Perfect, T. J., & Maylor, E. A. (2000). Rejecting the dull hypothesis: The relation between method and theory in cognitive aging research. In T. J. Perfect & E. A. Maylor (Eds.), *Models of cognitive aging* (pp. 1-18). Oxford: Oxford University Press.

Rejecting the dull hypothesis: The relation between method and theory in cognitive aging research.

Tim J. Perfect[1] and Elizabeth A. Maylor[2]

[1] Department of Psychology, University of Plymouth
[2] Department of Psychology, University of Warwick

Acknowledgements

This writing of this chapter was partially supported by a Medical Research Council grant (G9503572) to the first author.

Introduction

In this chapter, we begin by outlining a basic problem in cognitive aging research, namely that the statistical approach based on rejection of the null hypothesis is not appropriate for theory building. We then discuss two broad approaches that have been used in an attempt to overcome this shortcoming. We end with a discussion of potential ways forward, as outlined in the chapters that follow.

The dull hypothesis

Psychologists spend considerable time and effort attempting to establish the reliability of basic effects. Over the past century psychologists have utilised many sophisticated statistical techniques to determine whether the variability in performance they are measuring is likely to represent a real phenomenon to be explained, or whether it is just chance variation about the population mean. We inculcate this approach to our students, against considerable resistance. Students' intellectual opposition to this idea is entirely understandable: if one stops to consider, it is odd that we base our science on the setting up, and then rejection, of the hypothesis that there will be no effect in our studies. Nonetheless, after a few years formal teaching, most students learn to accept the principle that the first aim of an experiment in psychology is to reject the null hypothesis.

However, when one turns to the study of aging it is not easy to see the utility of the null hypothesis. If one were to ask members of the public whether the average 80 year old and the average 25 year old are the same, then it is doubtful whether many would believe so, whatever the measure chosen (e.g. memory, speed, strength, health, visual acuity). However, that is what the null hypothesis

requires us to assume. The standard approach would be to conduct a study comparing older and younger adults on a certain test and to use our statistical techniques to determine whether the two populations are reliably different. What would the public make of this? They would probably think that testing this null hypothesis was a somewhat pointless exercise, that our ability to reject the null hypothesis was extremely unimpressive, and they may start to ask questions about how taxpayers' money was being spent.

Clearly, the null hypothesis is not an appropriate starting point for studies in aging. In fact if the measures were reliable and valid, it would be more interesting to accept the null hypothesis than to reject it. That is, it would be more unexpected, more counter-intuitive and more unusual in the normative sense to find something that older adults are just as good at as the young (see the chapter by Burke, Mackay & James, this volume, for examples of exactly this). In this context the uninteresting finding is that there is a difference, i.e. that younger people in general can remember more, are quicker, stronger, healthier, have better vision and so on. If the psychology of aging is going to have anything to offer, it must go beyond merely demonstrating that younger adults outperform older adults, i.e we must be able to reject the *dull* hypothesis.

Going beyond the dull hypothesis

Nearly a century of research has allowed us to accumulate a reasonably comprehensive picture of the cognitive change that accompanies increased age. Setting aside caveats about methodology, it is clear that there are reasonably ubiquitous age differences in cognition across the lifespan, across most domains of cognitive activity. However, at the behavioural level it is clear that the magnitude of

age-related change in different across different domains. We do not wish to discuss which domains of intellectual activity show the most age related change in this chapter (but see chapters by Burke et al, and Light, Prull, La Voie & Healy, this volume, for examples), but rather to focus on the relationship between the methods used to support different theoretical positions in cognitive aging.

One approach to the rejection of the dull hypothesis has been to contrast the magnitude of age-related differences in different domains of cognition, or in different processing steps within a cognitive domain. The dull hypothesis is rejected if it can be shown that older adults are particularly poor at some classes of tasks compared to other tasks, or compared to their performance in general. Such an approach is based on analysis of age-group differences between behavioural measures, and utilising the cognitive-neuropsychological model, the aim has been to detect dissociations, or better double dissociations, between age and domain of cognition. Researchers adopting this approach have tended to use the ANOVA approach, with the aim of not merely finding a main effect of Age (the dull hypothesis), but of finding an Age x Task or Age x Condition interaction. Researchers using this approach are essentially testing a localist model of cognitive aging, that carries the assumption that aging consists of a series of local (or modular) deficits that can be isolated with the appropriate experimentation. We discuss the use of this method in greater detail below (see also Verhaeghen, this volume).

The second approach has been to study associations in age-related change across different domains of cognition. The logic of this approach is that it is possible to reject the dull hypothesis if it can be shown that age-related change in one domain of cognition is independent of age-related change in another domain. Thus, the focus here is not on the magnitude of age-related change, but whether

different changes are independent of one another. Thus, under this approach, two different domains of cognition may show the same average change, but if they are independent the degree of age-related change on one test will tell you nothing about the age-related change on the other test. Researchers using this approach generally hold a global (or general) view of cognitive decline with age. The purpose of their analyses is to determine the extent to which cognitive change across many domains can be explained by change in a single (or relatively few) 'primitive' measure, designed to tap a general aspect of cognition, such as neural speed, or working memory capacity (see Salthouse, this volume, for a full discussion of this approach). Thus, this approach restates the dull hypothesis in terms of whether age-related sensitivity on a task is greater, or less, than that expected from a general decline based upon age-related change in the 'primitive'. This is assessed by means of correlational techniques such as partial correlation, multiple regression and path analysis to control for the effects of age-related change in the primitive on the age-task association. This approach is discussed further below.

The use of dissociations to reject the dull hypothesis.

