Prior to paradigm integration, the task is to resolve construct definitions of gF and WM

Damian P. Birney, David B. Bowman, and Gerry Pallier

School of Psychology (A18), University of Sydney, Sydney, NSW, 2006 Australia

damianb@psych.usyd.edu.au   **davidb@psych.usyd.edu.au**
gerryp@psych.usyd.edu.au
http://www.psych.usyd.edu.au/staff/damianb/

Abstract: Blair's account, like the intelligence field in general, treats many distinct constructs as if they were practically interchangeable – this is not self-evident. Paradigm integration and rationalization of redundant nomenclature are important for the continued development of understanding. The prior task is to demonstrate where synonymy of constructs across paradigms occurs, and where it fails. We present arguments why this is the case.

Blair's account of the diverse research on fluid cognition and related functions is an ambitious attempt at clarification. Certainly there is a need to dissociate more clearly the various fluid functions from psychometric or mathematical notions of *g*. We concur that *g* has little explanatory power and that the fluid intelligence (gF) framework promises to provide a more informative account of the processes underlying the quintessential aspects of intellectual functioning. As such, we see no value in continuing discussion of *g* other than as a point of historical interest. However, what must not happen is for gF and other constructs that Blair treats as mostly synonymous, such as executive function (EF) and working memory (WM), to be confused and allowed to become increasingly devoid of meaning.

The issue for our commentary has to do with definitions. Definitions tend to be obscured when research paradigms are integrated, because subtle and careful theorising *within* a domain does not always translate well *across* domains. Greater clarity in understanding the issues that Blair presents will not be obtained until more fundamental definitional issues are addressed. For instance, Blair projects and perpetuates confusion in the field's understanding of general intelligence by treating terms such as *g*, psychometric *g*, IQ, and even sometimes gC, as synonymous (the latter by equating verbal-scale IQ with general intelligence). Blair's account, like the intelligence field in general, treats many distinct constructs as if they were practically interchangeable – this is not self-evident. Hence, our critique is based as much on the field as it is on this particular review.

The gF-WM-EF issue is a further case in point. gF has developed meaning from within the psychometric domain, where it is common to define constructs not only by what they are, but also by what they are not. Hence, using factor-analytic techniques, gF has been *empirically* defined as the latent trait extracted from a variety of reasoning-dominated tests. This gF trait is related to, but empirically (and therefore theoretically) distinct from, the gC latent trait, which is similarly extracted from various tests of (typically verbal based) acculturated knowledge. WM theory was developed within the cognitive-experimental paradigm, mostly using dual-task methodologies to dissociate various storage and processing systems. EF has a more recent history and has been endorsed most actively by cognitive neuropsychology. The tasks used in these related, yet distinct, research programs have been developed with different purposes in mind. Hence, it is not always clear how theorising within one paradigm should be compared and integrated with theorising in another; it is not always clear how core processes identified using different methodologies in different paradigms can be compared (e.g., dual-task or factor-analytic methods); and it is not always clear how task- and paradigm-specific differences might obscure the detection of common underlying processes. Careful investigation is first required to build a common frame of reference between paradigms. Theories must then be integrated so that *strong* predictions about shared and unique processes can be tested. However, such practices are rare and, in their stead, findings from different paradigms have often been used to bolster validity claims for one paradigm without any real attempt to synthesise theory from the other. We need conceptualisations that are more precise and theoretically "risky" than the typical nomological network of loosely defined theoretical constructs that is pitted against an empirical bed of roughly convergent and divergent correlations (Borsboom et al. 2003; 2004).

The need for careful consideration of definitions is evidenced repeatedly in the literature. For instance, early seminal research cited by Blair appeared to demonstrate that reasoning ability was little more than WM (Kyllonen & Christal 1990). However, this work turned out to be based on a *weak theory* of WM (Conway et al. 2003). That is, the absence of a comprehensive understanding of the processes entailed in WM and the pragmatic approach towards task selection, serve as a caveat in accepting these early signs of synonymity. Despite far-reaching implications of this point for our theoretical understanding of human cognition, and despite due caution at the time by Kyllonen and Christal, this fact has often been overlooked. Research using *stronger* theories based on a more complete understanding of WM has suggested that less than half of the variance in performance on gF tasks can be accounted for by WM capacity (e.g., see Engle et al. 1999b). Crucial to much of Blair's argument is that the tasks used by different researchers to represent variously named constructs are similar in the processes they entail, even if the labelling is different. However, correlations are not sufficient evidence for this (Borsboom et al. 2004). Careful consideration of task characteristics is required because ultimately it is the tasks which provide the operational definition of the latent trait. One must not pass this responsibility off to psychometrics and neglect the importance of careful theory-driven task selection by arguing that factor analysis will distil the uncontaminated error-free latent trait from task-specific characteristics (cf. Colom et al. 2004). Indeed, evidence that the Flynn effect holds in some gF tasks and not others (sect. 3.1 of the target article) indicates that it matters which task(s) one selects as an indicator of gF!

So how are we to proceed through this quagmire? Paradigm integration and rationalization of redundant nomenclature are important for the continued development of understanding. However, the prior task is to demonstrate where *synonymity* of constructs across paradigms occurs, and where it fails. Synonymity means that the processes or constructs operationalised under the different paradigms (gF, WM, or EF, say) are linked to referents that are one and the same – covariation is a necessary condition, but it is not sufficient. gF, as originally conceptualised, is a broad multifaceted factor that psychometrically captures the essence of what is common in tasks requiring, for instance, inductive and deductive reasoning, quantitative reasoning, cognitive flexibility, abstraction of common principles, the development of strategies, and manipulation of mental representations (Carroll 1993). There is no doubt that WM capacity plays some role in determining performance on such tasks. Some argue that what is common is *controlled attention* (Engle et al. 1999a). It also seems clear, however, that when one considers the tasks used to operationalise gF, there is fundamentally more to the conceptualisation than controlled attention. Demonstration of causality, and not simply covariation, must be our goal. This requires an intimate understanding of underlying processes, which in turn requires an intimate understanding of the tasks (Borsboom et al. 2004). Psychometric research has for too long been agnostic to process theories. Blair's account – indeed any account of the field that does not take into consideration definitional differences across studies and across research paradigms – is destined to reflect much more confusion than clarity.