THE RELATIONSHIP BETWEEN INDIVIDUAL ACHIEVEMENT AND CLASS-AVERAGE ACHIEVEMENT ON PUPILS' INTRINSIC MOTIVATION

Aoife Cassidy

Supervisors: Dr Stella Long and Dr Paul Mulcahy

Doctoral thesis submitted to the

Department of Educational Psychology, Inclusive and Special Education,

Mary Immaculate College,

In Partial Fulfilment of the Requirements for the Award of

Doctorate in Education and Child Psychology.

Submitted to Mary Immaculate College, September 2021

Word Count: 29977 (exclusive of tables, figures. references and appendices)

Abstract

Title: The Relationship Between Individual Achievement and Class-Average Achievement on Pupils' Intrinsic Motivation

Aoife Cassidy

Background: Three decades of research have investigated the relevance of the big-fish-little-pond effect (BFLPE) and the internal/external frame of reference model (I/E model) to pupils' academic self-concepts (ASCs). The BFLPE posits that pupils' ASCs are influenced by comparisons which pupils make between their performance and that of their peers. Consistent with this theory, research has found that class/school-average achievement is negatively correlated with pupils' ASCs. Research investigating the I/E model has found that pupils' individual academic achievement is positively correlated with their ASCs in corresponding domains (i.e. the same academic domain or subject area, such as reading achievement and reading self-concept), but negatively correlated with pupils' ASCs in contrasting domains (i.e. academic domains or areas that are distinctively different from each other, such as reading achievement and mathematics self-concept, and/or mathematics achievement and reading self-concept). Recently, a combined model which simultaneously investigates the predictions of both models has found that the predictions of both models are also supported within the unified model. However, a paucity of research has investigated the relevance of these models to other motivational constructs.

Aims: This research aims to investigate the relevance of the BFLPE, the I/E model, and the combined model to fourth-class pupils' intrinsic motivation for mathematics and reading. Subsequently, pupils' ASCs are explored as potential mediating variables between individual and class-average achievement and pupils' intrinsic motivation.

Sample(s): Cross-sectional Irish data from the 2011 Trends in International Mathematics and Science Study and Progress in International Reading Literacy Study Combined International Database.

Methods: A series of latent-manifest structural equation models were employed to analyse the data.

Results: All predictions of the combined model were not supported among the chosen sample. However, individual reading achievement was positively correlated with pupils' intrinsic motivation for reading, and class-average mathematics achievement was negatively correlated with pupils' intrinsic motivation for mathematics.

Conclusions: The findings extend current knowledge as limited research to date has investigated the predictions of the combined model on motivational constructs outside of ASC. The findings are also pertinent to informing practice, as recent education guidelines have emphasised the development of positive attitudes towards learning as key objectives.

Declaration

I declare that this research dissertation, entitled ‘The Relationship Between Individual Achievement and Class-Average Achievement on Pupils' Intrinsic Motivation’ is my own original work. All information from other sources has been cited and acknowledged appropriately.

Signed: 

Date: 01/09/2021

Acknowledgements

To the school communities who were initially approached to participate in the original research study, I am extremely grateful for your time and interest in the project. Although the intended study could not be completed due to the unprecedented pandemic, your time and kindness at that stage was truly valuable. Thank you to the pupils, teachers, parents, and principals who were invited to participate at that stage.

To my two research supervisors, Dr Stella Long and Dr Paul Mulcahy, thank you for sharing your expertise with me throughout this journey. Through undertaking this project with your supervision, my research skills have been greatly enhanced.

Thank you also to everyone in the wider community at Mary Immaculate College who has supported my learning during both my undergraduate and postgraduate studies. In particular, a special thanks to each member of the Professional Doctorate in Educational and Child Psychology (DECPsy) team for their encouragement and guidance across all aspects of the course.

To my friends, thank you for your understanding, patience, and encouragement throughout my studies. Thank you also for taking an interest in the project, and for bringing my attention to other interests when distractions from the thesis were needed. Sharing these times with you often meant that I reapproached the research with a fresh perspective.

I am also thankful to have undertaken this journey alongside my DECPsy classmates. I am grateful to have had the opportunity to get to know all of you and to have been able to share the ups and downs with you along the way.

Finally, I am indebted to my family for their continuous support throughout this journey. To my parents Celine and Brian, I am forever grateful for all of the educational opportunities which you have provided me with throughout my life. Thank you for your persistent encouragement and belief in me. Without you, none of this would have been possible. To my brothers, Cian and Eoin, thank you for your understanding and encouragement during my studies. And to my granny, Marie Lande, and my uncle, Kevin, thank you for your extremely welcoming and generous hospitality during the last 18 months of the DECPsy course.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

AIC	Akaike Information Criterion
ASC	Academic Self-Concept
BFLPE	Big-Fish-Little-Pond Effect
BIC	Bayesian Information Criterion
CFI	Comparative Fit Index
CI	Confidence Interval
DECPsy	Professional Doctorate in Educational and Child Psychology
DES	Department of Education and Skills
DH	Department of Health
ES	Effect Size
IEA	International Association for the Evaluation of Educational Achievement
I/E model	Internal/External Frame of Reference Model
MIREC	Mary Immaculate Research Ethics Committee
MLR	Robust Maximum Likelihood Estimator
NCCA	National Council for Curriculum and Assessment
NEPS	National Educational Psychological Service
POI	Person-Object-Conception of Interest
RMSEA	Root Mean Square Error of Approximation
SD	Standard Deviation
SDT	Self-Determination Theory
SE	Standard Error
SRMR	Standardized Root Mean Square Residual
TIMSS	Trends in International Mathematics and Science Study
TLI	Tucker–Lewis Index

PIRLS	Progress in International Reading Literacy Study
WoE	Weight of Evidence

Glossary of Terms

Academic self-concept (ASC): ASC is postulated to be a multidimensional construct and refers to an individual's beliefs about their ability to perform, learn and achieve in particular academic domains (Marsh, 1986, 1990a, 1990b; Marsh et al., 1988; Marsh & Craven, 2006). Thus, an individual may express a higher or lower level of ASC in different academic subjects or areas, such as mathematics, English, foreign languages, science, etc.

Big-fish little-pond effect (BFLPE): This theory posits that pupils' ASCs are influenced by comparisons which pupils make between their performance and that of their peers (Marsh, 1987; Marsh & Parker, 1984). Hence, pupils who achieve to a similar objective standard will have more positive/negative ASCs based on the overall standard of achievement within their learning environment; with pupils in higher-achieving environments presenting with a relatively lower level of ASC and pupils in lower-achieving environments presenting with a relatively higher level of ASC.

Combined model: First proposed by Chiu (2012), the combined model incorporates the theoretical underpinnings and predictions of both the BFLPE and the internal/external frame of reference model into a single framework. The predictions of this model can be seen in Figure 5.

Dimensional comparisons: This term is used to denote comparisons which pupils make between their own, individual performances in different academic domains, such as their performance in English compared to their performance in mathematics.

Doubly-manifest model (this may also be referred to as a 'doubly-manifest approach'): A doubly-manifest model is one in which measures/variables within the model are construed as manifest variables. As defined in the 'manifest variable' section of this glossary, manifest variables, such as a person's height, can be objectively measured. Thus, they contain one measurement score per variable. In multi-level doubly-manifest models, level two, or group scores, are computed via simple aggregation of participants' level one scores. For example, considering an individual measure such as an individual's height, the level two or group score (e.g. a class) would be computed via simple aggregation of participants' individual scores (e.g. class-average height).

The combined model (see Figure 5), which is the focus of the literature review (see Chapter 2) and which informed the empirical paper (see Chapter 3), typically contains two

measures/variables which are ASC and achievement. When construed as a doubly-manifest model, both ASC and achievement measures would be treated as manifest variables. Thus, class-average achievement would be computed via simple aggregation of pupils' individual scores. Likewise, class-level ASC would also be computed via simple aggregation (e.g. the average ASC). The current analysis employed a latent-manifest approach (see 'latent-manifest model' within the current glossary of terms).

HOUWGT: The current study used data from an international database, notably the TIMSS and PIRLS 2011 Fourth Grade Combined International Database (Foy, 2013; TIMSS and PIRLS International Study Centre, 2019). This database contains sample weights. HOUWGT, also known as the student house weight, is the sample weight which was used in the current research.

