

Testing the limits of self-assessment: A critical examination of the developmental trajectories of self-assessment processes

Abstract

In 1998, Black and Wiliam's landmark review of the literature on assessment revealed that summative assessment was overemphasised, whereas formative assessment was underemphasised. It was argued that the latter was 'at the heart of effective teaching', with self-assessment being an integral component (p. 2). Consequently, Black and Wiliam's (1998b) seminal work set in motion a new impetus towards self-assessment in a plethora of recent policy documents, such as in the renowned Organisation for Economic Co-operation and Development (OECD) report, *Synergies for Better Learning* (2013). Accordingly, at a practical level, new approaches to assessment became embedded in an effort to create more opportunities for self-assessment. Such prioritisation may also have been enticed by the wealth of research that has highlighted the reputed benefits of self-assessment, chiefly, self-regulated learning. Hence, self-assessment has been implemented into policy and practices with forceful conviction. However, is the 'trust' bestowed on self-assessment granted somewhat prematurely and unsystematically? For example, the policymakers themselves have had trepidations regarding a possible disconnect between assessment policies and practices. Furthermore, it has been argued that we have deficient knowledge of children's engagement in the self-assessment process (Andrade & Du, 2007). Implicit in this issue is the notion that some children may be unable to engage in accurate self-assessments of their academic work. Specifically, it appears that some developmental pathways render some children more susceptible to making more inaccurate self-assessments than others, with children's academic abilities and gender also resulting in variability in the self-assessment process. Unfortunately, self-regulated learning, the most coveted by-product of the self-assessment process, has been found to be dependent on accurate self-assessments (Nicol, 2009). Thus, the literature reviewed here tells a cautionary

tale for policymakers and practitioners alike, hereby highlighting the need for a reformation of current assessment policy directives and practice in classrooms worldwide.

Keywords: self-assessment; development; policy; psychology; Piaget

Introduction

According to the *Organisation for Economic Co-operation and Development* (2012), assessment of students is ‘integral to the work of teachers’ and is consequently ‘uncontested and widely supported’ (OECD, 2012, p. 10). This quotation encapsulates the new impetus towards assessment globally, and the public trust accorded to its application in classrooms. In Ireland, the *National Council for Curriculum and Assessment* (NCCA; 2007) have defined assessment as ‘the process of gathering, recording, interpreting, using, and reporting information about a child’s progress and achievement in developing knowledge, skills and attitudes’ (p. 7). In recent years, assessment has emerged at the forefront of educational research, discourse and policies; noted as a central process of teaching and learning within the *Irish Primary School Curriculum* (Department of Education and Skills; DES, 1999). Accordingly, the NCCA’s (2007) document, *Assessment in the Primary School: Guidelines for Schools*, and the DES’ (2011) *National Strategy to Improve Literacy and Numeracy among Children and Young People*, have prioritised the use of assessment in classrooms, with the DES (2011) envisaging that assessment could improve literacy and numeracy standards. In fact, following Black and Wiliam’s (1998b) landmark review of the literature on assessment, formative assessment has been increasingly implemented in classroom practice in order to inform teaching and learning. Black and Wiliam (1998b) argued that formative assessment is ‘at the heart of effective teaching’, with self-assessment as an integral component (p. 2). Children can self-assess by identifying the strengths and weaknesses of their work and setting personal learning targets accordingly (NCCA, 2007). Hereby, self-assessment can be used to inform learning.

Black and Wiliam’s (1998b) seminal work resulted in global policy shifts which advocated the use of formative assessment methods, such as self-assessment in classrooms. This placed formative assessment at the forefront of legislation and policies in an effort to

create more opportunities for self-assessment (Broadfoot & Black, 2004). However, it appears that such policy shifts may have occurred somewhat prematurely; the concept of self-assessment may have been accepted by policymakers rather haphazardly, and hereby, uncritically. In fact, Broadfoot and Black (2004) explained that assessment methods that enjoy public legitimacy are often accorded trust by individuals, and so are not subject to the scrutiny that they should be. Thus, the current paper will critically examine the utility and reliability of self-assessment in classrooms. In particular, in line with Piaget's Theory of Cognitive Development (1978), the current paper will 'put to the test', the NCCA's (2007) assertion that self-assessment can be used 'by children of all ability levels' and 'throughout the primary school' years (p. 14); hence, the accuracy of schoolchildren's self-assessments will be explored. Given the evidence that self-assessments must be accurate in order to evoke self-regulatory processes (e.g., Nicol, 2009), this paper seeks to address a crucial, yet relatively unaddressed question; do we know enough about children's engagement in the self-assessment process in order to safely authorise its prioritisation in classrooms and policies worldwide?

