EUROPEAN JOURNAL OF COGNITIVE PSYCHOLOGY, 2001, *13* (1/2), 29–46

A research strategy for investigating group differences in a cognitive construct: Application to ageing and executive processes

Timothy A. Salthouse
Department of Psychology, University of Virginia, USA

A strategy involving five distinct phases is proposed as a means of obtaining the most informative evidence about group differences in a particular aspect of cognitive functioning. Perhaps because the phases require different types of analytical methods, most prior research has focused on only a few of the proposed phases. This is unfortunate because each phase provides valuable information about the nature of the relevant construct, and its role in group differences in cognition. The strategy is illustrated with research on adult age differences in executive processes.

How can one determine whether a specific cognitive construct is an important contributor to differences in cognitive functioning associated with membership in a particular individual differences category or group? In this paper I describe a strategy that could be used for this purpose, and I illustrate its application with research focused on adult age differences in executive processes. The strategy is designed to incorporate strengths of both experimental and correlational approaches, which have largely been pursued independently in past research on ageing and cognition.

The five phases of the proposed strategy are schematically illustrated in Figure 1. They consist of: (1) construct specification and operationalisation; (2) examination of group-related differences on a variable presumed to reflect the construct; (3) examination of unique group-related influences

Requests for reprints should be addressed to Timothy A. Salthouse, Department of Psychology, University of Virginia, Charlottesville, Virginia 22904–4400, USA.
Email: salthouse@virginia.edu

This research was supported by National Institute on Aging Grant R01 AG06826. I would like to thank Reinhold Kliegl and Dan Spieler for their helpful comments on an earlier version of this manuscript.

http://www.tandf.co.uk/journals/pp/09541446.html DOI:10.1080/09541440042000205

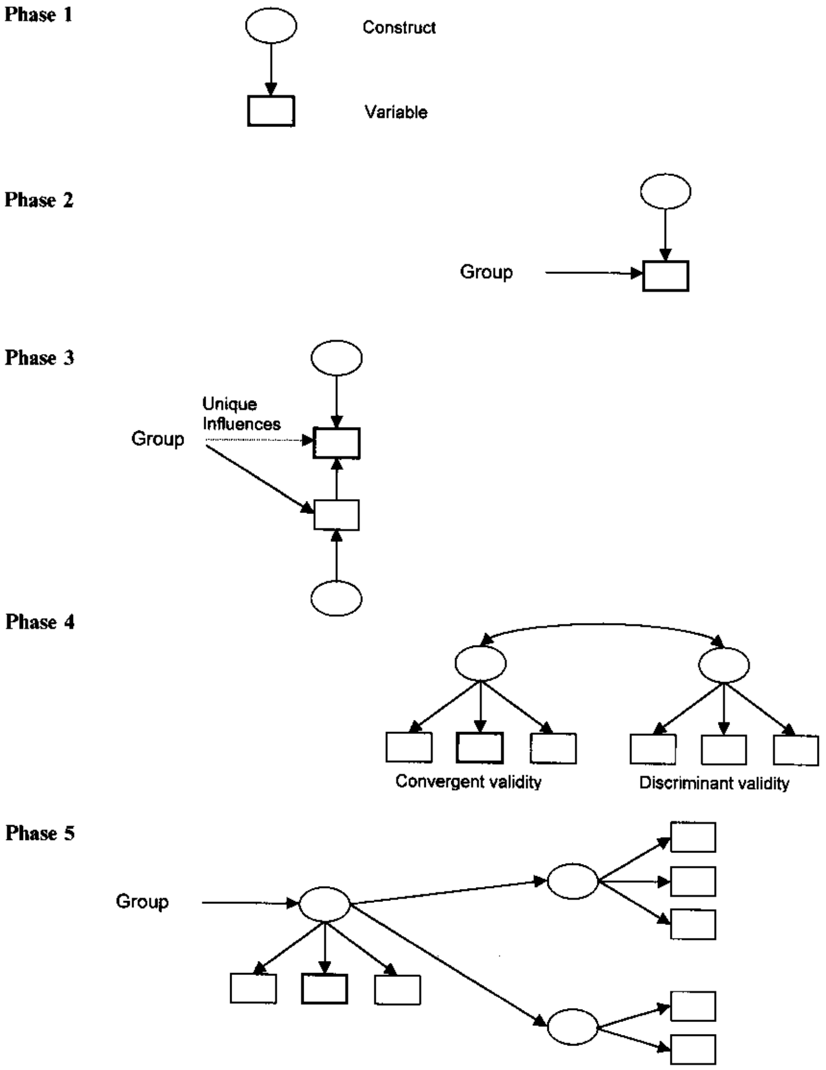

Figure 1. Schematic illustration of the five phases in the proposed strategy. Following convention in structural equation models, the squares represent manifest variables and the circles represent latent variables or constructs.

on the variable or construct; (4) investigation of convergent and discriminant validity of the construct; and (5) evaluation of the centrality of the construct in group differences in cognition. Although the phases are described sequentially, except for the first phase they need not proceed in any particular order and several may be combined in the same stage of data collection. This strategy is obviously not the only possible method of investigating group or between-person differences in cognition, but each of the phases can be justified as providing distinct and valuable information about the existence and nature of group differences in the relevant construct, and its role in other aspects of cognitive functioning.