A sample weight corrects for bias which may be present within a data sample. Specifically, this bias refers to how the composition of the collected data sample may differ from the composition of the actual target population. Outlining the rationale for sample weights within large scale datasets such as TIMSS and PIRLS, Rutkowski et al. (2010) noted that sample schools, pupils, or teachers may be selected with unequal probabilities. For instance, the sample may contain more or less pupils of a particular sex, or pupils from a particular socio-economic background, than would be actually representative of the target population. As explained by Rutkowski et al. (2010), sample weights essentially provide a value for each participant which reflects their "actual proportional occurrence in the population" (p.143). Thus, when employed within an analysis, this ensures that particular participants or groups are not over or under represented within the analysis, and that they are accurately represented in line with the target population. If certain participants or groups were over or under estimated within the data sample, this could lead to inaccurate results as the composition of the data sample would not be typical of the target population. Accordingly, the purpose of using the sample weights is to ensure that the study findings accurately reflect the target population.

HOUWGT is a student-level sample weight (Foy, 2013). HOUWGT is a transformation of another student-level weighting variable, TOTWGT. TOTWGT, also known as the total student weight, "sums to the student population size in each country" (Foy, 2013, p. 16). HOUWGT "ensures that the weighted sample corresponds to the actual sample size in each country" (Foy, 2013, p. 16). Thus, while TOTWGT takes the population size at the country

level into account, HOUWGT additionally weights for the sample size within each education system. Further information on why this sample weight was chosen can be found in section 4.3.2.4.

Internal/external frame of reference model (I/E model): This theory proposes that pupils' ASCs are shaped by comparisons which they make between their relative performances in different academic domains, such that individual achievement in one area will be positively related to a pupil's performance in that area but negatively related to a pupil's performance in a contrasting area, as pupils compare their relative performances across different academic domains when constructing their domain-specific ASCs (Marsh, 1986). For example, according to the I/E model, a student's ASC in mathematics would be positively affected by their performance in mathematics, but negatively affected by their performance in a different subject, such as English.

Intrinsic motivation: In the present study, intrinsic motivation aligns with the conceptualisation of intrinsic motivation outlined by Deci and Ryan (1985). Thus, it is a force which drives pupils' behaviour, and pupils who are intrinsically motivated to learn demonstrate greater enjoyment and interest towards learning.

Latent construct: A latent construct pertains to a construct which cannot be objectively measured. Examples include pupils' ASCs or their intrinsic motivation. As these constructs cannot be measured objectively, they are typically measured using scales which contain item indicators. Pupils' responses to these item indicators provide insight into their overall level of the construct.

Latent-manifest model (this may also be referred to as a 'latent-manifest approach'): Key statistical models in the current thesis, such as the BFLPE (see Figure 3) and the combined model (see Figure 5), typically contain two types of measures/variables such as ASC and academic achievement. In the context of these models, a latent-manifest model is one in which one of the measurement/variable types (e.g. ASC scales) is construed as a latent variable, and the other measurement/variable type is construed as a manifest variable. Therefore, for the latent variable, the scale items are input as indicators of the latent construct within the model and thus the indicators provide insight into pupils' overall level or score of the variable. While for the manifest variable, the overall, or total, objective scores which pupils obtain on the relevant measure (e.g. height or the overall score from an achievement test) are used within the model. For the level two component in multilevel latent-manifest

models, level two (or 'group') scores are computed via simple aggregation of participants' level one manifest scores. For example, considering an individual measure such as an individual's achievement score, the level two or group score (e.g. a class) would be computed via simple aggregation of participants' individual scores (e.g. class-average achievement).

Manifest variable: Unlike latent constructs, manifest constructs can be objectively measured. Examples include a person's height or weight.

Multilevel structural equation modelling: Structural equation modelling essentially seeks to model causal relationships between variables (Field, 2000). In an introductory article on structural equation modelling, Hox and Bechger (1999) noted that structural equation modelling may be understood as a "combination of factor analysis and regression or path analysis" (p. 1). Field (2000) explained that structural equation models comprise of two components, a measurement component/model and a structural component/model. In the measurement component/model, latent constructs and indicators of these constructs are assigned. This part of the structural equation model may be considered akin to factor analysis. In the structural part of the structural equation model, relationships between the latent constructs (that are measured in the measurement part of the structural equation model) and other variables within the structural model (which are not indicators of other latent constructs) are specified. Thus, structural equation modelling facilitates exploration of relationships between latent variables and other variables. Examples of structural equation models which are pertinent to the current research can be seen in Figure 3, Figure 4, Figure 5, and Figure 7. These models were proposed on the basis of prior research and theory, and the predictions of the models can be tested by specifying and computing the models.

Multilevel modelling refers to statistical models which include clustered, or hierarchical, data (Field, 2018). For example, pupils may be clustered within classes, and classes may be clustered within schools. This particular example depicts a three-level hierarchy. The levels within the hierarchy are commonly referred to as level one, level two, level three etc. In this case, level one would refer to the pupils, level two would refer to the classes, and level three would refer to the schools. Statistically, multilevel modelling differs from conventional linear modelling as it takes within-cluster dependence into account by explicitly modelling a given clustering variable (e.g. class) as a coefficient in a linear equation.

In the current thesis, the structural equation models that are depicted in Figure 3, Figure 5 and Figure 7, are multilevel structural equation models. This is because they are structural

equation models which contain clustered data, specifically students are clustered within classes. Thus, the models have a two-level hierarchy, with pupils (level one) nested within classes (level two). The terms, level one and level two, are used in the current thesis to describe the levels within the hierarchical models. Therefore, level one achievement would refer to individual achievement and level two achievement would refer to class-average achievement.

Reciprocal effects model: As outlined by Marsh et al. (2018) and Seaton et al. (2015), the reciprocal effects model advocates that there is a reciprocal relationship between ASC and academic achievement, such that pupils' prior achievement impacts their future ASCs, and pupils' prior ASCs impacts their future academic achievement. Unlike the BFLPE, the I/E model, and the combined model, which are typically investigated using cross-sectional study designs, the reciprocal effects model is typically investigated through longitudinal study designs. Research has supported the reciprocal effects model; however, results have indicated that the effect of prior academic achievement on pupils' ASCs is greater in size than the impact of pupils' ASCs on their future academic achievement (Marsh et al., 2018; Seaton et al., 2015; Valentine & DuBois, 2004).

Social comparisons: Social comparisons refer to comparisons which pupils make between their performance and that of their peers or classmates.

Subject-interest: In the present paper, subject-interest aligns with the person-object-conception of interest theory outlined by Krapp (2002). Essentially, this theory proposes that interest is characterised by the relationship between a person and an object of interest. This relationship embodies both an emotional aspect, such as the enjoyment a person feels when interacting with the object of interest, and a value aspect, notably the personal value which one attributes to the object of interest. Therefore, subject-interest in the current study pertains to both the enjoyment and personal value which pupils attribute to a particular academic subject.

Chapter One: Introduction

In this chapter, an overview of the thesis will be provided. This will commence with an account of factors which influenced my decision to study motivational constructs. Next, key theories and motivational constructs which informed the research will be outlined. This will include an overview of the nature and background of self-determination theory (SDT; Deci & Ryan 1985; Ryan & Deci, 2017), the person-object-conception of interest (POI) theory (Krapp, 2002), and academic self-concept (ASC). Three theories which account for factors which may influence pupils' ASCs will also be introduced. Following this, the current research, which is documented in the empirical paper, will be discussed, and the rationale for the research in line with the aforementioned theories and motivational constructs will be outlined. Subsequently the research process, and in particular the influence of the COVID-19 pandemic on the current research, will be summarised. Epistemological considerations and theoretical perspectives which informed the research will also be considered. To conclude, the overall structure of the subsequent thesis chapters will be outlined.

1.1. Rationale for Choosing to Study Motivational Constructs

In choosing a topic for the current thesis, I first reflected on my own personal experiences, interests, and professional practice within the field of education and child psychology to date. While working as a primary school teacher, I was particularly interested in motivational factors or constructs which may be implemented to enhance learning. Fundamentally, I believed that these factors had an essential role to play in optimising pupils' educational gains and achievements, beyond what purely pedagogical teaching approaches which focus solely on academic content could offer. From my own practice I became acquainted with the work of Carol Dweck (1999, 2006, 2014). Dweck (1999, 2006) essentially proposed that pupils' thinking regarding their abilities influences their learning behaviour. Specifically, she proposed that pupils with a 'growth mindset' believe that their academic abilities are malleable, and can be improved if they put sufficient practice and effort into learning. In contrast, pupils with a 'fixed mindset' believe that their overall level of ability is largely set or pre-determined, and that there is little they can do to change this (1999, 2006). In line with Dweck's (1999, 2006) theory, pupils who possess a 'growth mindset' are more inclined to engage in and demonstrate more persistence when faced with new learning tasks. Indeed, Dweck's work in this area has been described as "one of the most influential motivational theories in education" (Dai & Rinn, 2008, p. 296). I personally found

Dweck's (1999, 2006) work to be beneficial in my classroom practice. Through fostering a growth mindset among the pupils with whom I worked, I believe that this encouraged them to believe in themselves and their learning ability, and accordingly to focus on and engage in classroom learning activities at a deeper level. From a practical perspective, I was intrigued to experience how knowledge of this motivational construct could be used to promote the development of positive dispositions towards learning and to enhance pupils' progress.