The idea behind this approach is simple: one measures age-differences across two (or more) measures, and attempts to determine whether the age-differences are the same magnitude across these tasks. If not, then one expects to find an Age x Task interaction in an analysis of variance (ANOVA). There are numerous examples that could be used to illustrate the ANOVA based approach to cognitive aging, but only one will be used here. In 1982, Salthouse and Somberg examined age-related change in performance at encoding, storage and output in a Sternberg search paradigm, with the stated aim of determining which stages of

processing are most age-sensitive. They manipulated encoding by having participants study intact or visually degraded stimuli. Storage was manipulated by varying the set size for comparison (one vs four items) and response was manipulated by varying the difficulty of the manual response required (simple button press vs complex use of two keyboards).

In the study participants were presented with digits to be remembered (1 vs 4) for 1.5 seconds, followed after a further 1.5 seconds by a target digit (intact vs degraded). The task was to indicate as quickly as possible whether the target digit had been present in the initial set, and responses were made via button press (simple vs complex). Using response time as the dependent variable, there were main effects of age for all stages of processing (encoding, storage, response) and there were also age x complexity interactions. That is older adults were particularly slowed by having four items to search rather than one, by the use of degraded stimuli, and by the use of a complex response.

The presence of age x complexity interactions appear to offer a clear rejection of the dull hypothesis. This is not what Salthouse and Somberg (1982) conclude however. They argue that the data are entirely consistent with the idea that there is a single monolithic change with increased age. Their argument is based on two facts: firstly that the age x complexity interactions were ubiquitous rather than specific to one processing stage. Second, that in proportionate terms, the complexity effect is equivalent for young and old. The essence of this argument is captured in Figure 1.1.

Figure 1.1 here.

Figure 1.1 illustrates the classic age x task interaction often reported in the literature. There are several important features to this data. First there is not only a main effect of age overall, but there is an effect of age in the baseline task. Second, there is a complexity effect for both groups. In terms of mean performance, these two facts mean that in absolute terms, the age difference is greatest in the more complex task, and hence there is a significant age x task interaction. However, if one takes differences in baseline response time into account (by using logarithms, or proportionate change) then the age x task interactions disappear.

The issue of taking baseline differences into account is central for cognitive aging research, because baseline differences are almost inevitably found. However, there is not a clear agreement as to how baselines should be used, and none of the methods currently used are entirely satisfactory. One common solution is to control for baseline task reaction time by calculating proportional increase due to the complex task. (In the area of dual tasks, this is known as divided attention costs, see Somberg & Salthouse, 1982; Perfect & Rabbitt, 1992 for examples).

However, dividing the increase in response time due to the complex task by the baseline measure carries with it several assumptions which may not be warranted. Foremost of these is that response time is a linear scale. i.e. that a change in RT from 200ms to 300ms is equivalent to a change in RT from 2 seconds to 3 seconds. Verhaeghen discusses this issue in Chapter 3, and makes clear that such an assumption is unjustified. Speed accuracy trade off functions are non-linear, and so division by baselines may produce data that are difficult to interpret.

The second problem is that division by baseline does not produce a pure measure of costs if the behavioural index (response time) is not determined solely by the process under manipulation. This is probably a ubiquitous problem. Consider the

visual degradation factor in Salthouse and Somberg's (1982) experiment, which was assumed to be a manipulation of encoding factors (rather than storage or response factors). In the visually intact condition, response time will be a combination of motor speed and central processing time. If we assume, for the sake of simplicity, that the motor and central processes are independent and additive, then the response time (RT) for younger adults on the baseline task can be described as:

$$RT_{base} = RT_{Motor} + RT_{Central(base)} \quad (1)$$

In a complex version of the task (e.g. visually degraded stimuli, which are assumed not to produce slower motor responses) then, response time will be:

$$RT_{complex} = RT_{Motor} + RT_{Central(complex)} \quad (2)$$

Taking the proportionate costs gives us:

$$PCosts = (RT_{Central(complex)} - RT_{Central(base)}) / (RT_{Motor} + RT_{Central(base)}) \quad (3)$$

Equation 3 tells us that proportionate costs due to a change in visual degradation is in part a function of the motor response latency, a factor which is assumed not to vary across conditions, and so should not influence costs due to an increase in visual complexity. It also tells us that the extent to which true costs are under-estimated for changes in task complexity will be greater where motor speed

becomes a lager proportion of the total response time.

Now let us consider the effects of aging, to see what effect this will have on our ability to draw conclusions about proportional costs with increased age. If we assume that the effect of aging on motor speed is a slowing factor of m, whilst for central processes it is a factor of c, substitution of the slowing factors into Equation 3 to get the costs for older adults gives us

$$Pcosts_{old} = c\ (RT_{Central(complex)} - RT_{Central(base)}) \ /\ (m\ RT_{Motor} + c\ RT_{Central(base)})$$

$$Pcosts_{old} = (RT_{Central(complex)} - RT_{Central(base)}) \ /\ ((m/c)\ RT_{Motor} + RT_{Central(base)}) \qquad (4)$$

Equation 4 tells us that if the only effects of age are multiplicative on the underlying processes then the proportional costs for older adults will approach the proportional costs for younger adults as the ratio m/c approaches unity. As m/c decreases from 1, so we would expect proportional costs for older adults to exceed those of the young, whilst increases in m/c over 1 would produce reduced estimates of proportional costs in older adults. As a first step towards estimating the ratio m/c, we can use the estimates of central and peripheral slowing from the metaanalysis by Cerella (1985). His estimates were that central slowing (c) is more marked than peripheral (motor) slowing (m), with a ratio of approximately 1.1/1.4, or 0.79. Thus, according to this analysis one might expect to find greater estimates of age-related costs in the elderly, *even when the 'true' costs are matched across age groups.* However, the magnitude that this differential effect has on the estimate of change due to complexity will itself be proportional to the relative influence of the motor speed element. If motor speed is a larger proportion of total baseline response time, one would expect greater age-differences in proportional costs. If the motor speed

element is a relatively small proportion of total baseline response time, one might expect the age differences in proportional costs to be negligible.