The first phase consists of the development of theoretical assumptions or speculations about why the construct is important, and how it can be most meaningfully assessed. The primary purpose of this phase is to justify the relevance of the construct, and provide a rationale for a particular operationalisation of it. No data are collected or analysed in this phase, and instead the purpose is to establish a conceptual foundation for the tasks and variables used in subsequent phases.

The second phase in the proposed strategy consists of the collection of evidence that the groups differ in the level of the variable used to operationalise the construct. Several iterations of phases 1 and 2 are often needed to rule out construct-irrelevant influences and optimise the purity of the assessment because assumptions may have to be altered and tasks modified as data are collected. Because it is difficult to interpret the meaning of group differences with coarse variables that are likely to reflect a variety of different constructs, the refinement aspect of phases 1 and 2 is viewed as critical by many researchers. Depending on the theoretical assumptions, a variety of analytical procedures (e.g., subtraction, process dissociation, etc.) and derived measures (e.g., differences, proportions, etc.) could be used to try to achieve process purity. However, in all cases the major goal of phases 1 and 2 is to isolate the purest possible measure of the relevant construct, and then determine whether the groups differ in that measure.

The third phase in the proposed strategy involves obtaining evidence that the group differences in the relevant variable are at least partially independent of the group differences in other variables. Although frequently neglected, this phase is potentially quite important because many variables are often found to exhibit group differences, and it is desirable to establish that at least some of the group-related effects on the relevant variable are unique to that variable. That is, because organismic variables, or pre-existing individual difference categories such as group membership, typically cannot be randomly assigned or manipulated; people who differ on one variable are likely also to differ on other variables. However, if none of the group differences on the relevant variable or construct are independent of the group differences on variables that reflect other constructs, it is possible that the differences on the target variable are merely another manifestation of a more general phenomenon. A variety of statistical control procedures can be used to rule out this possibility, and establish that there are unique group differences on the relevant variable. Each procedure is designed to control the variation on other variables to allow a determination of the extent to which the group-related effects on the target variable are independent of the group-related effects on other variables.

The fourth phase in the proposed strategy consists of investigating construct validity by determining whether different operationalisations of the construct have moderately high correlations with one another, but relatively low correlations with variables presumed to reflect other constructs. Efforts in phases 1 and 2 are designed to maximise the correspondence between the construct and one particular variable, but they are unlikely to be completely successful because almost all constructs are broader than any single variable or else they would be of very limited interest. That is, even if the phase 2 research was successful in yielding a relatively pure assessment of the construct, purity needs to be complemented with breadth in order to ensure that the construct has meaning beyond one specific paradigm.

The idea that the most valid assessment of a construct should be based on several different operationalisations of that construct is similar to the recommendation by Garner, Hake, and Eriksen (1956) that converging operations be used in the investigation of theoretical constructs. The primary modification from the Garner et al. proposal is that in the context of individual differences research convergence is typically assessed with correlations. The rationale is that if two variables represent the same theoretical construct, then they should vary in a systematic fashion, such that individuals high on one variable should be high on the other variable reflecting that same construct. In contrast, much weaker relations should be apparent between variables representing different constructs. Correlations among variables hypothesised to reflect the same construct should therefore be moderately high to establish convergent validity, and substantially greater than the correlations among variables hypothesised to reflect different constructs to establish discriminant validity.

It is instructive to compare the approach to construct assessment in phases 1 and 2 with that in phase 4. As noted previously, phases 1 and 2 typically rely on an isolation approach by attempting to eliminate the contribution of construct-extraneous influences through task manipulations and various types of experimental control. In contrast, the approach in phase 4 relies more on the principle of aggregation by examining several variables that are all presumed to reflect the same construct, and then defining the construct by what is common among them. The two approaches are complementary because subtracting out irrelevant influences can be expected to yield a relatively pure but narrow construct, whereas averaging out irrelevant influences is likely to yield a less precise but broader and potentially more generalisable construct.

Notice that there are two key aspects to the proposed phase 4: (1) examining group differences in several different variables hypothesised to reflect the same construct to determine if the groups differ in the expected direction on each variable, and (2) examining the correlations among the variables to determine whether the variables reflect a distinct and unitary construct with respect to the pattern of individual differences. The first aspect is therefore necessary but not sufficient because group differences could exist in a variety of variables, and yet it cannot be determined whether the variables reflect the same construct or different constructs unless all the variables are available from the same individuals to allow correlations among them to be examined. If the variables reflect the same construct then not only should they all exhibit a similar pattern of group differences, but the correlations among the variables should be high, particularly relative to the correlations with variables hypothesised to reflect different constructs.