I also became aware that the importance of capitalising on motivational factors to promote pupils' development is increasingly prevalent in recent policy and guidelines, such as recent Irish curricula set out by the National Council for Curriculum and Assessment (NCCA; 2018, 2019), best practice teaching guidelines set out by the National Educational Psychological Service (NEPS; 2019, 2020), and wellbeing guidelines (Government of Ireland, 2019). Indeed, these documents have emphasised that fostering positive attitudes and self-beliefs towards learning are critical for optimising educational outcomes for pupils and for promoting pupils' mental wellbeing. Therefore, given both my personal interest in the area and the increasing significance of motivational factors within current practice-related literature, I was keen to discover more about motivational constructs which may be influential on pupils' learning behaviour and progress. Hence, when I initially embarked on choosing a research topic, I decided to explore this area further.

Through engaging with relevant literature pertaining to motivation and learning, I learned about many theories which provide conceptual frameworks for understanding motivational constructs and factors which influence pupils' motivation. These theories were critical to informing the current research and are referenced throughout the thesis. Thus, an understanding of these theories and constructs provides deeper insight into the core research questions. The theories and constructs, and how they are connected to and related to learning, are documented below.

1.2. Motivational Constructs and Theories That Informed the Current Thesis

1.2.1. SDT: A Theoretical Framework for Understanding Human Motivation.

SDT is an empirically-based theory of human motivation and psychological wellness (Ryan & Deci, 2017; 2020). Drawing on literature and research from a range of psychological perspectives, the central tenets of SDT were first synthesised in Deci and Ryan's (1985) seminal book *Intrinsic Motivation and Self-Determination in Human Behavior*. Since then, the theory has been continuously developed and refined, and its practical application to a

range of areas of human life, such as education, work, sport, and health has been noted (Ryan & Deci, 2017, 2020). Following Deci and Ryan's (1985) work, the authors have retained a pivotal role in researching and advancing SDT. Synthesising a wealth of research and literature which emerged since their original work (Deci & Ryan, 1985), Ryan and Deci (2017) provided a comprehensive and updated conceptualisation of SDT.

1.2.1.1. Types of Motivation within SDT. A key feature of SDT is that it distinguishes between different types of motivation; specifically, these are intrinsic motivation, four types of extrinsic motivation (see Figure 1), and amotivation (Ryan & Deci, 2000a, 2017). As depicted in Figure 1, these types of motivation may be considered on a continuum from 'non-self-determined' or 'controlled', to 'self-determined' or 'autonomous'. Essentially, 'non-self-determined' or 'controlled' motivation refers to motivation which is determined, contingent on, or 'controlled' by external factors; whereas 'self-determined' or 'autonomous' motivation is not determined by external factors, but by a person's own initiative and free will (Ryan & Deci, 2000a, 2017). SDT posits that more 'self-determined' or 'autonomous' forms of motivation result in more favourable outcomes, such as enabling a person to thrive, to reach their full potential, and to experience heightened psychological wellbeing (Ryan & Deci, 2017). Further information on the conceptualisation of intrinsic motivation, extrinsic motivation, and amotivation within SDT is documented below.

1.2.1.1.1. Intrinsic motivation. As indicated in Figure 1, intrinsic motivation is the most self-determined or autonomous type of motivation. As outlined by Deci and Ryan (1985) and Ryan and Deci (2000a, 2017, 2020), when an individual is intrinsically motivated, they engage in an activity or pursuit as they find it be to inherently satisfying, interesting and enjoyable. Importantly, a distinctive feature of intrinsic motivation is that the positive affect which an individual experiences towards engaging in the activity or topic of interest is not contingent on separate external events, such as personal gains (e.g. eating well to achieve a secondary personal weight loss goal), rewards or punishments. A classic example of intrinsic motivation provided by Deci and Ryan (1985), is that of children engaging in play. While there may be secondary benefits associated with play, such as personal growth and learning, children spontaneously engage in self-directed play out of inherent interest, enjoyment and curiosity. Deci and Ryan (1985) noted that adults can likewise engage in similar behaviours, such as deciding to learn an instrument or to learn about history due to inherent curiosity. Considering the context of learning and education, SDT advocates, and research has shown, that pupils who are intrinsically motivated demonstrate enhanced outcomes in terms of their

engagement, performance, academic achievement, and psychological wellbeing (Froiland & Worrell, 2016; Ryan & Deci, 2017, 2020; Taylor et al., 2014). The concept of intrinsic motivation is central to the research reported in Chapter 3, the empirical paper, as this paper seeks to investigate the relevance of comparison processes on pupils' intrinsic motivation.

1.2.1.1.2. Extrinsic motivation. In contrast to intrinsic motivation, extrinsic motivation refers to motivation which is driven by separate, external factors or events, such as rewards or punishments for engaging in a particular activity/behaviour (Deci & Ryan, 1985; Ryan & Deci, 2000a, 2017, 2020). Thus, when an individual is extrinsically motivated, they engage in a behaviour as they believe that it will bring about favourable external consequences. Ryan and Deci (2000a, 2017, 2020) documented that SDT conceptualises four different types of extrinsic motivation. These are external regulation, introjected regulation, identified regulation and integrated regulation.

As outlined by Ryan and Deci (2000a, 2017, 2020), the four types of extrinsic motivation differ in terms of how non-self-determined/controlled or self-determined/autonomous the behaviour is (see Figure 1). The least autonomous type of extrinsic motivation is external regulation, and with this type of motivation an individual's behaviour is 'controlled' or contingent on external demands which need to be met to gain rewards or to avoid punishment; hence in responding to external demands the person feels that they have little control over their behaviour. Introjected regulation is also deemed to be a largely controlled form of extrinsic motivation and this type of motivation is commonly associated with being contingent on self-esteem, self-worth and avoiding negative feelings such as fear, anxiety and guilt. Accordingly, the external 'reward' for following through with internal demands or values (e.g. 'I must', 'I should' etc.) is to preserve one's ego or identity, whereas the punishment for not following through may be feelings of guilt or shame.

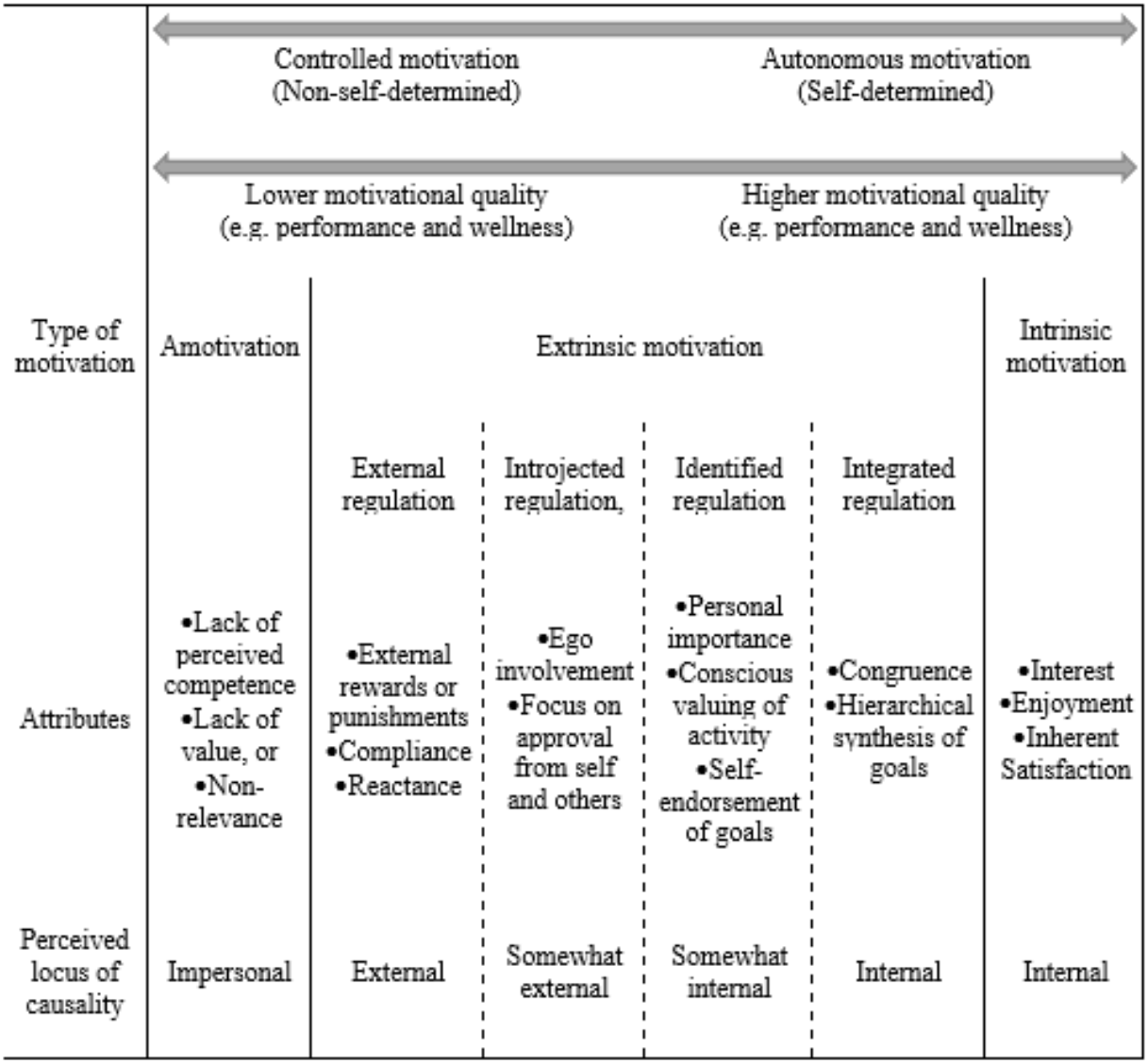
Within SDT's continuum of motivation (Ryan & Deci, 2000a, 2017, 2020), identified regulation and integrated regulation are more autonomous and internalised types of extrinsic motivation, and with these types of motivation behaviour is driven by goals and values which are believed to be worthwhile and of personal importance (e.g. eating well to maintain a personal goal of being healthy). Essentially, with these types of motivation there are no external demands on the person to perform or act in a certain way, it is the personal value of potential extrinsic rewards which motivates the person (e.g. the desire to be healthy). Integrated regulation is regarded as the most autonomous form of extrinsic motivation. While

