and different types of data. If only cross-sectional data are available, some types of mediation analyses involving decomposition of covariances will often be more informative than simple correlations about the cross-sectional interrelations among the variables, particularly if alternative interpretations are considered. Longitudinal data are essential for drawing conclusions about changes occurring within the individual, but it is important to recognize that, because they are observational rather than experimental, even longitudinal data are limited with respect to their ability to identify causes. The limitation is probably even greater when all of the observations are obtained at some unknown point after the critical relations originated.

The commentators had quite different perspectives on what researchers in this area should do in the future. Rabbitt (2011) suggested that the only option was to "Keep calm and carry on" (p. 787), whereas Raz and Lindenberger (2011) argued that "no practical reason can justify the continuation of business as usual" (p. 794). My view lies between those of the commentators, as I feel that greater progress can be achieved if researchers recognize potential limitations of all data collection and analysis methods and attempt to base conclusions on results from multiple methods of data collection and analysis whenever possible.

References

Abe, O., Yamasue, H., Aoki, S., Suga, M., Yamada, H., Kasai, K., . . . Ohtomo, K. (2008). Aging in the CNS: Comparison of gray/white matter volume and diffusion tensor data. *Neurobiology of Aging, 29,* 102–116. doi:10.1016/j.neurobiolaging.2006.09.003

Allen, J. S., Bruss, J., Brown, C. K., & Damasio, H. (2005). Normal neuroanatomical variation due to age: The major lobes and a parcellation of the temporal region. *Neurobiology of Aging, 26,* 1245–1260. doi:10.1016/j.neurobiolaging.2005.05.023

Charlton, R. A., Barrick, T. R., McIntyre, D. J., Shen, Y., O'Sullivan, M., Howe, F. A., . . . Markus, H. S. (2006). White matter damage on diffusion tensor imaging correlates with age-related cognitive decline. *Neurology, 66,* 217–222. doi:10.1212/01.wnl.0000194256.15247.83

DeCarli, C., Massaro, J., Harvey, D., Hald, J., Tullberg, M., Au, R., . . . Wolf, P. A. (2005). Measures of brain morphology and infarction in the Framingham Heart Study: Establishing what is normal. *Neurobiology of Aging, 26,* 491–510. doi:10.1016/j.neurobiolaging.2004.05.004

Ecker, C., Stahl, D., Daly, E., Johnston, P., Thomson, A., & Murphy, D. G. M. (2009). Is there a common underlying mechanism for age-related decline in cortical thickness? *NeuroReport, 20,* 1155–1160. doi:10.1097/WNR.0b013e32832ec181

Fotenos, A. F., Mintun, M. A., Snyder, A. Z., Morris, J. C., & Buckner, R. L. (2008). Brain volume decline in aging: Evidence for a relation between socioeconomic status, preclinical Alzheimer disease and reserve. *Archives of Neurology, 65,* 113–120. doi:10.1001/archneurol.2007.27

Fotenos, A. F., Snyder, A. Z., Girton, L. E., Morris, J. C., & Buckner, R. L. (2005). Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD. *Neurology, 64,* 1032–1039. doi:10.1212/01.WNL.0000154530.72969.11

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N. A., Friston, K. J., & Frackowiak, R. S. J. (2001). A voxel-based morphometric study of ageing in 465 normal adult brains. *NeuroImage, 14,* 21–36. doi:10.1006/nimg.2001.0786

Grieve, S. M., Williams, L. M., Paul, R. H., Clark, C. R., & Gordon, E. (2007). Cognitive aging, executive function, and fractional anisotropy: A diffusion tensor MR imaging study. *American Journal of Neuroradiology, 28,* 226–235.

Hsu, J.-L., Leemans, A., Bai, C.-H., Lee, C.-H., Tsai, Y.-F., Chiu, H.-C., & Chen, W.-H. (2008). Gender differences and age-related white matter changes of the human brain: A diffusion tensor imaging study. *NeuroImage, 39,* 566–577. doi:10.1016/j.neuroimage.2007.09.017

Kennedy, K. M., Erickson, K. I., Rodrigue, K. M., Voss, M. W., Colcombe, S. J., Kramer, A. F., . . . Raz, N. (2009). Age-related differences in regional brain volumes: A comparison of optimized voxel-based morphometry to manual volumetry. *Neurobiology of Aging, 30,* 1657–1676. doi:10.1016/j.neurobiolaging.2007.12.020

Kraemer, H. C., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology, 27*(Suppl. 2), S101–S108. doi:10.1037/0278-6133.27.2(Suppl.).S101

Lemaitre, H., Goldman, A. L., Sambataro, F., Verchinski, B. A., Meyer-Lindenberg, A., Weinberger, D. R., & Mattay, V. S. (2010). Normal age-related brain morphometric changes: Nonuniformity across cortical thickness, surface area and gray matter volume? *Neurobiology of Aging.* Advance online publication. doi:10.1016/j.neurobiolaging.2010.07.013