Self-assessment: The policy and legislative context

Andrade and Du (2007) have argued that it is difficult to ascertain the most effective methods of using self-assessment in the classroom because not enough is known about children's engagement in the self-assessment process. Such concerns have been acknowledged by international bodies. For example, the OECD and Centre for Educational Research and Innovation (2008) have commented that although the principles of formative assessment have been applied at policy level, there are barriers which hinder its wider practice. In the OECD's influential report, *Synergies for Better Learning* (2013), it was also stated that securing a link between policy and practice is one of the main challenges in assessment. Accordingly in

Ireland, perhaps influenced by such global shifts, Grogan (2013) asserted that policies are idealistic and difficult to implement, and so are being applied in an ad-hoc manner.

Nevertheless, despite the recognition of these barriers to assessment, self-assessment is continuously being implemented into a plethora of recent policy documents. From an international perspective, the OECD and the Programme for International Student Assessment (PISA) have promoted the use of formative assessment in educational contexts in order to facilitate student learning and achievement (e.g. Looney, 2011; Nusche, Halász, Looney, Santiago & Shewbridge, 2011). In fact, in an OECD review of *Evaluation and Assessment in Education*, Nusche et al. (2011) stated that formative assessment needs ongoing attention with a particular emphasis on developing students' skills for self-assessment. This globalisation has resulted in countries having similar assessment policies and practices (Broadfoot & Black, 2004). Accordingly, the Irish Primary School Curriculum (1999) encouraged teachers to adopt a formative view of assessment, followed by the aforementioned NCCA (2007) document which encouraged the integration of formative assessment into educational practices. DES school inspectors have also commended the use of formative and self-assessment methods, with one department inspectorate (2014) advising that formative assessment practices 'should be developed further and implemented consistently throughout the school' (p. 4). In all, the predominance of formative assessment methods is very evident amongst Irish and global policy directives, with the intention that child-centred methods of assessment, such as self-assessment will enhance the teaching and learning experience.

Why is self-assessment being advocated?

Indeed, it does appear that self-assessment has the potential to enhance teaching and learning and so the prioritization of self-assessment in educational documents may be

warranted. Research illustrates how self-assessment encourages student involvement, motivation, learning and responsibility. It can contribute to a positive classroom environment and develop students' metacognitive skills by allowing them to recognise their strengths and weaknesses with regards to a piece of work (Andrade & Du, 2007; NCCA, 2007). Evidence also indicates that it can increase students' problem-solving abilities, reduce disruptive behaviour, as well as maintain high levels of student self-efficacy (Brookhart, Andolina, Zuza, & Furman, 2004; Paris & Paris, 2001; Ross, 2006). In turn, Bandura (1994) stated that high self-efficacy can increase student motivation, persistence at a task and target setting. Most importantly, research illustrates how self-assessment can improve children's learning to learn and metacognitive skills, and thus, develop children's self-regulatory skills (Brookhart et al., 2004). In fact, many of the perceived benefits of self-assessment are reputed with the benefits generated by self-regulation (Nicol, 2009). Studies have shown that self-regulation can predict children's academic success more powerfully than IQ, as well as literacy and mathematical attainment (Hutchinson, 2013; McClelland, Acock & Morrison, 2006).