identified regulation stems from an individual valuing certain goals and considering them to be personally worthwhile; with integrated regulation the motivational goals and values behind the behaviour are more internalised and considered to be assimilated, or congruent, with the person's overall sense of self and values. Integrated regulation may be viewed as most akin to intrinsic motivation due to the high level of autonomy associated with both types of motivation. Hence, in line with SDT, the type of extrinsic motivation that will result in optimum learning outcomes for pupils is integrated regulation.

1.2.1.1.3. Amotivation. Within the SDT framework, Ryan and Deci (1985, 2000a, 2017, 2020) described one other state known as amotivation. Essentially, the term refers to a state in which an individual is not motivated. Ryan and Deci (2017) propose that this lack of motivation may stem from a lack of perceived competence, a lack of value or interest towards the behaviour/activity, or a lack of perceived incentives for completing particular behaviours.

Figure 1

Self-Determination Theory's Continuum of Motivation



Note. Figure adapted from the Center for Self-Determination Theory (2017).

1.2.1.2. SDT: Three Basic Psychological Needs. A fundamental assumption underlying SDT is that people have an inherent disposition towards learning, positive psychological growth, and in line with SDT's continuum of motivation, more autonomous forms of motivation (Ryan & Deci, 2017, 2020). However, this disposition can be facilitated or hindered by the presence or absence of three basic psychological needs. These three needs are autonomy, competence, and relatedness. Each of these needs were outlined by Ryan and Deci (2017), and the explanation of each basic need which is presented here is informed by

their work. Thus, in line with Ryan and Deci (2017), autonomy pertains to an individual's need to be in control of their own initiative, growth and actions. Hence, it may be supported by having free rein to choose and engage in activities of interest, and it may be hindered by having to comply with an external regime. Competence refers to the need for people to feel that they can function, succeed, grow, and develop mastery in their environment. It may be enhanced by completing developmentally appropriate challenges and/or by receiving unexpected positive feedback for performance on a challenging task. In contrast, it may be hindered by receiving negative feedback. The third need, relatedness, pertains to the need to feel a sense of belonging and connectedness with others, and this encompasses both feeling valued and cared for by others and feeling affection and care towards others. The importance of supporting pupils' basic psychological needs in school for fostering intrinsic motivation towards learning, and for optimising pupils' development and learning, has been emphasised within recent research literature which considered how SDT may be applied to education/school systems (e.g. Niemiec & Ryan, 2009; Reeve, 2012; Ryan & Deci, 2017, 2020).

1.2.2. The Person-Object-Conception of Interest Theory

Drawing on his earlier work (e.g. Krapp 1993; Krapp, 1999), and noting a perceived lack of a coherent conceptualisation of interest within educational literature and research, Krapp (2002) proposed the POI theory as a theoretical framework which may be used to inform future understanding and research pertaining to pupils' interests. In acknowledging the potential benefit of this theoretical framework for future research and practice, Krapp (2002) recognised that SDT (e.g. Deci & Ryan, 1985) similarly provided an advantageous framework for conceptualising motivational constructs.

Krapp (2002) differentiates between individual interest and situational interest. Individual interest pertains to a person's natural disposition or preference towards the object/subject of interest, and is considered to be a relatively stable trait. Conversely, situational interest is not proposed to be a stable trait, but refers to interest which arises from experiencing an object/subject of interest in favourable and intriguing circumstances. In line with Krapp's (2002) theory, literature and research has advocated that both individual interest and situational interest can result in valuable learning and performance outcomes, such as greater attention when engaging in interest-related activities, greater persistence when faced with challenges, enhanced performance, and enhanced learning (Alexander & Jetton, 1996;

Durik, 2017; Hidi, 1990; Hidi & Renninger 2006; McDaniel et al., 2000; Renninger et al., 2002; Renninger & Hidi, 2002). Reflecting on both individual interest and situational interest, the POI theory (Krapp, 2002) proposes that interest may be understood as a relationship between a person and the object/subject of interest. In essence, a person's prior experience and individual interest influences their disposition towards engaging with the object/subject of interest, and simultaneously the environment in which one experiences the object/subject of interest (i.e. situational interest) can impact the development a person's individual interest.

POI theory (Krapp, 2002) also specifies that interest encompasses both feeling and value-related aspects. Specifically, the feeling aspect pertains to the individual's affect (e.g. feelings of enjoyment and satisfaction) towards the object/subject of interest, whereas the value-related aspect pertains to the personal value or personal importance which the individual attributes to the goals or content that are related to the object/subject of interest. Accordingly, the value-related aspect of interest, as defined within the POI theory, may be considered to be on the extrinsic motivation continuum of SDT (e.g. Ryan & Deci, 2017). POI theory has underpinned previous research which investigated pupils' subject-interests (e.g. Schurtz et al., 2014) and was considered when deciphering the theoretical underpinning for the current research. Subsequently, and as outlined in section 1.3.1., it was decided that the current research would explore pupils' intrinsic motivation as conceptualised within SDT.

1.2.3. Academic Self-Concept

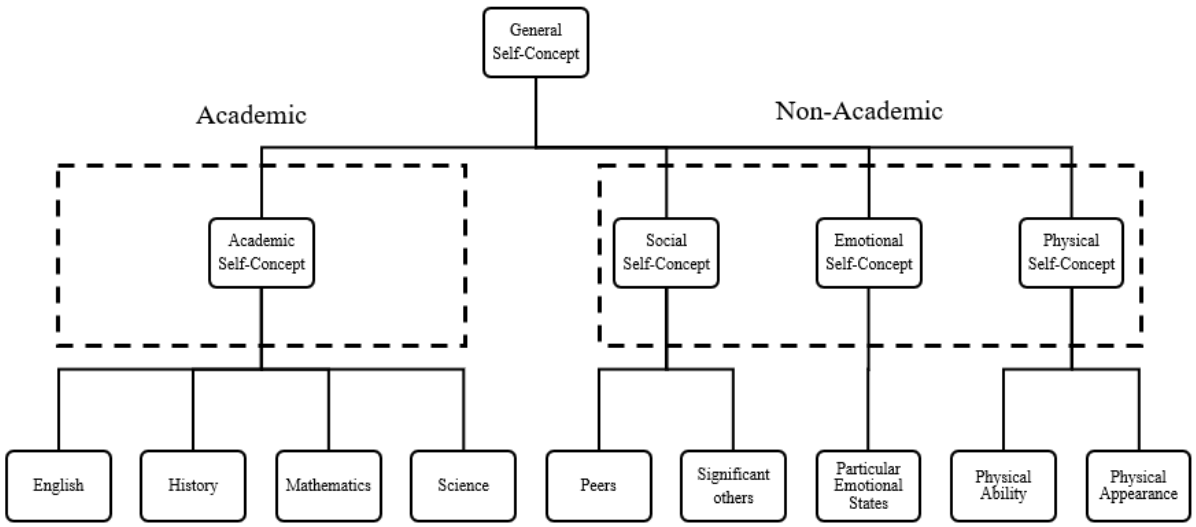
ASC is a motivational construct which is pertinent to the present research. The conceptualisation of ASC that is presented here is particularly relevant to, and is a key component of, the models pertaining to comparison processes which are briefly outlined in subsequent sections. This conceptualisation of ASC was established from the work of Shavelson et al. (1976) and Marsh and Shavelson (1985).