The foregoing argument rests upon the assumption that the age differences can be expressed as simple linear multiplicative effects. However, this may not be true. If the effects of age are non-linear (e.g. Hale, Lima & Myerson, 1991), then proportional costs become harder to interpret. Likewise, if older adults solve more complex tasks in fundamentally different ways from younger adults, then the assumption of multiplicative effects would be false. Thus if one then wishes to compare younger and older adults on this measure in any sensible way, one needs to ensure that the proportionate measure means the same thing for both groups. This is an assumption that is rarely tested.

A development of the proportionate costs idea is the plotting of older adults response latencies against the response latencies of younger adults. Under this technique, the proportionate costs emerge as the slope of the function which relates younger and older response times across experimental conditions. However, in addition to the slope, there is an intercept term in a linear function, or other parameters in more complex functions (see Hale, Lima & Myerson, 1991 for a summary of a range of functions that have been proposed). Such functions - known as Brinley plots after their originator (Brinley, 1965) - have become very influential in cognitive aging, although their use has not gone unchallenged (Perfect, 1994: Fisk, Fisher & Rogers 1992). We return to the issue of Brinley plots in the section on regression analysis below, since Brinley plots are correlational in nature.

Salthouse (e.g. 1991) has argued in favour of converting older adults performance into z-scores based on the distribution of scores in the younger

population, in order to compare across different tasks. This escapes the assumption of linearity inherent in using proportionate measures. However, it involves another assumption that is not justified; it assumes that all variation about a mean is systematic, ie. that the standard deviation measure for a task is a measure of true variance in that process (Chapman & Chapman, 1973). Since no measure has perfect reliability, this assumption is false, and variance in a test is always in part due to error. Thus, if the average older adult is one standard deviation below the younger average on a simple task and on a complex task, this does not mean that the age effect is necessarily equivalent across the two tasks. If a large proportion of the variance in a task is random (i.e. the reliability is low) it is harder to obtain a group difference. Thus 1 S.D. of age-related change on a task of low reliability is indicative of greater age-sensitivity in the underlying process than 1 S.D. of age-related change on a task of high reliability. This is problematic in aging research when comparing the magnitude of age-effects in tasks that vary widely in their complexity. Simple speeded measures are highly reliable, and so estimates of the magnitude of age-effects on such tests are likely to be accurate. However, more complex tasks, such as memory tasks or tests of executive function have either lower, or unknown, reliability, and hence are probably less likely to show age-sensitivity. For example, Salthouse, (1996b) reported that a battery of speed measures had estimated reliabilities ranging from 0.56 to 0.96 with an unweighted mean of 0.85, whilst tests of short and long term memory had estimated reliabilites between 0.37 and 0.89, with an unweighted mean of 0.62.

Another difficulty with the interpretation of interactions comes with the realisation that no task is process pure. No behavioural measure tests only what it purports to, and nothing else. Thus the reason that tests may or may not be age-

sensitive may not be because of what they are supposed to measure, but because of something else that is being tested at the same time. For example, fluency tasks requiring written responses may show age effects not because of the generational aspect of fluency, but because of the effect of writing speed on performance. Likewise, the Wisconsin Card Sort Test is not merely tapping the ability to change rule based behaviour, but also comprehension of the instructions, memory for prior responses and many other factors. Thus comparing two tasks which are chosen to differ on a specified construct may produce interactions with age for reasons that are nothing to do with the construct itself. Alternatively, two tasks that do tap the same construct may give age interactions either because of what else they tap, or because they differentially load on the construct itself. In the latter case, it is possible then that an interaction may still emerge because of a single factor. Salthouse discusses this issue in greater detail in the next chapter in the present volume.

The use of correlational data to reject the dull hypothesis.

The idea behind the correlational approach is as simple as the ANOVA based approach. If one measures age-change in a range of cognitive abilities, then one can reject the dull hypothesis if correlations between age and cognitive ability in one domain are independent of correlations between age and ability in another domain. The logic of this approach is illustrated in Figure 1.2, which uses the conventional Venn diagram notation to indicate the shared variance between variables. (For a fuller account, see Salthouse, 1994). In this example Age (A), cognitive ability B, and cognitive ability C are all interrelated. The question of interest is the extent to which the A-B association, shown as the region (a+b) in Figure 1.2, is

independent of the A-C association, shown as the region (b+c). The correlational technique used is to measure the A-B association, and then see how much variance in this association is reduced by controlling for C. Salthouse (1994,1996a) has promoted the use of proportional reduction in age-associated variance, calculated as the proportion b / (a+b), and use of this technique has repeatedly shown that controlling for speed measures in this way markedly reduces the age-association with measures of higher order cognitive function, such as learning and memory (Salthouse & Dunlosky, 1997), intelligence (Hertzog,1989) and executive function (Salthouse, Fristoe & Rhee, 1996).

Figure 1.2 about here.

As with the ANOVA approach, we begin with an example before discussing difficulties of interpretation. Baltes and Lindenberger (1997) collected 14 separate psychometric assessments of cognitive ability, from a total of 687 individuals aged between 25 and 103 years of age. These 14 assessments were combined to provide estimates of 5 intellectual factors - Perceptual Speed, Reasoning, Memory, Knowledge and Fluency. In addition they collected data on visual and auditory acuity. As expected there were age-related declines across all the measures. The question of interest for Baltes and Lindenberger was whether the cognitive change could be predicted by the sensory change. That is, if one controls for age-change in sensory acuity, does this result in the age-change in cognition being reduced, or even removed altogether. Rather than discuss each cognitive ability separately, we will illustrate the approach with their data based upon a composite of all 5 factors.