Self-assessment: The challenges

However, despite the host of benefits and policies pertaining to self-assessment, a review of the literature illustrates that students' often overestimate their abilities, and as such, their self-assessments only hold a tenuous relationship with their actual performance. Mabe and West's (1982) meta-analysis of 55 studies revealed a weak overall relationship between predicted and actual ability ($r = .29$), while Freund and Kasten's (2012) revealed a moderate overall relationship ($r = .33$). Studies using children have also revealed weak correlations between students' predicted and actual abilities (e.g., Bouffard, Vezeau, Roy, & Lengelé, 2011; Sadler & Good, 2006; Sung, Chang, Chang & Yu, 2010). Most importantly, Zimmerman

(1990) contended that in order for children to develop self-regulatory skills, they must be aware if they possess a skill or not. Hereby, self-regulatory processes may be dependent on accurate self-assessments (Boseovski, 2010; Nicol, 2009). Evidence has also indicated that overestimations of ability can lead to external attributions of failure, maladjustment, poor social skills, narcissism, defensiveness and self-defeating behaviours. Försterling and Morgenstern (2002) found that overestimation of ability led to 'ability-insensitive time allocation' in a task, which resulted in participants allocating too little time to subtasks for which they had low abilities (p. 584). Conversely, those who underestimated their performance allocated excess time on material that they had actually mastered (Bol & Hacker, 2012). Research indicates that those who underestimate may also set lower goals for themselves, thus inhibiting achievement strivings (Ackerman & Wolman, 2007; Freund & Kasten, 2012). In sum, it appears that distorted self-assessments can result in negative educational outcomes and unfortunately may inhibit self-regulatory processes and their associated benefits, such as those aforementioned.

Self-assessment: A developmental viewpoint

Nevertheless, it is important to note that an analysis of the literature suggests that inaccuracies in self-assessments may be restricted to certain developmental periods (Boseovski, 2010; Folmer et al., 2008; Stipek & Mac Iver, 1989). In line, this questions the proposal made by the NCCA (2007) that self-assessment can be used throughout the primary school years. In fact, research illustrates that as children develop, they become more negative, albeit more accurate, in their self-assessments (e.g., Freedman-Doan et al., 2000; Pomerantz & Saxon, 2001). Interestingly, the changes in self-ratings may correspond with *Piaget's Stages of Cognitive Development* (1976). Piaget (1976) explained that cognitive changes occur at four different stages - the sensorimotor stage (infancy), the pre-operational stage (toddler and early

childhood), the concrete operational stage (elementary and early adolescence) and the formal operational stage (adolescence and adulthood). Pre-operational thinking is typically characterised by egocentrism, which Berk (2013) has defined as ‘the failure to distinguish others’ symbolic viewpoints from one’s own’ (p. 244). However, egocentrism diminishes during the concrete operational stage, and children begin to think concretely, where they require concrete materials as objects of thought. During the formal operational stage, concrete thinking is replaced with abstract thinking, whereby individuals are capable of more complex thinking (Berk, 2013). Each stage is characterised by a number of cognitive traits, which render certain children more likely to make inaccurate self-assessments than others.

For example, research outlines how the egocentric nature of young children could limit their perspective-taking abilities, resulting in overestimations of performance (Schneider, 1998; Stipek & Mac Iver, 1989). In addition, the wishful thinking hypothesis states that pre-operational thinkers often subordinate reality for wishful thinking which could lead to overestimations in performance (Butler, 1990). Schneider (1998) has provided direct evidence for the wishful-thinking hypothesis; he found that young children’s self-evaluations of how many balls they threw into a basket were not reflective of how they performed, but rather on how they wished they could have performed. Another explanation for young children’s overestimations of ability arises from the effort-ability paradigm (Stipek & Mac Iver, 1989). Young children are unable to differentiate effort from ability and consequently judge their performance based on the amount of effort they put into a task resulting in optimistic self-assessments (Pomerantz & Saxon, 2001).

An analysis of the literature has also revealed that there are changes in children’s self-appraisals during the middle childhood years, coinciding with the transition from pre-operational thinking into concrete operational thinking (Blatchford, 1997; Wigfield et al., 1997; Freedman-Doan et al., 2000). During the concrete operational stage, egocentrism diminishes

and concrete thinkers become less consumed by wishful thinking (Berk, 2013; Bulter, 1990). However, concrete thinking is still characterised by certain cognitive limitations which can lead to invalid self-assessments. Similarly to pre-operational thinkers, those at the concrete operational stage may also lack certain metacognitive processes (i.e., hypothetico-deductive reasoning) which would prevent them from making accurate self-assessments (Berk, 2013; Veenman, Wilhelm and Beishuizen, 2004). Pomerantz and Saxon (2001) have also reported that concrete thinkers lack an understanding of the relationship between effort and ability, leading to more optimistic self-assessments; this is similar to pre-operational thinkers.