Essentially, ASC refers to an individual's beliefs regarding their academic learning and performance abilities (Marsh, 1990a, 1990b; Marsh et al., 1988; Marsh & Craven, 2006; Marsh and Shavelson, 1985; Shavelson et al. 1976). Shavelson et al. (1976) proposed that ASC is a multidimensional construct, such that pupils' ASCs may be higher or lower in different academic areas. For example, a pupils' mathematics self-concept may be higher than their reading self-concept, or vice versa. In line with Shavelson et al.'s (1976) model, which was later modified by Marsh and Shavelson (1985), ASC is conceptualised within an overall global/general model of self-concept which is hierarchical in nature. As can be seen in

Figure 2, this model places global/general self-concept at the top of the hierarchical model. At the next level, the model distinguishes academic and non-academic self-concepts. As depicted in Figure 2, ASC may be further sub-categorised into different academic areas. Indeed, this empirically-based model has been highly influential in underpinning a comprehensive and widely accepted conceptualisation of ASC which continues to inform ASC research over four decades since it was originally proposed (Kadir & Yeung, 2016; Marsh, 1990a, 1990b; Marsh et al., 1988; Marsh & Craven, 2006). Consistent with this model of ASC, studies reviewed for the current thesis (of note those documented in Chapter 2) investigated domain-specific ASCs (e.g. English self-concept, mathematics self-concept, etc.).

Figure 2

The Multidimensional and Hierarchical Nature of Self-Concept



Note. Figure adapted from Shavelson et al. (1976) to demonstrate the multidimensional hierarchical structure of Shavelson et al.'s (1976) and Marsh and Shavelson's (1985) self-concept model. As can be seen in Figure 2, ASC is distinct from other non-academic types of self-concept. Academic self-concept is also sub-categorised into different academic domains.

Since the proposition of Shavelson et al.'s (1976) and Marsh and Shavelson's (1985) conceptualisation of ASC, a plethora of research has emphasised that higher ASCs are linked with a range of advantageous learning and educational outcomes (Kadir & Yeung, 2016). For example, literature has documented that more positive ASCs are associated with higher academic achievement, pupils' future educational and career aspirations, and increased

engagement, effort and persistence from pupils when learning (e.g. Eccles, 2009; Fang et al., 2018; Guo et al., 2015; Kadir & Yeung, 2016; Marsh & Craven, 2006; Marsh & Yeung, 1997; Yeung, 2011; Yeung, Kuppan, Foong, et al., 2010; Yeung, Kuppan, Kadir, et al., 2010). Given the range of valuable educational outcomes that are associated with higher ASCs, it is important to consider how more positive ASCs may be fostered. Linking this to Dweck's (1999, 2006) conceptualisations of fixed and growth mindsets, it is interesting to note that research has shown that pupils with a growth mindset (as opposed to a fixed mindset) demonstrate more positive ASCs, and thus encouraging pupils to develop a growth mindset may in turn enhance their ASCs (Carvalho & Skipper, 2020; Ommundsen et al., 2005).

Another distinction which has emerged within the literature since Shavelson et al.'s (1976) and Marsh and Shavelson's (1985) framework for understanding ASC is the separation of competency and affective components on scales which measure pupils' ASCs. This stance is supported by both theoretical arguments and empirical studies, notably the results of confirmatory factor analyses which have demonstrated that items which pertain to pupils' affect towards particular subject areas are distinctive from items in which pupils rate their competencies (Arens et al., 2011; Arens et al., 2013; Marsh et al., 1999). Indeed, in earlier studies investigating ASC, measurement scales may have commonly included items which pertained to pupils' affect towards learning in a particular subject area, such as 'I like [name of subject]'. However, literature has since specified that measurement scales focus on either affective or competency components. Accordingly, recent research, and in particular research pertaining to the theoretical models relating to comparison processes that are explored in the current study, has advocated excluding affective items and to focus solely on competency items when investigating pupils' ASCs (e.g. '[name of subject] is easy for me'; e.g. Guo et al., 2018; Lohbeck & Möller, 2017; Marsh et al., 2009; Pinxten et al., 2015). Therefore, in line with the aforementioned literature and definition of ASC, in the current study ASC refers to an individual's belief about their ability to perform, learn and achieve in a particular academic area. It does not pertain to affective components such as liking or having interest in a particular subject, as these aspects would be more indicative of intrinsic motivation (Deci & Ryan, 1985; Ryan & Deci, 2017) and/or interest (Krapp, 2002) as outlined above.

1.2.4. Theories Pertaining to Comparison Processes and Pupils' ASCs

Since the mid-1980s, literature has asserted that pupils' ASCs are shaped by comparisons which pupils make between their performance and that of their peers (social comparisons), and comparisons which pupils make between their own individual performances in different academic domains (dimensional comparisons). This literature is underpinned by empirically supported theoretical models. Specifically, these models are the big-fish-little-pond effect (BFLPE; Marsh, 1987; Marsh & Parker, 1984), the internal/external frame of reference model (I/E model; Marsh, 1986), and the combined model (Chiu, 2012). While a brief outline of each model will be provided here, each model is critiqued in further detail in Chapter 2.

1.2.4.1. The Big-Fish-Little-Pond Effect. Essentially, the big-fish-little-pond effect (BFLPE; Marsh, 1987; Marsh & Parker, 1984) theory posits that pupils' subject-specific ASCs are influenced by comparisons which pupils make between their performance in a particular subject and that of their peers. Consistent with this theory, research has found that class-average achievement is negatively correlated with pupils' ASCs (Fang et al., 2018; Marsh & Hau, 2003).

1.2.4.2. The Internal/External Frame of Reference Model. The internal/external frame of reference model (I/E model; Marsh, 1986) posits that pupils' ASCs are shaped by comparisons which pupils make between their relative performances in different academic domains, such that individual achievement in one area will be positively related to a pupil's ASC in that area but negatively related to a pupil's ASC in a contrasting area, as pupils compare their relative performances across different academic domains when constructing their domain-specific ASCs. For example, according to the I/E model, a student's ASC in mathematics would be positively affected by their own performance in mathematics, but negatively affected by their performance in a different subject, such as English.

1.2.4.3. The Combined Model. As the name suggests, the combined model (Chiu, 2012) built on the BFLPE and the I/E model by integrating both models into a single theoretical framework. In Chapter 2 a literature review pertaining to the combined model is presented.

1.3. The Current Research

As documented in Chapter 2, the BFLPE, the I/E model, and the combined model are supported by empirical evidence when tested in relation to pupils' ASCs, thus indicating that

pupils' ASCs are influenced by social and dimensional comparison processes (e.g. Kavanagh, 2019; Pinxten et al., 2015). However, limited research has investigated the relevance of the models to other motivational constructs. Accordingly, the current research, which is documented in the empirical paper, investigated the relevance of these models to Irish fourth-class pupils' intrinsic motivation for mathematics and reading. An additional model was also proposed and investigated. This was an extension of the combined model which was tested in relation to pupils' intrinsic motivation. This model hypothesised that within the combined model, any effects on pupils' intrinsic motivation may be mediated by their ASCs.

It was reasoned that exploring comparison processes, ASC, and intrinsic motivation was a valuable area to explore in line with the recent Irish curricula (NCCA; 2018, 2019), best practice teaching guidelines (NEPS; 2019, 2020), and wellbeing guidelines (Government of Ireland, 2019). Of note, while these guidelines underscore the need to foster positive attitudes and self-beliefs towards learning, information regarding the potential impact of social and/or dimensional comparison processes on pupils' attitudes and self-beliefs, and how the potential impact of comparison processes may be overcome in practice, is largely absent from this literature. Hence, I decided that this was the research area which I wished to pursue for my thesis.

1.3.1. Rationale for Investigating Intrinsic Motivation

1.3.1.1. A Gap in Previous Research. Based on a literature review that is presented in Chapter 2, only one study to date investigated the relevance of the BFLPE, the I/E model, and the combined model to a motivational construct other than pupils' ASCs. This was completed by Schurtz et al. (2014) who investigated the relevance of the models to pupils' subject-interests. Their conceptualisation of subject-interest was in line with Krapp's (2002) POI theory. Hence, Schurtz et al.'s (2014) subject-interest measures pertained to both pupils' feeling or affect towards, and the personal value or importance which they placed on, learning mathematics and English as an additional language. In line with the conceptualisation of intrinsic motivation within SDT, the current study focused solely on exploring the models in relation to pupils' feelings or affect towards reading and mathematics. It was reasoned that this was necessary to explore as within the framework of SDT, the aspect of subject-interest relating to the personal value or personal importance which a pupil attributes to a particular subject is theoretically distinctive from the aspect which pertains to pupils' feelings or affect towards the subject. Of note, the aspect relating to

pupils' feelings or affect towards the subject aligns with intrinsic motivation, however, the aspect pertaining to the personal value or personal importance which pupils place on the subject falls within the extrinsic motivation continuum of SDT. Accordingly, the current study sought to investigate the relevance of the BFLPE, the I/E model, and the combined model to pupils' intrinsic motivation.