There was an age correlation with the composite measure of r = -0.79. In

addition, there was a non-linear (quadratic) association, such that the rate of intellectual decline increased at the later end of the age span. This quadratic effect correlated with age at r = -0.23, bringing the total age-related variance in the composite measure to 67.4%. However, if age-related change in hearing was controlled (by partialling out scores on the auditory acuity test) then the age-related variance in the cognitive measure was reduced to 7.0%. The equivalent figure for visual acuity was 16.7%, and the effect of controlling for both vision and hearing was to reduce the age-related variance in the composite cognitive score to a mere 3.0%. On the basis of such evidence, Baltes and Lindenberger (1997) conclude:

> ".. a large portion of the mechanisms that drive negative age differences in sensory performance also bring about the aging of complex cognition. This finding has implications regarding the search for "psychological primitives" of negative age differences in cognition.... The very high degree of commonality between the age-related variance of the two domains is consistent with the notion that at least a major portion of these primitives is operating at a relatively global, rather than modular or domain-specific, level." (p.20).

Thus these authors are clear in indicating that they subscribe to the view that a single factor can explain the majority of age-related variance. That is that age-related change in one domain is not independent of age-related change in another domain, and so one must conclude that the age-related change observed across domains does not represent modular, independent change, but is probably the result of a single process operating throughout the central nervous system. Baltes

and Lindenberger (1997) do not specify what they believe this 'primitive' to be, but clearly the implication is that age related change in hearing loss and reasoning (for example) are highly inter-related, then the most plausible conclusion is that some single biological process underpins them both.

Other authors have gone further and proposed specific primitives that might explain age related change. Salthouse (1996a, and this volume) is most associated with the view that age-related differences in neural speed can explain most of the age-related variance in cognition. Others have proposed alternative primitives, such as processing resources (e.g. Fastenau, Denburg & Abeles, 1996), working memory capacity (Kirasic, Allen, Dobson & Binder, 1996), loss of connections in a network (Cerella, 1990), or information loss in transmission through a network (Myerson, Hale, Wagstaff, Poon & Smith, 1990). It is not within the scope of the present chapter to review the relative merits of these different theoretical claims. Rather, we will focus on the methodologies used to support these positions. The accounts based on speed, attentional resources and working memory capacity have all been based on the regression technique as used by Baltes and Lindenberger, described above. The theoretical accounts based upon properties of networks (cell loss and information loss) are principally based upon analyses of Brinley functions. We discuss each of these in turn.

Problems with regression analyses.

It may sound trite, but one should not forget the old chestnut that correlation does not imply causation. The regression based techniques, including hierarchical regression and path analytic techniques, used to support single factor accounts of aging are based on correlations in cross-sectional data. Finding an association

between age related variance in a 'primitive' and in a more complex cognitive task does not mean that the former causes the latter, but merely that they are associated in the data set analysed. There may be many intervening variables that can account for the association between the two measures. This issue is discussed fully in the chapter by Salthouse, and will not be discussed further. It is worth noting, however, that a convincing demonstration that changes in a primitive cause changes in more complex cognitive measures can only come with longitudinal designs, i.e. demonstrations that the degree to which an individual slows down is predictive of how much cognitive change they show in other domains.

Related to the point about causation is the issue of task purity. As was discussed in relation to ANOVA based analyses, no task is process pure, and this has important consequences for interpretation of correlational designs. Regression techniques are based upon individual differences, and the relation between individual differences and cognitive structure is by no means always obvious. This point was made in an analysis of the properties of motor cars by Lykken (1971). He asked the question of whether correlational analysis of the performance measures of cars would tell us something about how cars worked. Based on manufacturers' specifications, he found, for instance, that heavier cars tend to be able to accelerate faster and have a higher top speed than lighter cars. If we believed that correlation told us about structure, then we would be forced to argue that massive objects accelerate more easily than lighter ones, which is contrary to the laws of Physics. (In the motoring world, large cars have disproportionately large engines).

The difficulty in interpretation of behavioural data reflects the fact that none of the measures used directly tap the putative primitive that is used to explain age-related change in performance (see chapter by Rabbitt, this volume, for a fuller

articulation of this view). Neural speed is not measured directly, but instead is inferred from behavioural indices of speed in simple tasks. Likewise, working memory capacity is not measured directly, but is inferred from perfomance measures using specific memory tests. One might quibble that pure measures can never be obtained, but the point remains that individual differences in the behavioural measure cannot uniquely be attributed to individual differences in the primitive assumed to underlie performance.

A consequence of the fact that tasks do not uniquely measure what they purport to measure, and the fact that regression analyses are conducted in cross-sectional aging designs means that the potential for overestimating the effect of change in one variable on change in the other variable is high. In cross sectional designs - of the kind used by Baltes and Lindenberger (1997) - the measures are not just indications of domain specific ability, but also markers of the aging process and cohort differences. The older the individual then the more likely they are to be either chronically or acutely ill, to be less well educated, to be less familiar with testing, to be further away from formal education and so forth. It is also likely that they will bring different attitudes and motivations into the test session. All of these factors will mean that any measures of cognitive ability collected across a large age range will intercorrelate to some degree, whether the underlying constructs are related or not. For instance, memory performance on a recognition task and accuracy of responding in an inspection time task may be theoretically distinct, but education may affect both, and since older adults are in general less well educated than younger ones, the measures will intercorrelate. More specifically, the age-related variance in both may be highly related.