However, the formal operational stage sees a sophistication of children's cognitive processes, which is accompanied by a steep decline in self-ratings, resulting in more accurate self-assessments (Archambault, Eccles & Vida, 2010; Stipek & Mac Iver, 1989). Evidence suggests that formal operational thinkers understand the relationship between effort and ability and so make more valid self-assessments (Stipek and Mac Iver, 1989). Dweck and Leggett (1988) argued that such an understanding results in individuals associating increased effort with lower abilities. Contrastingly, those in the pre- and concrete-operational stages of development may believe that increased effort is indicative of higher abilities. Furthermore, in accordance with *Piaget's Theory of Cognitive Development* (1976), Flavell (1992) stated that formal thinkers are capable of hypothetico-deductive reasoning. Hence, when faced with a challenge, they hypothesise which variables may affect an outcome and then deduce logical and testable inferences, enabling them to systematically isolate and combine variables in order to establish which inferences are true (Berk, 2013). This process is dependent on metacognitive control, which is required for making accurate self-assessments, indicating that formal operational children are indeed capable of higher self-accuracy (Veenman, Wilhelm and Beishuizen, 2004).

In sum, it appears that distinct characteristics of each of *Piaget's Stages of Cognitive Development* (1976) can influence the self-assessment process. Young children's biased self-assessments do not seem to be purposeful, but rather, a product of their cognitive limitations (Boseovski, 2010). However, as children develop, their cognitive processes become more sophisticated, and thus less susceptible to biased foresight, resulting in more accurate self-assessments. Hereby, given the research that has indicated that self-regulatory processes are dependent on accurate self-assessments (e.g., Nicol, 2009), the evidence reviewed challenges the well documented notion that children of all ages are able to engage in accurate self-assessments of their work.

Self-assessment: Prior academic attainment

Moreover, Folmer et al. (2008) has suggested that within-cohort factors may impact on self-assessment processes. Specifically, it appears that prior academic attainment may have an effect on self-assessments, albeit such an effect could be restricted to certain developmental periods. Kruger and Dunning (1999) explained that low achievers may lack the metacognitive skills to realise that they are unskilled. This well-established phenomenon, in which the 'unskilled are unaware', is known as the Dunning-Kruger effect (Kruger & Dunning, 1999, p. 1121). The bias has received much empirical support (e.g., Boud, Lawson and Thompson, 2013; Kwon & Linderholm, 2014; Pazicni, & Bauer, 2014). Studies using child participants have revealed similar trends (e.g., Kasperski & Katzir 2013; Sung et al., 2010). However, more empirical child studies are needed in order to withstand the evidence to the contrary; it has been debated that low academic attainment may lead to more nuanced self-assessments (Archambault et al., 2010). Furthermore, there is no available literature which has studied the interaction between development and prior academic attainment. Implicit in this issue is

evidence which suggests that cognitive variations characterising different developmental pathways render some cohorts more vulnerable to the effects of prior academic attainment than others.

For example, from a Piagetian perspective, it has been argued that preoperational thinkers are unable to seriate and so are unable to ‘order items along a quantitative dimension’, rendering them oblivious to grades (Berk, 2013, p. 250). Instead, they may rely on salient information such as praise and symbols as indicators of their competence (Stipek & Mac Iver, 1989). However, as children develop, they begin to compare their grades to those of their peers (Bouffard et al., 2011). In fact, it has been disputed that one of the most important developmental changes in self-assessment processes are social comparisons (Bouffard et al., 2011; Ruble, Boggiano, Feldman & Loebel, 1980). Thus, it may not be until children reach the proceeding developmental milestones for the interaction between academic ability and self-assessments to become apparent. Therefore, it is questionable that children of all ability levels are capable of engaging in accuracy self-assessments of their academic work, as had been proposed by the NCCA (2007) and other educational bodies.