1.3.1.2. The relationship between ASC, Competence, and Intrinsic Motivation.

Deci and Ryan's (1985) and Ryan and Deci's (2017) conceptualisation of intrinsic motivation and the broader theoretical framework of SDT was also critical to informing and underpinning the proposed models in the present research. Of note, SDT advocates that the presence or absence of three basic psychological needs can enhance or thwart pupils' intrinsic motivation. These needs are autonomy, competence and relatedness. The basic psychological need of competence, which is particularly relevant to the current research, refers to an individual's (e.g. a pupil) need to feel that they are able to function, learn and succeed in their environment. In line with SDT, a higher sense of competence can enhance pupils' intrinsic motivation, whereas a lower sense of competence may thwart pupils' intrinsic motivation. Within the context of education and school systems, pupils' sense of competence may be considered to be closely related to the conceptualisation of pupils' ASCs, which essentially pertains to pupils' beliefs regarding their abilities to learn, succeed and perform academically (Marsh, 1990a, 1990b; Marsh et al., 1988; Marsh & Craven, 2006; Marsh and Shavelson, 1985; Shavelson et al. 1976). Therefore, consistent with SDT, it was hypothesised that if previous research demonstrated that the BFLPE, the I/E model, and the combined model are relevant to pupils' ASCs, the models may also be relevant to pupils' intrinsic motivation. In line with the theoretical framework underpinning the BFLPE, the I/E model, and the combined model, this may in turn indicate that social and dimensional comparison processes impact not only pupils' ASCs, but also their intrinsic motivation. Building on this link between ASC and intrinsic motivation, it was also posited that any effects on pupils' intrinsic motivation may be mediated by their ASCs.

1.4. The Impact of the COVID-19 Pandemic on the Direction of the Current Research

The direction of the current research was significantly impacted by the COVID-19 pandemic. Following completion of a systematic review of relevant literature relating to the combined model in July and August of 2019, which is documented in Chapter 2, a research proposal was created. However, this project was different to the research presented in the

empirical paper of this thesis. It sought to investigate the relevance of the BFLPE, the I/E model, and the combined model to Irish second-class pupils' ASCs in mathematics and reading. Ethical approval was sought and granted for this study in January 2020 (see Appendix A), and prior to the school closures which commenced on March 13th 2020, a pilot study was completed and official data collection had commenced.

Due to the school closures, data collection for original study with second-class pupils could not continue. Consequently, in consultation with my supervisors, I had to consider how the research would progress. After exploring a range of potential options, in line with available data sources and consideration of other pertinent areas for future research which emerged from the literature review, it was decided that the current research investigate the relevance of the BFLPE, the I/E model, and the combined model to another motivational construct, notably pupils' intrinsic motivation. This could be investigated in relation to Irish fourth-class pupils' intrinsic motivation for mathematics and reading using the 2011 Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS) combined international database. Although it was hoped that the research be conducted on primary empirical data, as research in this area is largely based on secondary analysis of large-scale survey datasets, this alternative was deemed superior to using the limited data which was collected prior to July 2020. Based on the number of classes and participants which could be analysed from the initial data collection, there was an insufficient number of participants in the sample to adequately address the research questions (Field, 2018).

This significant change in the direction to the research occurred during the Spring and Autumn semesters in 2020, at a stage when much groundwork for the intended research and analysis had already been configured. However, while this generated a number of unforeseen setbacks, the challenges also greatly enhanced my psychology and research skills. For example, I learned about different types of motivational constructs and theories, such as SDT (Deci & Ryan, 1985; Ryan & Deci, 2017) and the POI theory (Krapp, 2002). It also heightened my understanding of how large-scale, nationally representative studies are conducted and analysed. These skills will be advantageous in my future work, both as a practitioner and researcher. As a practitioner, the skills will be valuable when considering the findings from research which used similar large-scale complex survey data to inform my practice. As a researcher, I now have a more comprehensive understanding of how to

complete research analyses using large-scale complex survey datasets, making this a more viable option when considering undertaking research in the future.

1.5. Epistemological Considerations

The current research is embedded within a robust theoretical literature base (Marsh, 1986, 1987; Marsh & Parker, 1984; Chiu, 2012), of which the principal concepts were briefly detailed above, and are presented and analysed further in Chapter 2 of this thesis. The research sought to extend prior literature in this area by investigating whether the predictions of the BFLPE, the I/E model, and the combined model would be upheld when tested in relation to pupils' intrinsic motivation. Therefore, as the research sought to prove or disprove the significance of particular theories among a certain cohort of pupils, it was reasoned that the research may be justifiably situated within the postpositivist paradigm.

As outlined by Mertens (2015), the postpositivist paradigm proposes that while the reality of the social world may not be understood with certainty, there is one reality which may be understood within a certain degree of probability. The postpositivist paradigm also recognises that the individual thoughts, experiences and opinions of the researcher may impact observations and/or research findings, and accordingly it is emphasised that the researcher maintain an objective stance throughout the research process (Mertens, 2015). Drawing on experimental methods from the positivist paradigm, quantitative methodologies are primarily employed within the postpositivist paradigm (Mertens, 2015). Consistent with the postpositivist paradigm, the current study analysed quantitative, cross-sectional data from an established international database. Thus, the researchers own views and/or biases did not impact the observations.

1.6. Overall Structure of the Thesis.

The subsequent sections of the thesis are divided into three main parts. In the first part, a systematic literature review which was undertaken to synthesise previous research to date and identify areas for future research is documented. Following this, the empirical paper is presented. The empirical paper documents the research analyses which were completed to investigate the principal research questions, which essentially focused on investigating relationships between individual achievement and class-average achievement on pupils' intrinsic motivation for mathematics and reading. A discussion regarding the results of the analyses is also provided in this section. Subsequently, in the final chapter a critical review of the research that was undertaken is presented. This includes an overview of the research

process as a whole, and challenges and ethical considerations which were encountered and overcome. It also critiques the design, measures, and analysis techniques which were employed, along with considering the implications of the findings for research and practice. To conclude, an impact statement is presented, which underscores how findings from the current research may be maximised for public benefit both inside and outside of academia.

Chapter Two: Literature Review

2.1. Introduction

Motivational factors, such as the development of positive self-beliefs and interest towards learning, are critical for scaffolding effective learning (Baten et al., 2019; Dweck, 1999; Lee & Sue, 2015; Mitchell, 2014). Indeed, over the past two decades this has been underscored in international studies, which have consistently highlighted robust positive correlations between motivational constructs, such as pupils' confidence and the value they attribute to learning, and educational attainment (Mullis, 2007; Organisation for Economic Co-operation and Development, 2004, 2007, 2014). Accordingly, it is perhaps unsurprising that the development of positive attitudes towards learning are now established as key objectives in both national and international curricula (National Council for Curriculum and Assessment [NCCA], 2018, 2019; Mullis, Martin, Ruddock, et al., 2009; Mullis, Martin, Kennedy, et al., 2009; Mullis & Martin, 2017, 2019).

Given the importance of pupils' attitudes and beliefs for optimising learning, much literature has sought to establish a more comprehensive understanding of specific constructs which may be particularly pertinent to facilitating and enhancing learning. In addition, literature has detailed theoretical frameworks which may account for factors which impact pupils' motivation and learning behaviour. Two particular theories which are prominent to the current review are the big-fish-little-pond effect (BFLPE; Marsh, 1987; Marsh & Parker, 1984) and the internal/external frame of reference model (I/E model; Marsh, 1986). Fundamentally, these theories were posited to account for comparison processes, notably comparisons which pupils make between their performance and that of peers and/or their own performances in different subject areas, which may impact the development of pupils' academic self-concepts (ASCs). More recently these two theories have been combined into a single unified model, which for the purpose of the present review will be referred to as the combined model (Chiu, 2012).

The current review seeks to explore if the combined model is supported by empirical evidence. As the combined model was built upon a plethora of previous literature, a synopsis of this theory and research will first be outlined. Specifically, this will commence with an overview of the conceptualisation of ASC. A sound understanding of this is necessary in the context of outlining the subsequent theoretical frameworks, notably the BFLPE and I/E

model, as these models were initially proposed to account for factors which influence the development of pupils' ASCs (Marsh, 1986, 1987; Marsh & Parker, 1984). Consequently, the vast majority of research investigating these theories to date has focused on investigating the relevance of the models to pupils' ASCs (Schurtz et al., 2014). Following an overview of the BFLPE and I/E model, the defining features and rationale for the combined model is appraised.