More recently, there has been a more direct criticism of the logic that

underpins the regression approach, including path analysis. Lindenberger and Pötter (1998) conducted a systematic analysis of the statistical properties of this approach. Their argument is best illustrated by reference to Figure 1.2. They were interested in whether controlling the strength of the age-independent relation between cognitive ability C and ability B, had an effect on the estimate of age-related variance that they share. Strictly, the logic of the argument that underpins the use of regression analyses, of the kind used by Baltes and Lindenberger (1997) is that the effect of controlling for ability C on the age (A) association with B (i.e. A-B) should be independent of the effect the age-independent relation between B and C. However, this was not the case. It was found that the influence that C had on the A-B relation was proportionate to the quadratic of the partial correlation between B and C. Thus, it is not possible to conclude that a particular primitive explains the age-related association between age and a dependent variable, because the estimate of the influence is confounded by the partial correlation between the dependent and the primitive, *that is age-independent*. i.e. corresponding to the area labelled "d" in Figure 1.2. Because the nature of the effect is quadratic, this can lead to outcomes that are unpredictable.

As an example, Lindeberger and Pötter (1997) show that path analyses which vary the partial correlation represented by "d" from zero to r = 0.572, whilst holding constant the relation between Age and B, can lead to radically different path models, which are either consistent with the view that all age-change is mediated by C, or are consistent with a positive effect of Age on B, independent of C, or a negative effect of Age on B, independent of C. i.e. with regards the independence of the Age - B association, all outcomes are possible, depending upon the partial correlation between B and C. Lindenberger and Pötter make clear

that such analyses technically do not tell us about how much age-related change in B is related to age-related change in C, but rather tell us whether age-related change in B is related to individual differences in C, which will include differences that are not age-related. Given that simple age correlations with measures are often in the region of 0.3 to 0.5, this means that only 9-25% of individual differences in test performance are age-related. With regards the general use of the technique, it is hard to disagree with Lindenberger and Pötter's (1998) conclusion that:

> "the decision to entertain the hypothesis that a certain variable mediates the causal effect of another should be based on theoretical considerations, *and not on the outcome of hierachical linear regression analyses.*" (p.227, italics added).

The use of Brinley plots

"Brinley plots" refer to scattergrams, plotted in Young-Old co-ordinate space. The name stems from Brinley (1965) who first represented his data in this fashion. Most conventionally, these are plotted using the mean response time from different experimental conditions (e.g. Cerella, 1985), although mean errors have also been used (e.g. Brinley, 1965). Sometimes the data represent pairs of individuals in a single experimental condition, with each point representing a younger and an older person in the same rank order in their respective population. Thus a point represents the response time for the fastest younger adult plotted against the response time for the fastest older adult. A second point represents the equivalent point for the second fastest individual in each population and so on (e.g. Maylor & Rabbitt, 1994)

Brinley plots have been used to address the dull hypothesis in the following

fashion. (We focus on the use of experimental means as the unit of analysis, but the argument applies to individual based analyses also). Data are collected across a range of experimental tasks, and plotted as a Brinley plot. Regression equations are then fit to these points, to determine the goodness of fit for simple functions. Age differences in such plots are revealed to the extent that points fall above the line y = x, i.e. to the extent that older response times are slower than younger ones. It has been found that, using such an approach, most of the variance between experimental conditions can be explained by simple functions. However, there is some debate as to whether a linear function, or a non-linear function best captures the regularity seen in the data (Hale, Lima & Myerson, 1991).

There has been considerable debate as to the interpretation of Brinley plots (e.g. Cerella, 1994, Perfect, 1994; Fisk & Fisher, 1994; Myerson, Wagstaff & Hale, 1994), and it is not the intention to revisit those arguments here. In any case the technique is discussed in full later in the volume (see chapters by Verhaeghen, Horn & Masunaga, Fisher, Duffy & Katsikopolous). However, we will note that the majority of Brinley plots are based on an analysis across experimental conditions, rather than across individuals. Thus estimates of the predictive power of these functions can be extremely misleading, because individual variance has been excluded. Such effects can be extremely powerful, and we illustrate with an example that is not age-related, to avoid clouding the issue.

Fienberg (1971) examined the relation between the probability of being drawn in the U.S. draft lottery, conducted in 1970, and when in the year a person's birthday fell. Across individuals, there was a correlation of rho = -0.226, which was moderately predictive, such that there was greater likelihood of being drawn for the draft if the person was born earlier in the year. However, Fienberg also collapsed the

data by month, and conducted another correlational analysis. Now the probability of being selected correlated with birthday, as measured by month at a level of rho = -0.839. Thus, using group means greatly increases the regularity of the data, and greatly inflates the apparent predictive power of the analysis; the analysis based upon means for each month suggest that most of the variance in probability of draft selection can be accounted for by birthday, but clearly, at the individual level this is not the case. Theory building based upon explanation of the regularity seen in the monthly figures would greatly overestimate the predictability of data at the individual level, and would likely lead to the adoption of single factor explanatory models. However, whilst such models would explain impressively high levels of variance at the level of the group, they would tell us little at the individual level.

The relation between method and theory in cognitive aging.

The previous sections have taken a critical look at the two kinds of methodology used to reject the dull hypothesis. These two methodologies have been linked with two broad theoretical approaches to understanding cognitive aging - the ANOVA based approach is associated with the modular or localist approach to aging, whilst the correlational approach is associated with the general or single factor approach to aging. Given the problems associated with both methodologies, what can be done to improve the understanding of our data and hence our theories of cognitive aging?

The work reported in the chapters that follow show a number of approaches to this question, ranging from empirical studies to mathematical modelling. Interestingly, despite the different approaches adopted by the authors in the present volume, there are consistent themes that run through the book. One such

theme is the relation between data interpretation and theory in aging. Too often in the past we have let the data, taken at face value, drive the theory. Now there is greater realisation that in fact the same data may be compatible with many theoretical approaches. Rather than worry about how our data inform our theories, many authors argue that our theories should inform our understanding of our data, and our drive to collect new data that is informative with regard theory.