Self-assessment: Gender

In addition to developmental stage and prior academic attainment, gender differences in the accuracy of self-assessments have been a popular focus of many studies. These studies have produced mixed findings, albeit such findings may be different for each developmental period. Notably, a host of studies have illustrated that boys were more likely than girls to overestimate their mathematical abilities (Blatchford, 1997; Sheldrake, Mujtaba and Reiss, 2014). In contrast, girls displayed higher competence beliefs for subjects such as music and reading (Archambault et al., 2010; Eccles, Wigfield, Harold & Blumenfeld, 1993). It appears

that participants' competence beliefs were influenced by gender-stereotypes. However, the aforementioned studies did not explore the relationship between gender and development, in contempt of the evidence which suggests that such research is warranted. For example, Kohlberg (1966) suggested that gender constancy does not develop until late in the primary school years. Accordingly, boys may be increasingly pessimistic about their literacy abilities as they develop, whereas girls may remain optimistic. Yet, no studies to date have examined gender differences in the self-assessment process from a developmental perspective. Hence, gender differences are relatively unaddressed in policy and practical guidelines.

Recommendations for research

Overall, it is evident that a paucity of research exists regarding the impact of developmental trajectories on the accuracy of children's self-assessments. The dearth of research in the area is noteworthy given that Irish policy documents, such as the NCCA's (2007) *Assessment in the Primary School: Guidelines for Schools*, and OECD documents such as *Synergies for Better Learning* (2013) have strongly advocated the use of self-assessment in classrooms and across age groups. Yet, Andrade and Du (2007) argued that we still do not know enough about children's engagement with self-assessment in order to construct a pragmatic theory of self-assessment, or to ascertain the most effective methods of using it in the classroom. Schunk (2008) has also acknowledged the scarcity of research in the area and has thus called for more developmental research to be conducted on children's metacognitive and self-regulative processes, with explicit reference to Piaget's developmental framework. Stipek and Mac Iver's (1989) Piagetian-based review on the accuracy of children's self-assessments has also cited some convincing evidence which suggested that self-assessment processes may be governed by specific developmental pathways. Nevertheless, in contempt of

these evidence-driven recommendations, few studies have studied self-assessment from a developmental perspective. And yet, policy directives are still advocating the use of self-assessment in classrooms across the globe and across development, in the absence of any substantial scientific evidence that has indicated that children of all ages can benefit from engaging in self-assessments of their work. Rather, a review of the extant literature indicates that children of all ages are incapable of engaging in accurate self-assessments of their work (e.g. Stipek & McIver, 1989).

However, Jambunathan and Burt's (2008) research has revealed that cognitive limitations cannot explain all of the variance in students' self-assessments. We cannot overlook the evidence which has suggested that gender and prior academic achievement can have a substantial impact on the accuracy of self-assessments (e.g., Sheldrake et al., 2014; Sung et al., 2010). Many researchers have recommended that these areas be explored (e.g., Andrade et al., 2008; Folmer et al., 2008; McQuillan, 2013). In addition, there have only been four studies that have researched rubric-referenced self-assessments (i.e., Butler, 1990; Higgins, Harris & Kuehn, 1994; Laveault & Miles, 2002; Sadler & Good, 2006). Moreover, these studies have not explored age or prior academic attainment as variables. This is remarkable seen as rubrics are a commonly used self-assessment tool (NCCA, 2007). Rubrics can be used by students in order to self-assess their academic work in literacy; they communicate expectations for the piece of writing by listing criteria describing the varying levels of quality across a scale (Andrade, 2000). However, very little is known about children's utility of rubrics for English writing and whether or not they can use them correctly. Therefore, the true potential of self-assessment tools, such as rubrics, in the learning and writing experience is questionable.

Overall, as a consequence of the gap in the extant literature, it is recommended that future research is conducted in order to examine the accuracy of schoolchildren's self-assessments across all ability levels, ages and across gender. Furthermore, it is recommended

that the reliability of self-assessment tools, such as rubrics, are also assessed. Accordingly, the findings from these studies could be used to better inform policy and practice.