The final phase of the proposed strategy involves obtaining evidence that group-related effects on the target construct partially mediate group-related effects on other aspects of cognitive functioning. The rationale for this phase is that variables or constructs are generally of little interest in isolation, but instead are interesting with respect to their interrelations with, and impact on, other types of cognition. In order to establish that the construct is related to important and interesting aspects of cognitive functioning, therefore, empirical linkages need to be established both between group membership and the construct, and between the construct and variables reflecting other types of cognitive functioning.

Phase 2 research is concerned with the first of these linkages, but it is limited to a single indicator of the construct. Because few variables exhaust the meaning of the relevant constructs, and are seldom exclusively influenced by a single construct, the linkages are likely to be most meaningful when each construct is represented by multiple variables or indicators. Furthermore, the construct of interest should be linked not only to group membership, but also to variables representing other aspects of cognitive functioning. The most convincing evidence for this type of linkage is likely to be based on experimental interventions, in which the level of the critical construct is altered and concomitant changes are produced in the group differences in a variety of other cognitive variables. However, because experimental interventions may not always be possible, weaker evidence for the hypothesised linkages can be obtained from correlational data, in the form of structural models of the relations among group membership, the critical construct and other cognitive variables. Correlational data also allow statistical control procedures to be used to investigate implications of the hypothesised linkages. For example, if the construct mediates some of the group-related effects in cognitive functioning, then statistical control of the variation in the construct should lead to an attenuation of the group differences in other cognitive variables.

It is important to emphasise that different analytical methods are required across the five phases of the proposed strategy. Phase 1 is primarily theoretical to establish the conceptual foundation for the construct and its assessment, phase 2 is primarily experimental to refine and purify assessment of the construct, phases 3 and 4 rely on correlational procedures because they are based on the examination of relations among variables, and phase 5 is based on either experimental (strong) or correlational (weak) methods. From the perspective of the proposed strategy, therefore, no particular research method is inherently superior to another. Furthermore, exclusive reliance on a single method of investigation would be inappropriate because experimental control is needed to ensure precise assessment and to provide the strongest evidence for a causal role of the construct in group differences in other aspects of cognition, but correlational procedures are needed to evaluate the degree to which variables are independent of one another.

APPLICATION OF THE PROPOSED STRATEGY TO AGE DIFFERENCES IN EXECUTIVE PROCESSES

A considerable variety of tasks have been claimed to reflect executive processes. For example, within the field of neuropsychology the following tasks have been used to assess executive processes, with the accompanying types of justifications: The Wisconsin Card Sorting Test, because the examinee needs to maintain a task set (i.e., follow a rule to select responses), and exhibit flexibility to change the rule according to feedback and changing contingencies; verbal and figural fluency tests, because the respondent must generate multiple responses that satisfy particular constraints while flexibly using different retrieval strategies; various tower and maze tasks, because they require the planning and execution of a sequence of actions to achieve a goal; random generation tests in which the research participant attempts to produce a novel sequence of responses, and thus must inhibit or suppress stereotyped responses while keeping track of the responses already produced; the Stroop Colour–Word Test, because a prepotent response (reading a word) must be inhibited when making another response (naming a colour); and the Trail Making Test, because the examinee must maintain the position in one sequence while switching attention to a different sequence.

Although at least superficially these tasks appear to have little in common, the idea of executive processes has been appealing because the term refers to control processes that oversee the operation of more specialised processes. In this respect the concept is intended to capture aspects similar to the functioning of a business executive who is not a specialist in any particular area, but instead is responsible for the smooth functioning of an entire organisation. He or she creates and executes plans, establishes priorities, and monitors the execution of action sequences.

A more precise description of the processes often encompassed by the term executive has recently been provided by Smith and Jonides (1999, p. 1659). The five processes in their list are: (1) focusing attention on relevant information and processes, and inhibiting irrelevant ones; (2) scheduling processes in complex tasks, which requires the switching of focused attention between tasks; (3) planning a sequence of subtasks to accomplish some goal; (4) updating and checking the contents of working memory to determine the next step in a sequential task; and (5) coding representations in working memory for time and place of appearance. Subjective decisions are still required with respect to whether these processes are involved in a particular task, but the Smith and Jonides taxonomy is useful as a first step in determining whether a task reflects executive processes.