In Chapter 2, Salthouse offers an overview of the methods used to answer theoretical questions in cognitive aging. He argues that dissociation based approaches, which seek to localize age effects in specific task-components do not offer explanations of cognitive aging, but merely refine what needs to be explained. He favours instead the broader approach to age-related change, and describes three methods for determining the extent to which age-related change in cognition is unique, or shared across tasks. His conclusion - that a relatively small number of factors can explain the majority of age-related cognitive change - offers a challenge to the remainder of authors in the volume. What Salthouse also makes clear in this chapter is the theoretical work that remains to be done: finding that age-related change is explained by a small number of cognitive primitives that can be called 'processing efficiency' does not in itself specify what those primitives are, nor what their explanatory status is. Are such primitives explanations in themselves, or are they merely markers of biological decline?

Chapter 3 (by Verhaeghen) makes the challenging statement that "everything we thought we knew about cognitive aging may be wrong" (p. xx) because we have used inappropriate methods. The chapter begins with a discussion of the problems with existing analytic techniques before suggesting the addition of a new technique to the armoury of cognitive aging researchers, namely

time-accuracy methodology. The starting point for this technique is the examination of state-traces, such as the time taken by younger and older adults to achieve the same level of accuracy, plotted across a range of accuracy levels. By manipulating response deadlines, functions describing the performance of younger and older adults can be constructed in terms of three underlying parameters, whose interpretation is theoretically driven. Verhaeghen makes two radical conclusions on the basis of the adoption of this approach. The first is captured in the quotation above - that interactions in behavioural data across tasks can arise even when the underlying parameters suggest no interaction, and conversely in some circumstances lack of interactions in behavioural data can mask interactions in the underlying parameters. Thus, Verhaeghen is making the very strong claim that behavioural data can only be understood in terms of an underlying theoretical model, and should not be taken at face value. The second radical conclusion is that dissociations with aging can be found, contrary to generalist accounts, and that such dissociations are caused by underlying 'quantum' states of complexity rather than cognitive domain (e.g. lexical vs non-lexical, executive vs non-executive). Verhaeghen concludes by arguing that viewing age-related change in this way represents a middle way between those who believe aging is all driven by a single underlying primitive, and those who believe in multiple independent or modular changes with age.

Fisher, Duffy & Katsikopolous' (Chapter 4) offer a theoretical analysis of a very different kind. This chapter is a formal mathematical analysis of the concept of slowing as it is used in cognitive aging research. They begin by arguing that previous research has conflated two distinct forms of slowing: chronological and chronocentric. By chronological slowing they mean the extent to which age-related

change is general, task- or process-specific. By chronocetric slowing they mean the extent to which the speed of baseline performance has an impact on criterion task performance, in younger and older adults. They go on to explore the extent to which these two constructs are related, and examine the effects of two estimates of chronocentric slowing - a novel measure they call P(speed) which is the proportion of the age difference in response times on criterion tasks that is due to differences in baseline speed, and the more conventional measure of how much age-related variance is explained by speed, P(shared). They argue that the former measure offers a more intuitive measure than the more commonly used technique, which can lead to misleading conclusions if there are task-specific age changes in slowing. Interestingly, although their methods are very different from the previous chapters, they end with a conclusion that would not be out of place in either of the preceding chapters. They argue that without a clear underlying theoretical model, data interpretation (in particular the estimate P(shared)) is problematic.

Horn & Masunaga (Chapter 5) bring a radically different theoretical perspective to the issue of human aging. They begin by reviewing the literature from the psychometric tradition that is supportive of the Gf-Gc theory of intelligence. They argue that the evidence favours the view that intelligence is multifactorial, and that a key part of the evidence for such a conclusion comes from the fact that across the adult lifespan some abilities decline whilst others either increase or are maintained. Since intelligence is not a single construct they further argue that it cannot be accounted for by a single construct such as 'g' or processing speed, and consequently, neither can age-changes in this construct be so explained. Instead they propose a very different way of thinking about intelligence and aging: they argue that the highest intellectual achievements are those reached by experts in a

domain. i.e. they are the result of extensive practice and specialization. They further argue that the evidence suggests that expertise takes years to achieve, and that older experts show maintenance of their skills. Thus, they argue, conclusions about intellectual change over the lifespan that are based on general tests may underestimate the abilities that older adults may reach in their areas of expertise.

Rabbitt (Chapter 6) offers a theoretical critique of the notion of single-factor accounts of cognitive aging. He begins with a critique of the construct of speed as a causal explanation in cognitive aging. He argues that speed measures are merely a measure of the efficiency of the cognitive system, rather than being a fundamental property of it. This is particularly true in simple tasks where the only way in which individuals can differ is the speed with which they reach asymtotic performance. In any case, argues Rabbitt, finding basic differences in speed merely begs the question of what causes the speed difference, since slower responses may stem from slower neurons, more impoverished neural networks, noisier systems and so forth.

Having discussed speed as a putative primitive for cognitive aging, Rabbitt goes on to offer an empirical demonstration that simple speed is not sufficient to capture all age-related individual differences. He shows that trial-to-trial variability is another stable individual characteristic that can explain individual differences in performance that mean response time cannot explain. However, the purpose behind this line of argument is not to argue for another primitive, but to demonstrate that single primitives are insufficient. Rabbitt is very clear that he does not believe in the usefulness of such an approach, instead favouring a model-driven strategy.

