Mental Retardation Diagnosis and the Flynn Effect: General Intelligence, Adaptive Behavior, and Context

Tomoe Kanaya[1] and Stephen J. Ceci[2]

[1]Claremont McKenna College, [2]Cornell University

ABSTRACT—***Due to the Flynn effect and the IQ cutoff score required for the diagnosis of mental retardation (MR), educational and social policies created to reduce inequities faced by at-risk children in theory may actually perpetuate them in practice. We argue for the importance of considering IQ, adaptive behavior, and context when diagnosing MR and conclude by recommending that more research is needed to determine individual differences in the Flynn effect among children who are at risk of academic failure and therefore, most vulnerable to such differences.***

KEYWORDS—*Flynn effect; IQ; policy; special education; school children*

A century ago, Alfred Binet and Theodore Simon, under the commission of the French government, developed a measure to identify which children could benefit from special education services (Binet & Simon, 1908). This was the impetus for the IQ tests that are used today and set in motion the practice of using measures of general intelligence as part of the mental retardation (MR) diagnosis. Within the last few decades, the MR diagnosis has undergone various revisions and the importance of IQ is still debated, which was the focus of our article. We argued that because of the Flynn effect and changing IQ norms, the IQ cutoff score of 70 is resulting in MR misdiagnoses and subsequent misallocations of resources for children who are at risk for academic failure.

The commentaries by Hagen (2007) and Widaman (2007) provide strong contributions to this topic by pointing out a number of important issues that add further complexity to those we raised. They both address the double-edged "stakes" involved with misdiagnoses due to the Flynn effect; school children tested on older, inflated norms will not receive the MR services that they are legally entitled to under the Individuals with Disabilities Act (IDEA, Public Law 105-17), whereas children tested on brand new, accurate norms will be stigmatized by the MR label that they would have avoided had they been tested on an outdated test.

Drawing on Burton Black's harrowing descriptions, Hagen (2007) illustrates the cyclical nature of the injustices caused when using IQ tests with at-risk children, including children from low-income neighborhoods, historically marginalized populations, and boys. Indeed, there is increased pressure to identify children in need of special services at an early age so they do not miss the fundamental skills that are an integral part of the more advanced curricula in later grades. Early testing, in turn, increases the number of subsequent, required reevaluations children will encounter. As Hagen intimates, given the environmental disadvantages that face at-risk children, policies that are created to reduce these inequities may actually perpetuate them.

Hagen (2007), however, states, "(Our) confidence in the underlying worth of IQ . . . is not shared by many professionals." We take his point; we did not intend to convey a confidence in any of the widely used IQ tests, including the Wechsler. They are laden with assumptions that we have challenged elsewhere, especially their presumed context generality, a point to which

Correspondence concerning this article should be addressed to Tomoe Kanaya, Claremont McKenna College, Department of Psychology, 850 Columbia Avenue, Claremont, CA 91711; e-mail: Tomoe.Kanaya@ClaremontMcKenna.edu.

we return below (Ceci, 1996). We also do not wish to assert that IQ is a biological marker that is free of cultural values or real-world experience. Our argument centers on the applied (as opposed to theoretical) importance placed on a child's IQ score, regardless of the issues that surround the validity of these measures or the theoretical nature of intelligence. We focused on IQ because it is central to the diagnosis of MR and is a predictor of achievement in schools but not because of a belief in its generality or underlying biological nature.

Widaman (2007) correctly points out that, in practice, the MR cutoff captures somewhat more than the bottom 2.27% of the population that the normal curve deviate would lead one to expect. Of course, this amplifies the problem of upwardly creeping IQ scores, resulting in even more cases of misdiagnosis due to aging norms. Indeed, the recommended interval approach of 75 points captures nearly twice as many children at the low end of the distribution, further exacerbating the problem. It is important for researchers, policymakers, and school administrators to recognize the exact magnitude of the impact of the Flynn effect on MR diagnoses due to the cutoff score, as Widaman notes in his cogent commentary. In addition, it is important to remember that almost all children suspected of academic difficulties are tested on IQ tests, given that 11 of the 12 other special education diagnoses have to rule out MR to establish their own diagnoses.

As for possible solutions to the "roving cutoff" dilemma, Widaman (2007) takes issue with our suggestion that adaptive behavior occupies a more central role. He argues that we should "not forget that mental retardation is, first and foremost, a phenomenon involving cognitive abilities, manifested by a lower rate of development of these abilities. Intelligence tests assess these cognitive abilities; adaptive skill measures do not." We agree that MR is a cognitive-based disability, but only insofar as it impacts adaptive behavior; cognitive abilities cannot be assessed independent of the context. Memory, attention, and reasoning, for example, are highly sensitive to context, and the same individuals who reason brilliantly in one context sometimes reason poorly in others, including IQ-testing contexts (Ceci, 1996). As such, adaptive behavior is the context for judgments of MR. Indeed, few children would come to the attention of a school psychologist if their adaptive behavior was adequate because adaptive behavior during schooling is academic in nature.

Widaman (2007) endorses the idea of subtracting points from an individual's IQ to compensate for the Flynn effect (approximately 0.3 points per year), even though there is some evidence to suggest that the Flynn effect may not be uniform across all IQ levels, thus resulting in some added error. He is correct, however, about the limitations of the recent research (e.g., Kanaya, Ceci, & Scullin, 2005), and we find his argument persuasive that "it may be better and fairer . . . to employ an adjustment that is presumably a little too much for some and a little too little for others than to employ no adjustment at all."

Although a uniform adjustment *will* be a better and fairer than no adjustment at all, it should only be regarded as a temporary solution. Although the studies that have shown differential Flynn effects at different ages and on different tests are limited, they are the best evidence we have at present. Clearly, more research must be dedicated to dissecting or "unpacking" potential individual as well as developmental differences in the Flynn effect. And, in turn, policies that impact the educational opportunities of at-risk children must reflect the findings from this research.

REFERENCES

Binet, A., & Simon, T. (1908). Le développement de l'intelligence chez les enfants. *L'Année psychologique*, *14*, 1–94.

Ceci, S. J. (1996). *On intelligence: A bioecological treatise on intellectual development*. Cambridge, MA: Harvard University Press.

Hagen, J. (2007). The label mental retardation involves more than an IQ score: A commentary on Kanaya and Ceci (2007). *Child Development Perspectives*, *1*(1), 60–61.

Kanaya, T., Ceci, S. J., & Scullin, M. H. (2005). Age differences in secular IQ trends: An individual growth modeling approach. *Intelligence*, *33*, 613–621.

Widaman, K. F. (2007). Stalking the roving IQ score cutoff: A commentary on Kanaya and Ceci (2007). *Child Development Perspectives*, *1*(1), 57–59.