To establish whether or not the combined model is supported by empirical evidence, a systematic search for relevant literature pertaining to the combined model was conducted to evaluate current empirical evidence investigating the theory. This was necessary as while the BFLPE and the I/E model have been widely studied and reviewed (e.g. Fang et al., 2018; Möller et al., 2009), the combined model is a more recent development and discourse examining the combined model is sparser. Following the overview of previous literature and the theoretical background underpinning the model, the process involved in completing the systematic search is documented and a critical review of the literature is presented. Consequently, findings are analysed, evaluated in line with relevant implications for theory and future practice, and potential areas for future research are outlined.

2.2. Overview of Previous Literature and Models.

2.2.1. Academic Self-Concept

ASC is postulated to be a multidimensional construct and refers to an individual's beliefs about their ability to perform, learn and achieve in particular academic domains (Marsh, 1986, 1990a, 1990b; Marsh et al., 1988; Marsh & Craven, 2006). Thus, an individual may express a higher or lower level of ASC in different academic subjects or areas, such as mathematics, English, foreign languages, science, etc. In addition to being positively correlated to academic achievement, ASCs can have a significant impact on pupils' future career aspirations, and the initiative, effort and persistence which pupils exert when learning new tasks (Eccles, 2009; Fang et al., 2018; Guo et al., 2018; Guo et al., 2015; Parker et al., 2013). Naturally, these traits are fundamental to maximising pupils' learning.

In earlier studies investigating ASC, measurement scales commonly included items which pertained to pupils' affect towards learning in a particular subject area, such as 'I like [name of subject]'. However, literature has since advocated that ASC measurement scales exclude affective items and pertain solely to pupils' perceptions regarding their abilities and competencies in a particular area, such as '[name of subject] is easy for me' (Lohbeck &

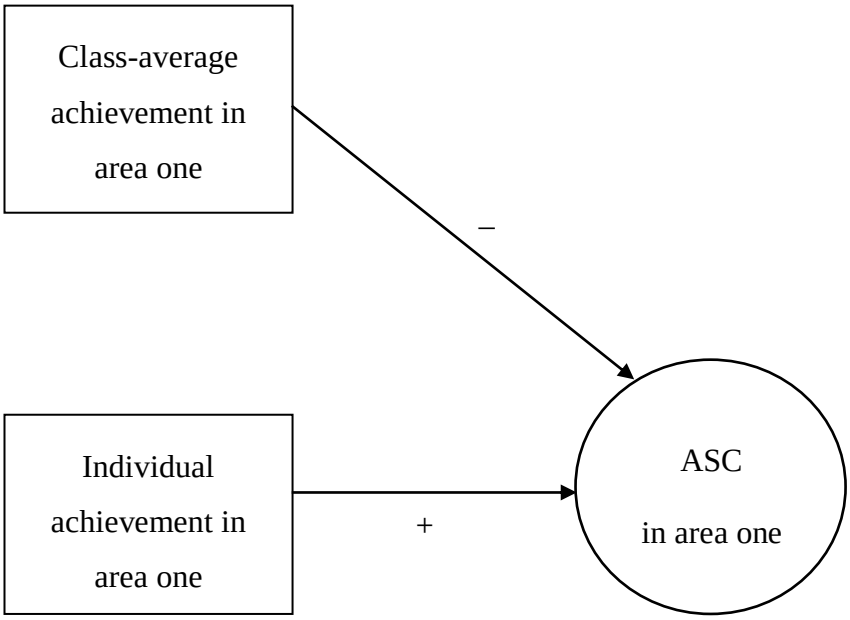
Möller, 2017; Marsh et al., 2009). Indeed, this stance is supported by both theoretical arguments and empirical studies, notably the results of confirmatory factor analyses which demonstrated that items which pertain to pupils' affect towards subject areas are distinctive from items in which pupils rate their abilities and/or competencies (Arens et al., 2013; Marsh et al., 1999).

2.2.2. Big-Fish-Little-Pond Effect

First proposed by Marsh and Parker, the BFLPE posits that pupils' ASCs are shaped by comparisons which they make between their performance and that of their peers (Marsh, 1987; Marsh & Parker, 1984; Thijs et al., 2010). Essentially, the performance of peers is used as a point of reference by pupils when they establish their own ASCs. Accordingly, pupils who achieve to a similar objective standard will have more positive/negative ASCs based on the overall standard of achievement within their learning environment; with pupils in higher-achieving environments presenting with a relatively lower level of ASC and pupils in lower-achieving environments presenting with a relatively higher level of ASC. In line with this theory, and as depicted in Figure 3, research has demonstrated that while individual achievement is positively correlated with pupils' ASCs in a particular academic domain, class-average achievement is negatively correlated with pupils' ASCs (Marsh et al., 2000). This robust finding has been generalised across countries and age ranges (Fang et al., 2018; Marsh & Hau, 2003). Conversely, a minority of exceptions have been reported within individual studies (Liou, 2014; Sung et al., 2014). As discussed by Fang et al. (2018), the prominence and size of the BFLPE is impacted by factors such as the subject domain, age and contextual factors within different countries and cultures.

Figure 3

Predictions of the Big-Fish-Little-Pond Effect



Note. Area one refers to a particular subject or academic domain, as ASC is domain-specific.

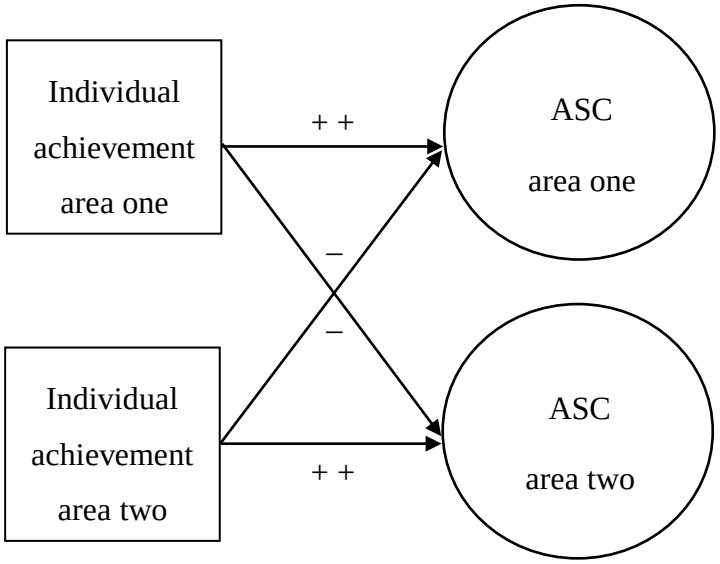
2.2.3. Internal/External Frame of Reference Model

The I/E model, as proposed by Marsh (1986), was initially posited to account for the observed discrepancy of high correlations between pupils' verbal and mathematic achievement, but lower correlations between pupils' verbal and mathematic ASCs. In accordance with this model, an individual's ASC is shaped by internal, or dimensional, comparisons which an individual makes between their performance in different academic domains, and external comparisons which an individual makes between their performance and perceived external environmental stimuli such as the performance of their peers and/or the feedback/grades they receive (Möller et al., 2009). As the model theorises that pupils' ASCs are shaped by comparisons which they make between their performances in different academic domains, research examining the model typically investigates academic domains that are considered to be distinctive of each other, such as English and mathematics (Chiu, 2012). Predictions of the I/E model can be seen in Figure 4. In accordance with this model, an individual's achievement in one domain has a positive effect on the individual's ASC in that domain; conversely, internal comparisons which the pupil makes between their individual performances in different academic areas results in achievement in one area having a

negative impact on the pupil's ASC in the contrasting area (e.g. pupils' individual mathematics achievement has a negative association with pupils' individual English ASC).

Figure 4

Predictions of the Internal/External Frame of Reference Model



Note. Area one refers to a particular subject or academic domain, as ASC is domain-specific.

A robust evidence base has found the predictions of the I/E model to be generalisable across countries and age ranges (Chiu, 2008; Marsh & Hau, 2004; Möller et al., 2009). Conversely, researchers have been critical of studies testing the I/E model, as they have typically focused solely on dimensional comparisons, as depicted in Figure 4 above, and did not investigate social comparisons at a class or school group level (Chiu, 2012; Parker et al., 2013; Pinxten et al., 2015). Noting this limitation, advancing research has commenced explicitly investigating social and dimensional comparisons simultaneously in line with the combined model, which is discussed below.