The following chapter (7) by Parkin and Java also focuses on the issue of general accounts of age-related decline. However, rather than discussing the issue

of general versus local accounts, Parkin and Java instead focus on contrasting three of the most commonly proposed 'primitives' that have been advanced as explaining age-related decline (frontal functioning, processing speed and fluid intelligence). Thus their approach can be seen as an attempt to answer some of the questions raised by Salthouse's chapter, in terms of what (set of) putative primitives might best explain cognitive aging. They then report a small-scale study which directly contrasts the explanatory power of the three primitives to explain age-related memory loss. They report a series of regression analyses in which performance on the digit symbol substitution test (DSST) (a measure of perceptual speed) is the strongest predictor of memory, with no residual effects of IQ or executive function. Thus, their empirical findings are consistent with those reported previously by Salthouse as supporting the idea that perceptual speed is the best 'primitive' for cognitive aging. However, these authors take a very different view of the DSST, and their chapter finishes with speculation as to what that test is measuring, and why it is so predictive of memory performance, and general intelligence. They argue that the DSST should instead be considered to be a form of working memory task, and that this may explain its success in predicting variance in memory performance.

Several of the chapters conclude with an appeal for clearer theoretical models within which to work: the chapter by Burke, Mackay and James (Chapter 8) represents just such an approach. This chapter focuses on one specific area of cognition - language functioning - as a test bed for theories of cognitive aging. They offer a clear theoretical model, and clear empirical dissociations with which to test their model against other theoretical accounts. The model - Node Structure Theory - in which aging is instantiated as a weakening of connections between units within

the language-processing system is tested with data from three domains of language production, namely tip of the tongue experiences, retrieval of names and spelling. Broadly, they report dissociations between language comprehension (input) and language production (output), with age-insensitivity for the former and age-decrements for the latter. They argue that whilst their model can account for such a pattern of findings, general accounts, such as generalised slowing, or decreased inhibition with age cannot.

The final chapter (9) by Light, Prull, La Voie and Healy also focuses on a particular theoretical model - the dual process model of memory - but uses a different analytic approach. These authors use meta-analysis to test whether effect sizes observed across many studies are compatible with a single-factor view of memory change across the lifespan. Thus this approach tests the general model of aging across many studies, thereby overcoming the problems associated with any single study that may result from particular samples, particular materials, or particular sets of instructions. Three theoretical areas of memory functioning are examined: implicit memory, recollection versus familiarity-based recognition memory (using the Tulving, 1985 technique), and intentional versus automatic processes (using Jacoby's, 1991 process-dissociation procedure). As well as examining whether age-related change in each area is compatible with single or dual-process models, the authors also attempt to relate the different areas of research.

Light et al.'s chapter is a thorough and scholarly piece of research in which the authors rigorously test alternative conceptualisations of their data sets in order to test alternative theoretical models. We do not wish to recapitulate the alternative classificatory schemes used in the metaanalysis at this point. However, there are two

clear themes that are noteworthy, and fit with the themes that emerge from the other chapters. First, it is clear that age differences are not equal in magnitude across the different data sets. However, more interesting is the fact that how one understands or interprets these effects is by no means clear or straighforward. Depending upon one's underlying theoretical model (e.g. should recollection and familiarity be seen as independent or redundant processes?), then the conclusions about age-sensitivity alter. Thus, even in large meta-analytic data-sets, the importance of a theoretical model emerges as paramount.

References

Baltes, P. B. and Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? *Psychology and Aging, 12*, 12-21.

Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birren (Eds.), Behavior, aging and the nervous system (pp.191-216). Springfield, IL: Charles C. Thomas.

Burke, D. M., Mackay, D. G. & James, L. E. (1999). Theoretical approaches to language and aging. In T. J. Perfect & E. A. Maylor (Eds.) *Models of Cognitive Aging*. Oxford: Oxford University Press.

Cerella, J. (1985). Information processing rates in the elderly. *Psychological Bulletin, 89*, 67-83.

Cerella, J. (1990). Aging and information processing rates. In J. Birren & K. W. Shaie, (Eds.) Handbook of the Psychology of Aging (3rd ed., pp 201-221). San Diego, CA: Academic Press.

Cerella, J. (1994). Generalized slowing in Brinley plots.. *Journal of Gerontology: Psychological Sciences*, 49, P65-P71.

Chapman, L. J. & Chapman, J. P. (1973). Problems in the measurement of cognitive deficit. *Psychological Bulletin, 79*, 380-385.

Dunlosky, J. & Salthouse, T. A. (1996). A decomposition of age-related differences in multitrial free recall. *Aging, Neuropsychology and Cognition, 3*, 2-14.

Hale, S., Lima, S. D. & Myerson, J. (1991). General cognitive slowing in the nonlexical domain: an experimental validation. *Psychology and Aging*, *6,* 512-521.

Horn, J. L. & Masunaga, H. (1999). New directions for research into aging and intelligence: the development of expertise. In T. J. Perfect & E. A. Maylor (Eds.)

Models of Cognitive Aging. Oxford: Oxford University Press.

Fastenau, P. S., Denburg, N. L. & Abeles, N. (1996). Age differences in retrieval: further support for the resource-reduction hypothesis. *Psychology and Aging, 11*, 140-146.

Fienberg, S. E. (1971). Randomization and social affairs: the 1970 draft lottery. *Science, 171*, 255-261.

Fisher, D. L., Duffy, S. & Katsikopolous, K. V. (1999). In T. J. Perfect & E. A. Maylor (Eds.) *Models of Cognitive Aging*. Oxford: Oxford University Press.

Fisk, A. D. & Fisher, D. L. (1994). Brinley plot and theories of aging: the explicit, muddled and implicit debates. *Journal of Gerontology: Psychological Sciences*, 49, P81-P89.

Fisk, A. D., Fisher, D. L. & Rogers, W. A. (1992). General slowing cannot explain age-related search effects: Reply to Cerella (1991). *Journal of Experimental Psychology: General*, *121*, 73-78.