In the subsequent discussion I will focus on three sets of tasks each containing pairs of similar conditions designed to allow the isolation of a critical executive process: the Stroop Test, the Trail Making Test, and experimental switching tasks. In the familiar Stroop Test, the examinee is instructed to name colours as rapidly as possible when they are in the context of congruent words, incongruent words, or neutral stimuli (i.e., a row of Xs). The additional time in the incongruent condition compared to the neutral condition has been interpreted as a reflection of the failure to inhibit irrelevant information (i.e., Smith & Jonides criterion 1). One study in my laboratory was conducted with the traditional Stroop Colour–Word task (Salthouse, 1996). In another study (Salthouse & Meinz, 1995) three different versions of the Stroop task were administered: the standard version with colour names in different colours, a position version with location names in different positions, and a number version with digits in different quantities.

The stimulus materials in the Trail Making Test consist of a page containing an array of circled targets, and the task for the participant is to connect targets in either numerical order (version A), or in alternating numerical and alphabetical order (version B), as rapidly and accurately as possible. The additional time to complete the alternating sequences compared to the non-alternating sequences has been interpreted as a reflection of inefficient switching (Smith & Jonides criterion 2), and possible difficulty in the updating of working memory to maintain the correct positions in each sequence (Smith & Jonides criterion 4). Two studies in my laboratory were conducted with the standard version of the Trail Making Test (Salthouse, Fristoe, & Rhee, 1996; Salthouse, Toth, Hancock, & Woodard, 1997), one study (Salthouse et al., 2000) was conducted with an alternative paper-and-pencil version (Connections), and another study (Salthouse & Fristoe, 1995) was conducted with a computer-administered version.

In addition to tasks based on these neuropsychological tests, in another study in my laboratory (Salthouse, Fristoe, McGuthry, & Hambrick, 1998) specially created experimental tasks were designed to investigate the task switching process (i.e., Smith & Jonides #2). Three pairs of tasks with reaction time responses were used, and within each pair the tasks had different rules for mapping the stimuli to responses. The rules were: right/left—respond by either typing the digit on the right or on the left; more/odd—respond with one key if the digit is more than 5 and with another key if it is less than 5, or respond with one key if the digit is odd and with another key if it is even; and add/subtract—type the sum of, or the difference between, two digits. The additional time for the response when a switch to a different rule is required compared to when no switch is required can be assumed to reflect processes involved in switching such as redirection of attention, change in task set, reconfiguration of production, rules, etc.

The second phase of the proposed strategy consists of documenting age differences in measures of the relevant executive processes. Because the tasks just described each involve a pair of similar conditions, the differences between the scores in the two conditions can be postulated to reflect a specific executive process. Age differences could be examined by comparing the means of groups of young and old adults, but because all of the studies to be described involved samples with a continuous distribution of age, correlations between age and the difference score are reported. These values are presented in Table 1, where it can be seen that they ranged from .18 to .60. There were between 124 and 259 participants in each study, and thus all of the correlations were significantly different from 0. Because the contributions of other processes are presumably eliminated by subtraction of another similar variable, these results provide clear evidence of age differences in variables that can be hypothesised to reflect relevant executive processes.

The third phase in the proposed strategy involves estimating the unique age-related influences in the target variables. The estimates could be obtained with any of several different methods. For example, one possible method consists of statistically controlling the variance in a similar

TABLE 1
Age correlations in potential measures of executive processes

Variable	*Age correlation*	*Sample size*	*Source*
Stroop Colour Incongruent–neutral	.60	178	Salthouse (1996)
Stroop Colour Incongruent–neutral	.47	242	Salthouse and Meinz (1995)
Stroop Position Incongruent–neutral	.18	242	Salthouse and Meinz (1995)
Stroop Number Incongruent–neutral	.27	242	Salthouse and Meinz (1995)
Trail Making Test B–A	.53	259	Salthouse et al. (1996)
Trail Making Test B–A	.32	124	Salthouse et al. (1997)
Connections B–A	.41	207	Salthouse et al. (2000)
Computer-administered Trail Making B–A	.29	167	Salthouse and Fristoe (1995)
Switching—more/less Switch RT–Pre-switch RT	.38	161	Salthouse et al. (1998)
Switching—right/left Switch RT –Pre-switch RT	.38	161	Salthouse et al. (1998)
Switching—add/subtract Switch RT–Pre-switch RT	.41	161	Salthouse et al. (1998)

Note that the age of correlations are with difference scores created by subtracting the second variable from the first.

variable before examining the age-related effects in the target variable. With the three tasks described earlier, the controlled variable could be the neutral variable in the Stroop Test (i.e., the time to name the colours of Xs), the time in the non-alternating (A) condition in the Trail Making Test, and the reaction time on pre-switch trials in the switching tasks. (It is worth noting that this type of analysis of residuals is often recommended instead of comparisons of difference scores because unlike difference scores, the residual is independent of the original scores [e.g., Cohen & Cohen, 1983].)