Recommendations for policies and classroom practice

In the interim, it is notable that in the absence of this prerequisite research, policy documents such as the *Assessment in the Primary School: Guidelines for Schools* (NCCA, 2007) continue to advocate the use of self-assessment in classrooms. In line, it is suggested that their guidelines on self-assessment should be revised in consideration of children's ages, gender and prior literacy attainment; particularly given the research that has indicated that young children of lower abilities are particularly inaccurate in their self-assessments, thus inhibiting self-regulatory processes. Subsequently, it is strongly recommended that educational bodies overcome this impediment by developing age-appropriate self-assessment tools and instructional practices, with the intention of maximising students' ability to accurately self-assess. This indicates the need for both qualified and pre-service teachers to receive specific training in formative, and self-assessment, to ensure its correct application in the classroom. Indeed, Ross (2006) has acknowledged that self-accuracy can be improved through student training and other teacher methods.

In consideration of these recommendations at policy level, there are a number of practical considerations that may improve the accuracy of children's self-assessments. For example, Ross (2006) noted that children often make biased self-assessment innocently, and as such, it is advised that teachers illuminate the importance of being as honest as possible in self-assessments. This is in line with Freund and Kasten's (2004) recommendations for self-assessment instruction. Furthermore, it is imperative that teachers ensure that children undergo the self-assessment process under optimal conditions. Hereby, Ross (2006) has outlined a

number of dimensions for self-assessment instruction. First of all, it is important that the criteria for self-assessments are ‘child-friendly’ (e.g., appropriate language, criteria that they consider important). Secondly, children should be explicitly shown how to apply the criteria (e.g., teacher modelling). Third, providing children with feedback on their self-assessments can increase the validity of self-assessments. For example, Ross (2006) refers to a process of triangulating self-assessments with teacher, and peer assessments of the same work. He considered this process especially important for children who perceive effort as more important than actual performance, and therefore, this could be useful for children at the earlier stages of cognitive development. In sum, these recommendations should increase children’s opportunities for making accurate self-assessments. However, as has been aforementioned, future research into the most effective methods of increasing self-accuracy is welcomed.

Conclusion

Overall, it appears that children may be inaccurate assessors of their academic work. Given this general inability of children to self-assess their work with accuracy, the widespread practice of self-assessment in classrooms across the globe is questionable. In particular, this paper calls into question the true potential of self-assessment for enhancing teaching through self-regulated learning. Furthermore, it appears that there is a bias in the literature towards the promotion of the benefits of self-assessment, whilst failing to examine its prospective vulnerabilities. Therefore, more research which examines the self-assessment process of children of all ages and abilities is required before ‘self-assessment’ can be safely implemented into our policies and practices. Notwithstanding these limitations, and the impressive benefits that accompany accurate self-assessments, the development of age-appropriate self-assessment tools and child-centred instructional practices that enhance self-accuracy are also important tasks for researchers, educationalists and policy-makers alike.

In light of this critical reflection on the literature, it is evident that the sentiments expressed in Black and Wiliam's influential work (1998b) may still be applicable over fifteen years later. They stated:

Pupils can assess themselves only when they have a sufficiently clear picture of the targets that their learning is meant to attain. Surprisingly, and sadly, many pupils do not have such a picture, and they appear to have become accustomed to receiving classroom teaching as an arbitrary sequence of exercises with no overarching rationale (p. 143).

In light of this statement, it appears incumbent that all educational practitioners, policy-makers and researchers collaborate to address this issue, in order to clarify this picture for the child. In doing so, parties must test the limitations of self-assessment, and indeed, test the limitations of children of all ages and ability levels and how such shortcomings impede on their self-assessments. In essence, the literature reviewed tells a cautionary tale of the inadequacies of self-assessment, a tale which follows a somewhat rhetorical and disjointed narrative. However, until we extend beyond the rhetoric, we need to be vigilant of employing self-assessment strategies in classrooms, in the absence of substantial knowledge of children's engagement in the self-assessment process.