Hertzog, C. (1989). Influences of cognitive slowing on age differences in intelligence. *Developmental Psychology, 25*, 636-651.

Jacoby, L. L. (1991). A process dissociation framework: separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*, 513-541.

Kirasic, K. C., Allen, G. L., Dobson, S. H., & Binder, K. S. (1996). Aging, cognitive resources and declarative learning. *Psychology and Aging, 11*, 658-670.

Light, L. L., Prull, M. W., La Voie, D. J. & Healy, M. R. (1999). Dual process theories of memory in old age. In T. J. Perfect & E. A. Maylor (Eds.) *Models of Cognitive Aging*. Oxford: Oxford University Press.

Lindenberger, U. & Pötter, U. (1998). The complex nature of unique and shared effects in hierarchical linear regression: implications for developmental

psychology. *Psychological Methods, 3*, 218-230.

Maylor, E. A. & Rabbitt, P. M. A. (1994). Applying Brinley plots to individuals: effects of aging on performance distributions in two speeded tasks. *Psychology and Aging, 9*, 224-230.

Myerson, J., Hale, S. Wagstaff, D. Poon, L. & Smith, G. A. (1990). The information-loss model: a mathematical model of age-related cognitive slowing. *Psychological Review*, *97*, 475-487.

Myerson, J., Wagstaff, D. & Hale, S. (1994). Brinley plots, explained variance and the analysis of age differences in response latencies. *Journal of Gerontology: Psychological Sciences*, 49, P72-P80.

Parkin, A. J. & Java, R. I. (1999). Determinants of age-related memory loss. In T. J. Perfect & E. A. Maylor (Eds.) *Models of Cognitive Aging*. Oxford: Oxford University Press.

Perfect, T. J. (1994). What can Brinley plots tell us about cognitive aging? *Journal of Gerontology: Psychological Sciences, 49*, P60-P64.

Perfect, T. J. & Rabbitt, P. M. A. (1993). Age and the divided attention costs of category exemplar generation. *British Journal of Developmental Psychology*, *11*, 131-142.

Rabbitt, P. M. A. (1999). Measurement indices, functional characteristics and psychometric contructs in cognitive aging. In T. J. Perfect & E. A. Maylor (Eds.) Models of Cognitive Aging. Oxford: Oxford University Press.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Salthouse, T. A. (1994). How many causes are there of aging-related decrements in cognitive function? *Developmental Review, 14*, 413-437.

Salthouse, T. A. (1996a). The processing speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403-428.

Salthouse, T. A. (1996b). General and specific mediation of adult age differences in memory. *Journal of Gerontology: Psychological Sciences, 51*, P30-P42.

Salthouse, T. A. (1999). Steps toward the explanation of adult age differences in cognition. In T. J. Perfect & E. A. Maylor (Eds.) Models of Cognitive Aging. Oxford: Oxford University Press.

Salthouse, T. A., Fristoe, N. & Rhee, S. H. (1996). How localized are age-related effects on neuropsychological measures? *Neuropsychology, 10*, 272-285.

Salthouse, T. A. & Somberg, B. L. (1982). Isolating the age deficit in speeded performance. *Journal of Gerontology, 37*, 59-63.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1-12.

Verhaeghen, P. (1999).The parallels in beauty's brow. Time-accuracy functions and their implications for cognitive aging theories. In T. J. Perfect & E. A. Maylor (Eds.) Models of Cognitive Aging. Oxford: Oxford University Press.

Figure 1.1: Typical Age x Task interaction plot showing a larger age difference in response latencies between young and old for a complex task than a simple task.

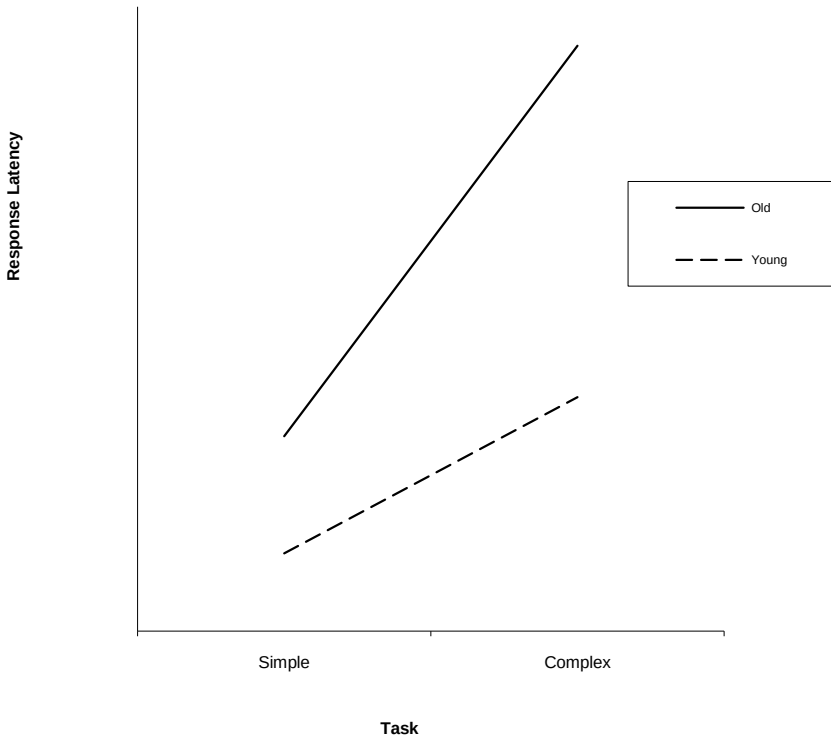

Figure 1.2: Venn diagram illustrating shared and unique variance associated with Age, and cognitive abilities B and C.