A second possible method for estimating the unique age-related effects on a variable involves statistically controlling the variance in a different type of variable before examining the age-related effects on the target variable. Because letter comparison and pattern comparison tasks have been administered in many studies in my laboratory, results will be reported with a composite perceptual speed variable (formed by averaging the two *z*-scores) as the different type of controlled variable. However, it is important to note that virtually any variable could be used as the control variable because the question of independence of age-related influences could be applied to any combination of variables. A third method that could be used to derive estimates of unique age-related effects involves controlling an estimate of the variance common to many variables before examining the age-related effects on the target variable. Results from this method will not be described here, but it should be noted that several studies using this method have found results consistent with those from the other methods (e.g., Salthouse, 1996; Salthouse et al., 1996, 1997, 1998).

Table 2 summarises the proportions of variance associated with age in the target variable, and in residuals created by partialling either the variance of a similar variable or the variance of a perceptual speed composite. Because when a simple variable is partialled from a complex variable the age-related effects on the residual represent effects on the unique aspects of the complex variable, the purest assessment of the relevant executive processes construct might have been expected when the similar variable was controlled. Because the perceptual speed construct appears to have little in common with the complex Stroop interference or switching variables, the meaning of the residual formed by partialling the composite perceptual speed variable from these complex variables is somewhat ambiguous. Nevertheless, the two columns on the right of Table 2 indicate that there was considerable reduction of the age-related effects in the complex variable after partialling either the similar variable or the perceptual speed composite variable. The average total age-related variance across the 11 variables in Table 2 was .254, but the average unique age-related variance was only .032 both after control of a similar variable, and after control of the composite perceptual speed variable.

These results imply that an average of less than 13% (i.e., .032/.254 = .126) of the total age-related variance on the variables presumed to represent the relevant executive processes construct is unique to those variables. Because most of the age-related effects on the target variables are shared with age-related effects on other variables, these results raise the possibility that age-related effects on variables postulated to reflect executive processes are merely another manifestation of a broader phenomenon. In other words, these data suggest that, at least in terms of age-related influences, there may be very little that is special or unique about the purported executive process variables.

Because there is frequently confusion about the interpretation of statistical control results such as these, several points should be noted. First, it

TABLE 2
Total and unique age-related variance in target variables presumed to include executive processes

Variables		*Age-related variance*		
Target	*Similar*	*Total*	*After similar*	*After PSpd*
Stroop Colour Incongruent	Neutral	.429	.068	.057
Stroop Colour Incongruent	Neutral	.323	.038	.037
Stroop Position Incongruent	Neutral	.234	.009	.007
Stroop Number Incongruent	Neutral	.193	.001	.006
Trail Making Test B	A	.348	.056	.060
Trail Making Test B	A	.182	.033	.013
Connections B	A	.161	.00	.007
Computer-administered Trail Making B	A	.293	.041	.086
Switching—more/less Switch RT	PreSwitch RT	.159	.052	.025
Switching—right/left Switch RT	PreSwitch RT	.265	.016	.045
Switching—add/subtract Switch RT	PreSwitch RT	.202	.033	.012

Note: The sample sizes and sources of these comparisons are the same as those listed in Table 1. PSpd refers to the perceptual speed composite formed by averaging *z*-scores for the letter comparison and pattern comparison variables.

is important to recognise that proportions of variance are always equal to or greater than zero, and thus some of the residual R^2 values may actually reflect a reversal of the original age relation, and this possibility cannot be detected without inspecting other parameters from the regression analysis such as the beta coefficients. That is, the beta coefficient may change from –.5 to +.2 after statistical control, but the corresponding R^2 values would only change from .25 to .04, which is misleading with respect to the total change in the direction and magnitude of the age relation. Second, both statistical significance and the power to be able to detect an effect as significant need to be considered before interpreting the results. For example, a decrease from a significant R^2 associated with age to a nonsignificant R^2 is not particularly informative unless the research design had sufficient power to have been able to detect medium to small effects as significant. Third, the residual R^2 values should be examined in both relative and absolute terms. Relative comparisons, such as the residual age-related variance versus the initial age-related variance, are informative about the proportional contributions of different types of influences. However, because the residual age-related variance represents the age-related effects that are statistically independent of effects shared with other variables, it can be important regardless of its size relative to the initial age-related variance since it presumably requires a unique explanation. Finally, it should be recognised that results of statistical control analyses are not explanations, but rather are outcomes of methods intended to clarify the nature of the phenomenon that needs to be explained. A discovery that, for example, 87% of the age-related variance on a variable was shared with other variables therefore implies that unique or task-specific explanations are necessary, but that they are likely to contribute to a maximum of about 13% of the total age-related variance on the variable. Other types of explanations are consequently needed to account for the remainder of the age-related effects that are shared across several variables, and to indicate how those effects are manifested in particular variables.