Michielse, S., Coupland, N., Camicioli, R., Carter, R., Seres, P., Sabino, J., & Malykhin, N. (2010). Selective effects of aging on brain white matter microstructure: A diffusion tensor imaging tractography study. *NeuroImage, 52,* 1190–1201. doi:10.1016/j.neuroimage.2010.05.019

Rabbitt, P. (2011). Between-individual variability and interpretation of associations between neurophysiological and behavioral measures in aging populations: Comment on Salthouse (2011). *Psychological Bulletin, 137,* 785–789. doi:10.1037/a0024580

Raz, N., & Lindenberger, U. (2011). Only time will tell: Cross-sectional studies offer no solution to the age–brain–cognition triangle: Comment on Salthouse (2011). *Psychological Bulletin, 137,* 790–795. doi:10.1037/a0024503

Rovaris, M., Iannucci, G., Cercignani, M., Sormani, M. P., de Stefano, N., Gerevini, S., . . . Fillippi, M. (2003). Age-related changes in conventional, magnetization transfer, and diffusion-tensor MR imaging findings: Study with whole-brain tissue histogram analysis. *Radiology, 227,* 731–738. doi:10.1148/radiol.2273020721

Salat, D. H., Buckner, R. L., Snyder, A. Z., Greve, D. N., Desikan, R. S. R., Busa, E., . . . Fischl, B. (2004). Thinning of the cerebral cortex in aging. *Cerebral Cortex, 14,* 721–730. doi:10.1093/cercor/bhh032

Salat, D. H., Tuch, D. S., Greve, D. N., van der Kouwe, A. J. W., Hevelone,

N. D., Zaleta, A. K., . . . Dale, A. M. (2005). Age-related alterations in white matter microstructure measured by diffusion tensor imaging. *Neurobiology of Aging, 26,* 1215–1227. doi:10.1016/j.neurobiolaging.2004.09.017

Salthouse, T. A. (2010). *Major issues in cognitive aging.* New York, NY: Oxford University Press.

Salthouse, T. A. (2011a). Cognitive correlates of cross-sectional differences and longitudinal changes in trail making performance. *Journal of Clinical and Experimental Neuropsychology, 33,* 242–248. doi:10.1080/13803395.2010.509922

Salthouse, T. A. (2011b). Neuroanatomical substrates of age-related cognitive decline. *Psychological Bulletin, 137,* 753–784. doi:10.1037/a0023262.

Soubelet, A., & Salthouse, T. A. (2011). Correlates of level and change in the Mini-Mental State Examination. *Psychological Assessment.* Advance online publication. doi:10.1037/a0023401

Sowell, E. R., Peterson, B. S., Thompson, P. M., Welcome, S. E., Henkenius, A. L., & Toga, A. W. (2003). Mapping cortical change across the human life span. *Nature Neuroscience, 6,* 309–315. doi:10.1038/nn1008

Stadlbauer, A., Salomonowitz, E., Strunk, G., Hammen, T., & Ganslandt, O. (2008). Age-related degradation in the central nervous system: Assessment with diffusion-tensor imaging and quantitative fiber tracking. *Radiology, 247,* 179–188. doi:10.1148/radiol.2471070707

Sullivan, E. V., & Pfefferbaum, A. (2006). Diffusion tensor imaging and aging. *Neuroscience & Biobehavioral Reviews, 30,* 749–761. doi:10.1016/j.neubiorev.2006.06.002

Voineskos, A. N., Rajji, T. K., Lobaugh, N. J., Miranda, D., Shenton, M. E., Kennedy, J. L., . . . Mulsant, B. H. (2010). Age-related decline in white matter tract integrity and cognitive performance: A DTI tractography and structural equation modeling study. *Neurobiology of Aging.* Advance online publication. doi:10.1016/j.neurobiolaging.2010.02.009

Walhovd, K. B., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., . . . Fjell, A. M. (2009). Consistent neuroanatomical age-related volume differences across multiple samples. *Neurobiology of Aging.* Advance online publication. doi:10.1016/j.neurobiolaging.2009.05.013

Westlye, L. T., Walhovd, K. B., Dale, A. M., Bjornerud, A., Due-Tonnessen, P., Engvig, A., . . . Fjell, A. M. (2010). Life-span changes of the human brain white matter: Diffusion tensor imaging (DTI) and volumetry. *Cerebral Cortex, 20,* 2055–2068. doi:10.1093/cercor/bhp280

Zimmerman, M. E., Brickman, A. M., Paul, R. H., Grieve, S. M., Tate, D. F., Gunstad, J., . . . Gordon, E. (2006). The relationship between frontal gray matter volume and cognition varies across the healthy adult lifespan. *American Journal of Geriatric Psychiatry, 14,* 823–833. doi:10.1097/01.JGP.0000238502.40963.ac