References

- Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and self-concept measures. *Journal of Experimental Psychology: Applied*, 13(2), 57.
- Andrade, H. & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment and Evaluation in Higher Education*, 32 (2), 159-181.
- Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, 57(5), 13–18.
- Andrade, H. L., Du, Y., & Wang, X. (2008). Putting Rubrics to the Test: The Effect of a Model, Criteria Generation, and Rubric-Referenced Self-Assessment on Elementary School Students' Writing. *Educational Measurement: Issues and Practice*, 27(2), 3-13.
- Archambault, I., Eccles, J. S., & Vida, M. M. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology*, 102(4), 804-816.
- Bandura, A. (1994). *Self-efficacy*. John Wiley & Sons, Inc. Retrieved from http://www.happyheartfamilies.citymax.com/f/Self_Efficacy.pdf
- Berk, L. E. (2013). *Child Development* (9th ed.). New Jersey, USA: Pearson.
- Black, P., & Wiliam, D. (1998b). *Inside the black box: Raising standards through classroom assessment*. Granada Learning. Retrieved from <http://faa-training.measuredprogress.org/documents/10157/15652/InsideBlackBox.pdf>
- Blatchford, P. (1997). Pupils' self assessments of academic attainment at 7, 11 and 16 years: Effects of sex and ethnic group. *British Journal of Educational Psychology*, 67(2), 169-184.

- Bol, L., & Hacker, D. J. (2012). Calibration research: where do we go from here? *Frontiers in Psychology*, 3(229). doi: 10.3389/fpsyg.2012.00229.
- Boseovski, J. J. (2010). Evidence for “rose-colored glasses”: An examination of the positivity bias in young children’s personality judgments. *Child Development Perspectives*, 4(3), 212-218.
- Nusche, D., Halász, G., Looney, J., Santiago, P., & Shewbridge, C. (2011). *OECD Reviews of Evaluation and Assessment in Education: Sweden 2011*. OECD Publishing.
- Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education*, (ahead-of-print), 1-16.
- Bouffard, T., Vezeau, C., Roy, M., & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research*, 50(4), 221-229.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education*, 11(1), 7-26.
- Brookhart, S. M., Andolina, M., Zuza, M., & Furman, R. (2004). Minute math: An action research study of student self-assessment. *Educational Studies in Mathematics*, 57(2), 213-227.
- Butler, R. (1990). The effects of mastery and competitive conditions on self-assessment at different ages. *Child Development*, 61(1), 201-210.
- Department of Education and Skills (1999). *Primary School Curriculum: Introduction*. Dublin: The Stationary Office.
- Department of Education and Skills (2011). *Literacy and Numeracy for Learning and Life: The National Strategy to Improve Literacy and Numeracy among Children and Young People 2011- 2020*. Dublin: The Stationary Office.

- Department of Education and Skills (2014). *Whole School Evaluation Report: Tralee Educate Together Primary School*. Retrieved from http://www.education.ie/en/Publications/Inspection-Reports-Publications/Whole-School-Evaluation-Reports-List/report1_20158A.pdf.
- Dunning, D., & Helzer, E. G. (2014). Beyond the correlation coefficient in studies of self-assessment accuracy commentary on Zell & Krizan (2014). *Perspectives on Psychological Science*, 9(2), 126-130.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95(2), 256.
- Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self-and task perceptions during elementary school. *Child Development*, 64(3), 830-847.
- Flavell, J. H. (1992). Cognitive development: Past, present, and future. *Developmental Psychology*, 28(6), 998.
- Folmer, A. S., Cole, D. A., Sigal, A. B., Benbow, L. D., Satterwhite, L. F., Swygert, K. E., & Ciesla, J. A. (2008). Age-related changes in children's understanding of effort and ability: Implications for attribution theory and motivation. *Journal of Experimental Child Psychology*, 99(2), 114-134.
- Försterling, F., & Morgenstern, M. (2002). Accuracy of self-assessment and task performance: Does it pay to know the truth? *Journal of Educational Psychology*, 94(3), 576.
- Freedman-Doan, C., Wigfield, A., Eccles, J. S., Blumenfeld, P., Arbreton, A., & Harold, R. D. (2000). What am I best at? Grade and gender differences in children's beliefs about ability improvement. *Journal of Applied Developmental Psychology*, 21(4), 379-402.

- Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin*, 138(2), 296.
- Grogan, L. (2013). *The success of self-assessment strategies in enhancing classroom instructional practices and pupil learning* (Master's thesis). Mary Immaculate College, University of Limerick.
- Higgins, K. M., Harris, N. A., & Kuehn, L. L. (1994). Placing assessment into the hands of young children: A study of student-generated criteria and self-assessment. *Educational Assessment*, 2(4), 309-324.
- Hutchinson, L. R. (2013). Young children's engagement in self-regulation at school. Retrieved from https://circle.ubc.ca/bitstream/handle/2429/44401/ubc_2013_fall_hutchinson_lynda.pdf?sequence=1.
- Jambunathan, S., & Burts, D. C. (2003). Comparison of perception of self-competence among five ethnic groups of preschoolers in the US. *Early Child Development and Care*, 173, 651-660.
- Kasperski, R., & Katzir, T. (2013). Are confidence ratings test-or trait-driven? Individual differences among high, average, and low comprehenders in fourth grade. *Reading Psychology*, 34(1), 59-84.
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex- role concepts and attitudes. In E. E. Maccody (Ed.), *The development of sex differences*. Stanford, CA: Stanford University Press.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121.

- Kwon, H., & Linderholm, T. (2014). Effects of self-perception of reading skill on absolute accuracy of metacomprehension judgements. *Current Psychology*, 1-16.
- Laveault, D., & Miles, C. (2002). *The study of individual differences in the utility and validity of rubrics in the learning of writing ability*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Looney, J. (2011). Integrating formative and summative assessment: Progress towards a seamless system? OECD Publishing: Paris. Retrieved from [http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/wkp\(2011\)4&doclanguage=en](http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=edu/wkp(2011)4&doclanguage=en).
- Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, 67(3), 280.
- McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, 21(4), 471-490.
- McQuillan, J. (2013). Urban Middle and High School Students Reading Attitudes and Beliefs: A Large-Sample Survey. *GJHSS-G: Linguistics & Education*, 13(7).
- National Council for Curriculum and Assessment (2007). *Assessment in the Primary School: Guidelines for Schools*. Dublin: NCCA.
- Nicol, D. (2009). *Transforming assessment and feedback: Enhancing integration and empowerment in the first year*. Glasgow: QAA Scotland.
- Organisation for Economic Co-operation and Development (2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. Retrieved from

http://www.oecd.org/edu/school/Synergies%20for%20Better%20Learning_Summary.pdf.

Organisation for Economic Co-operation and Development/Centre for Educational Research Innovation. (2008). *Assessment for learning: Formative assessment. International Conference, Learning in the 21st Century: Research, Innovation and Policy*. Paris: CERI/OECD. Retrieved from <http://www.oecd.org/dataoecd/19/31/40600533.pdf>.

Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89-101.

Pazicni, S., & Bauer, C. F. (2014). Characterizing illusions of competence in introductory chemistry students. *Chemistry Education Research and Practice*, 15(1), 24-34.

Piaget, J. (1976). Piaget's theory (pp. 11-23). Springer Berlin Heidelberg. Pomerantz, E. M., & Saxon, J. L. (2001). Conceptions of ability as stable and self-evaluative processes: A longitudinal examination. *Child Development*, 72(1), 152-173.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation*, 11(10). Retrieved from <http://pareonline.net/getvn.asp?v=11&n=10>.

Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebel, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology*, 16(2), 105.

Sadler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, 11(1), 1-31.

Santiago, P., Donaldson, G., Looney, A. & Nusche, D. (2012). *OECD Reviews of Evaluation and Assessment in Education: Portugal*. OECD Publishing. Retrieved from <http://www.oecd.org/edu/school/50077677.pdf>.

- Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science, 1*(2), 291-297.
- Schunk, D. H. (2008). Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review, 20*(4), 463-467.
- Sheldrake, R., Mujtaba, T., & Reiss, M. J. (2014). Calibration of self-evaluations of mathematical ability for students in England aged 13 and 15, and their intentions to study non-compulsory mathematics after age 16. *International Journal of Educational Research, 64*, 49-61.
- Stipek, D., & Mac Iver, D. M. (1989). Developmental change in children's assessment of intellectual competence. *Child Development, 521-538*.
- Sung, Y. T., Chang, K. E., Chang, T. H., & Yu, W. C. (2010). How many heads are better than one? The reliability and validity of teenagers' self-and peer assessments. *Journal of Adolescence, 33*(1), 135-145.
- Veenman, M. V., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*(1), 89-109.
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*(3), 451.
- Zell, E., & Krizan, Z. (2014). Do people have insight into their abilities? A metasynthesis. *Perspectives on Psychological Science, 9*(2), 111-125.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17.