The fourth phase in the proposed strategy focuses on investigating two aspects of construct validity. Convergent validity would be established if different variables hypothesised to represent the same construct have moderate to high correlations with one another, and discriminant validity would be established if those variables have much lower correlations with variables representing different constructs. Ideally, convergent validity would be evaluated with variables from tasks involving different methods, materials, and measures that are all presumed to reflect the same construct. Unfortunately, there is little evidence for this type of convergent validity for the executive processes construct because the correlations among variables from different tasks hypothesised to reflect executive processes are often no higher than the correlations with variables from tasks hypothesised to represent other constructs (e.g., Burgess, Alderman, Evans, Emslie, & Wilson, 1998; Duncan, Johnson, Swales, & Freer, 1997; Kopelman, 1991; Lehto, 1996; Rabbitt, 1997; Robbins et al., 1997; but see Della Sala, Gray, Spinnler, & Trivelli, 1998; and Hanes, Andrews, Smith, & Pantelis, 1996).

The analyses to be described therefore examine a very restricted type of convergent validity based on correlations among variables obtained from parallel versions of the Stroop and switching tasks. These data can be expected to provide upper limit estimates of the true convergent validity because the variables are derived from nearly identical tasks, and hence the resulting constructs are narrow due to the limited variation of methods, materials, or measures. In order to investigate discriminant validity, correlations were also examined with two variables hypothesised to represent a different construct, namely, perceptual speed. This also is not the optimum type of evaluation because the perceptual speed variables involved written responses, whereas the executive process variables were based on either vocal responses (for the Stroop tasks) or manual reaction time responses (for the switching tasks), and thus there is a confounding of theoretical construct and response mode. The confounding is likely to reduce the magnitude of the between-construct correlations, and hence the true degree of discriminant validity for the interference and switching constructs may be over-estimated with these data.

Although convergent and discriminant validity could be examined with raw correlations, or with reliability-adjusted correlations, a more powerful method relies on confirmatory factor analysis or structural equation modeling procedures. Two sets of analyses were therefore conducted on the data from Stroop interference (difference score) measures in three versions of the Stroop task (Salthouse & Meinz, 1995), and on the data from switching cost (difference score) measures in three switching tasks (Salthouse et al., 1998). Although difference scores frequently have low levels of reliability, that was not the case with these variables as the reliability estimates were all greater than .7. The first analysis was a confirmatory factor analysis with a perceptual speed factor represented by the letter comparison and pattern comparison variables, and either a Stroop interference factor or a switching factor each represented by three variables. A two-factor model fit the data in each data set quite well, which established that the constructs had convergent validity because the relevant variables loaded on the same factor, and discriminant validity because the correlation of the executive process factor with the perceptual speed factor was significantly less than 1.0.

The next analysis within each data set was a combination of the proposed phases 3 and 4. That is, this analysis consisted of a structural model with two latent constructs corresponding to perceptual speed and either Stroop interference or switching, and a path from the perceptual speed construct to the hypothesised executive process construct. In an initial model the age variable was related to both the speed and executive process constructs, but neither the path from age to the Stroop interference construct nor from age to the switching construct were significantly different from zero, and thus it was deleted from the model. The second model with each data set is illustrated in Figure 2, along with the fit statistics that indicate that the models provided excellent fits to the data.

Figure 2 provides information relevant to both phases 3 and 4 of the proposed strategy. That is, construct validity (phase 4) is established by the existence of two distinct constructs, but the absence of a direct path between age and either the Stroop interference construct or the switching

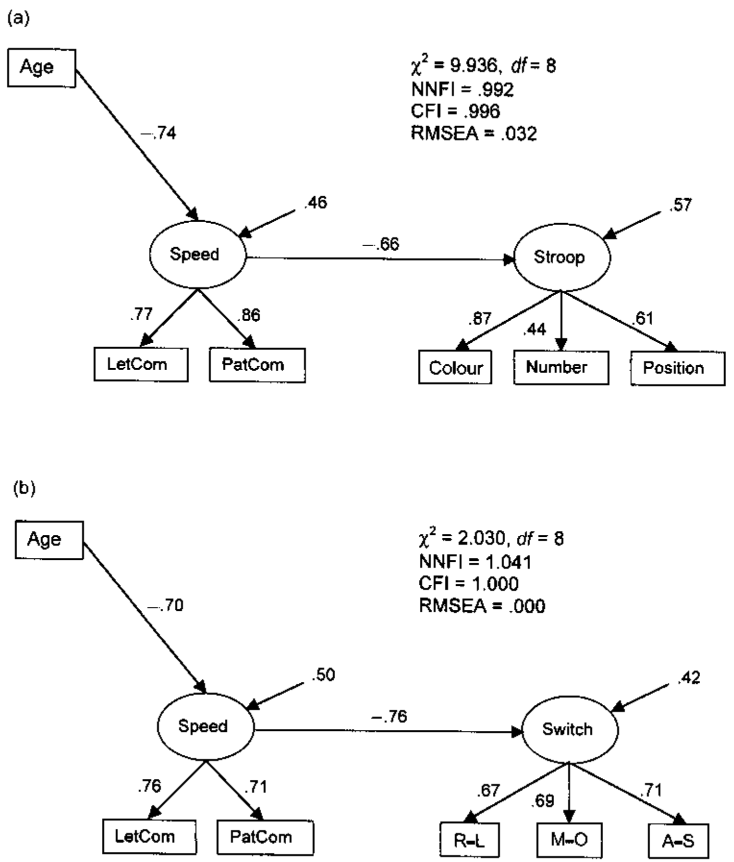

Figure 2. Structural models of the relations of age on perceptual speed and Stroop interference constructs (a), and of age on perceptual speed and switching constructs (b). No path is portrayed between age and the Stroop interference construct or between age and the switching construct because the path coefficient was not significantly different from zero.

construct means that there was no evidence of (statistically significant) unique age-related influences on the hypothesised executive processes constructs (phase 3).

The fifth phase in the proposed strategy consists of examining linkages between age and the construct, and between the construct and variables reflecting other aspects of cognitive functioning. The strongest and most convincing evidence for these linkages would be based on experimental manipulations of the critical construct because that would allow unequivocal conclusions about causal connections among the constructs. Unfortunately, there are not yet any experimental interventions of executive processes, in part because it is not obvious if, and how, constructs such as executive processes, or even more specific aspects such as resistance to interference or switching efficiency, might be altered.

A much weaker form of evidence for the hypothesised linkages is available from structural equation models based on correlations. The primary expectation with respect to the structural models is that executive process variables should function as a statistical mediator of age-cognition relations. That is, when executive process variables are included in the model the direct age-related effects on variables representing other aspects of cognitive functioning should be substantially smaller than the total age-related effects on those variables, because the total effects include both direct effects and effects that are mediated through the executive process constructs. A pattern such as this would be consistent with the hypothesis that the executive process construct may be mediating some of the age-related effects on other aspects of cognitive functioning. Evidence of this type is not definitive because there are many possible configurations of variables with correlational data, and it is well known that correlations by themselves do not imply causality. Nevertheless, correlational data do provide an opportunity to test critical implications of causal hypotheses. Specifically, if the target construct has little relation to age or to the other cognitive variables, or if there is little attenuation of the age-cognition relations when the construct is taken into consideration, then the results would be inconsistent with hypotheses that the construct is an important contributor to age differences in cognitive functioning.

Results from analyses reported in the original studies are consistent with the resistance to interference and switching constructs functioning as partial mediators of the age-related effects on other cognitive variables (Salthouse & Meinz, 1995; Salthouse et al., 1998, 2000). However, it is important to note that similar, and even stronger, patterns of mediation were evident with other potential mediators, such as perceptual speed. The interference and switching constructs are therefore not unique in exhibiting a pattern consistent with the mediation of age-related effects in other measures of cognitive functioning. Furthermore, because there was no evidence of unique age-related influences on the interference or switching constructs in the results summarised in Figure 2, all of the significant age-related effects on the executive process constructs in these studies appear to be mediated through perceptual speed.

CONCLUSION

What conclusion can be reached about age differences in executive processes based on the application of the proposed strategy? Perhaps the best one can say at the current time is that the evidence is still inconclusive. A major problem with respect to phase 1 is that it is not clear how one could determine whether executive processes were actually being assessed with a particular variable when no external criterion is available to validate that the variable truly reflects executive processes (e.g., Burgess, 1997; Tranel, Anderson, & Benton, 1994). In the absence of an explicit validation criterion, construct validity must be investigated with other techniques, such as examining the pattern of correlations with other variables, as in the proposed phase 4. There was clear evidence of age differences in variables hypothesised to reflect executive processes (phase 2), and the correlational evidence is consistent with the existence of Stroop interference and switching constructs distinct from a perceptual speed construct (phase 4), and with the involvement of these executive process constructs in the age differences in other cognitive tasks (phase 5). However, relative to the total age-related effects on these variables the unique age-related effects on the executive process variables were either very small or non-existent (phase 3), and similar or stronger patterns of mediation were evident with other types of variables and constructs.

Few if any phases in the proposed strategy are novel, but several of the phases have been neglected in prior research concerned with individual differences in cognitive functioning. For example, many researchers focus only on phases 1 and 2, and when group differences are found in the target variable it is often claimed that there are group differences in the relevant construct, and that the construct is responsible for at least some of the group differences in other types of cognition. Although evidence of group differences in the relevant variables is a necessary first step, it is insufficient to justify these other inferences. That is, with only data from phase 2 it is impossible to determine if the group-related effects on the relevant variable are unique to that variable or are shared with other variables that also exhibit group differences, the construct is defined very narrowly and can be considered to be confounded with a particular method of assessment, and there is no evidence that the construct is an important factor in group differences in other aspects of cognitive functioning.

Other researchers focus only on phases 4 and 5, and rely exclusively on correlational data. This type of evidence is also valuable, but it is incomplete in several respects. For example, interpretation of constructs may be difficult if none of the variables representing the construct is very precise, and if all are affected by an unknown mixture of different influences. Furthermore, because many alternative structural models are possible with correlational data, the data are often most informative in determining which models are *not* plausible rather than which of many possible models are plausible. Establishing that a particular model of the relations among constructs is the correct model may ultimately require other methods of investigation, such as experimental interventions.

It is not surprising that researchers from different backgrounds would emphasise different types of information or phases, but from the perspective of the proposed strategy this is undesirable because it results in the neglect of valuable information. There is clearly a need for variables that are precise and pure in reflecting the construct (phase 2), but there is also a need for evidence of construct validity in the form of moderately high correlations with other variables representing the same construct and low correlations with variables representing other constructs (phase 4). Moreover, in order to ensure that the relevant construct is both distinct and important with respect to group differences, the variable and construct should be found to have unique group-related effects (phase 3), and to have linkages to other variables and constructs (phase 5). As noted earlier, the order in which these different types of information is acquired is somewhat arbitrary, and in some circumstances it may be more efficient to combine several of the phases. Regardless of how the phases of the strategy are implemented, however, each aspect of the proposed strategy has the potential to contribute important information about the nature and role of group differences in cognitive constructs such as those intended to represent executive processes.

Manuscript received September 2000

REFERENCES

Burgess, P.W. (1997). Theory and methodology in executive function research. In P.M.A. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 81–116). Hove, UK: Psychology Press.

Burgess, P.W., Alderman, N., Evans, JH., Emslie, H., & Wilson, B.A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, *4*, 547–558.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Della Sala, S., Gray, C., Spinnler, H., & Trivelli, C. (1998). Frontal lobe functioning in man: The riddle revisited. *Archives of Clinical Neuropsychology, 13*, 663–682.

Duncan, J., Johnson, R., Swales, M., & Freer, C. (1997). Frontal lobe deficits after head injury: Unity and diversity of function. *Cognitive Neuropsychology, 14*, 713–741.

Garner, W.R., Hake, H.W., & Eriksen, C.W. (1956). Operationism and the concept of perception. *Psychological Review, 63*, 149–159.

Hanes, K.R., Andrewes, D.G., Smith, D.J., & Pantelis, C. (1996). A brief assessment of executive control dysfunction: Discriminant validity and homogeneity of planning, set shift, and fluency measures. *Archives of Clinical Neuropsychology, 11*, 185–191.

Kopelman, M.D. (1991). Frontal dysfunction and memory deficits in the alcoholic Korsakoff syndrome and Alzheimer-type dementia. *Brain, 114*, 117–137.

Lehto, J. (1996). Are executive function tests dependent on working memory capacity? *Quarterly Journal of Experimental Psychology, 49A*, 29–50.

Rabbitt, P.M.A. (1997). Introduction: Methodologies and models in the study of executive function. In P.M.A. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1–38). Hove, UK: Psychology Press.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., McInnes, L., & Rabbitt, P.M.A. (1997). A neural systems approach to the cognitive psychology of ageing using the CANTAB battery. In P.M.A. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 215–238). Hove, UK: Psychology Press.

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

Salthouse, T.A., & Fristoe, N. (1995). Process analysis of adult age effects on a computer-administered trail making test. *Neuropsychology, 9*, 518–528.

Salthouse, T.A., Fristoe, N., McGuthry, K.E., & Hambrick, D.Z. (1998). Relation of task switching to speed, age, and fluid intelligence. *Psychology and Aging, 13*, 445–461.

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., & Meinz, E.J. (1995). Aging, inhibition, working memory, and speed. *Journal of Gerontology: Psychological Sciences, 50B*, P297–P306.

Salthouse, T.A., Toth, J., Daniels, K., Parks, C., Pak, R., Wolbrette, M., & Hocking, K. (2000). Effects of aging on the efficiency of task switching in a variant of the trail making test. *Neuropsychology, 14*, 102–111.

Salthouse, T.A., Toth, J.P., Hancock, H.E., & Woodard, J.L. (1997). Controlled and automatic forms of memory and attention: Process purity and the uniqueness of age-related influences. *Journal of Gerontology: Psychological Sciences, 52B*, P215–P228.

Smith, E.E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science, 283*, 1657–1661.

Tranel, D., Anderson, S.W., & Benton, A. (1994). Development of the concept of "executive function" and its relationship to the frontal lobes. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 125–148). Amsterdam: Elsevier.