2.2.4. The Combined Model

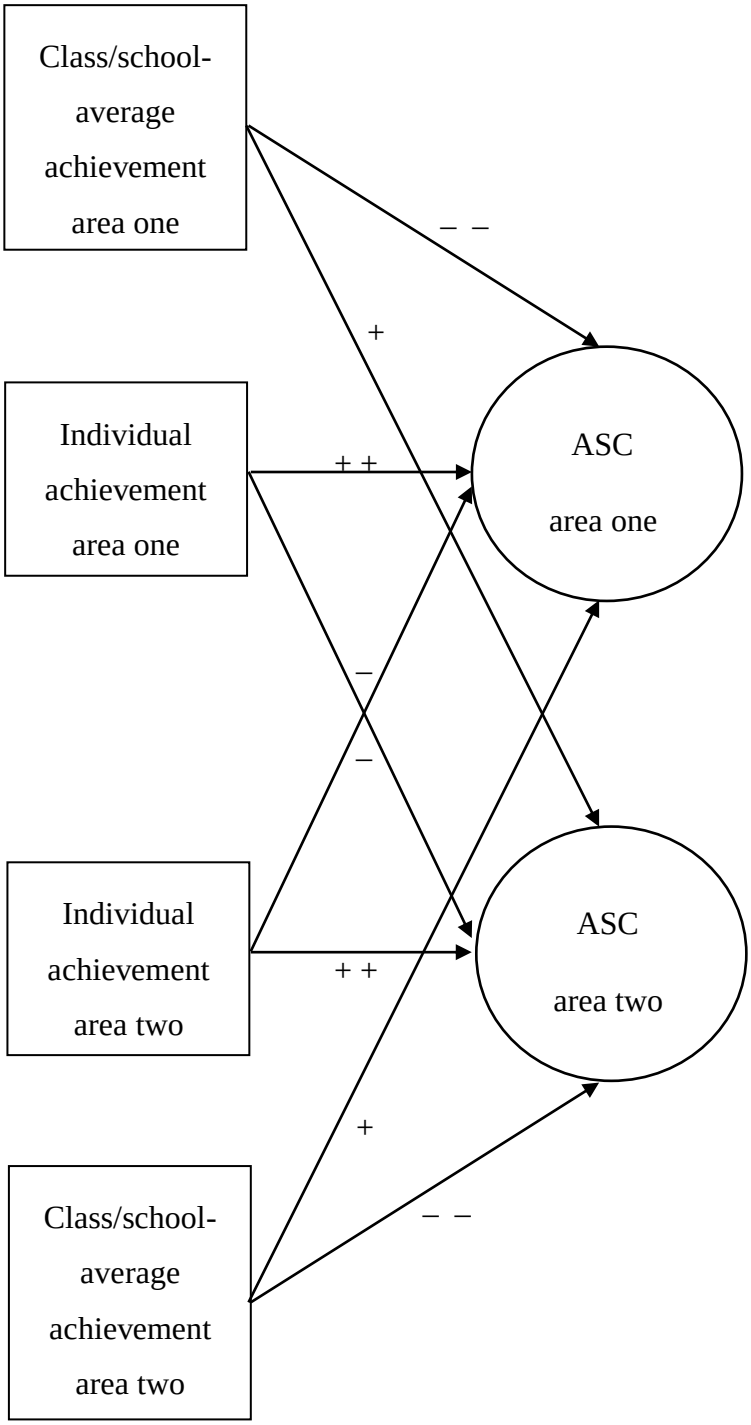
Chiu (2012) noted that both the BFLPE and the conventional I/E model were based on comparisons which individuals make between their individual performance, and associated frames of reference (i.e. social and dimensional comparisons). In line with the theoretical underpinning of the BFLPE, research investigating the BFLPE has principally focused on the impact of external, social comparisons on pupils' learning attitudes (Chiu, 2012; Schurtz et al, 2014). However, while it was theoretically proposed that the I/E model accounts for both

social and dimensional comparisons, research investigating the I/E model has principally investigated internal comparisons only (Parker et al., 2013). As documented by Chiu (2012), despite the similarities in the theoretical backgrounds for the BFLPE and I/E model, as both assert that pupils' ASCs are shaped by standards which pupils compare their performance to, these models have been investigated concurrently in two separate lines of research over the past three decades. Accordingly, Chiu (2012) posited that a theory which specifically conceptualises and investigates both social and dimensional comparison processes simultaneously was warranted. She reasoned that if both processes were still evident within the unified model, this would further support the independent contribution of both social and dimensional comparison processes to pupils' ASC formation. In addition, she advocated that integrating and testing the two theories simultaneously may further advance our conceptualisation and comprehension of comparisons processes which influence pupils' ASCs.

As outlined by Chiu (2012), the predictions of the combined model correspond with the predictions of the traditional I/E model and the BFLPE. Consistent with the I/E model, individual achievement in a particular domain will have a positive effect on pupils' ASC in the corresponding domain, but a smaller negative impact on pupils' ASC in the contrasting domain. In line with the BFLPE, class/school-average achievement in a particular domain will be negatively correlated with pupils' ASCs in the corresponding domain. In addition to these central predictions, the predictions of the combined model have been extended such that class/school-average achievement in one domain will have a positive effect on pupils' ASCs in a contrasting domain, but this effect will be smaller in size than the negative effect of the BFLPE (Parker et al., 2013; Pinxten et al., 2015). Parker et al. (2013) outlined that this prediction is in line with the conventional I/E model, as a positive/negative effect of achievement in one domain is hypothesised to have a contrasting effect on ASC in a contrasting domain. This prediction was not considered in prior research of the basic BFLPE or the I/E model, as studies exploring the impact of class/school-average achievement typically focused on only one academic domain (Chiu, 2012). A visual representation of all of the predictions of the combined model is presented in Figure 5.

Figure 5

Predictions of the Combined Internal/External Frame of Reference Model



Note. Area one and area two refer to particular subjects or academic domains, as ASC is domain-specific.

Considering the robust literature base supporting the conventional I/E model and the BFLPE, it is postulated that the combined model may further our conceptualisation and understanding of the development of ASC. However, despite the plethora of research that has investigated the BFLPE and the I/E model separately, research combining these models is a recent advancement (Chiu, 2012). There is of yet no consensus as to whether the predictions of both the I/E model and the BFLPE will both contribute as independent processes within the unified model. Studies have varied in terms of their support for this model (Lohbeck & Möller, 2017; Parker et al., 2013; Pinxten et al., 2015). Moreover, as the effects of the basic BFLPE and I/E models have been found to differ depending on factors such as culture, age and academic subject domain (Fang et al., 2018; Möller et al., 2009), it is necessary to critically evaluate and synthesise current literature investigating the combined model. This may provide valuable insights for informing future educational practices and areas for future research. Accordingly, this review seeks to investigate the following research question: Is the combined model supported by empirical research?

2.3. Critical Evaluation of Evidence

2.3.1. Systematic Literature Search

A preliminary reading of literature pertaining to ASC, the BFLPE, the basic I/E model, and the combined model took place during May and June of 2019. Subsequently, a systematic search of the literature was undertaken in July 2019. A final search for relevant literature was undertaken in March 2021. The searches sought to identify literature which investigated both the I/E model and the BFLPE simultaneously. Therefore, the search terms ‘big fish little pond effect’ AND ‘internal/external frame of reference’ were used. The databases British Education Index, Education Full Text (H.W. Wilson), Education Source, ERIC, PsycARTICLES and PsycINFO were searched using EBSCOhost. The databases Web of Science Core Collection, Current Content Connect, Data Citation Index, KCI-Korean Journal Database, Medline, SciELO Citation Index and Russian Science Citation Index were searched using Web of Science. These searches yielded 64 results, of which 37 were duplicates and subsequently removed. The 27 remaining studies were screened in line with the criteria documented in Table 1. On the basis of an initial title and abstract screening, 11 studies were excluded from the review. The remaining 16 studies were read in full to determine whether they met the inclusion criteria, and this resulted in a further seven studies being excluded. During the full-text screening, no additional citations which were likely to

have adhered to the review criteria were found. Hence, nine articles were included in this review. These are listed in Table 2. A flow diagram documenting the search process can be seen in Figure 6. Excluded studies are listed in Appendix B.

Table 1

Inclusion and Exclusion Criteria

Study Feature	Inclusion Criteria	Exclusion	Rationale
1. Type of publication	Peer-reviewed journal	Not peer-reviewed journal	Peer-reviewed journals are rigorously evaluated prior to publication; material likely to be of a higher standard
2. Language	Study published in English	Study not published in English	No access to translator
3. Type of study	Study explicitly investigates the BFLPE and the I/E model simultaneously, within a unified model (i.e. the combined model)	Study does not test for the presence of the BFLPE and the I/E model simultaneously, within a unified model Study simultaneously tests for additional theories/assumptions within the model, which are outside of the principal predictions of the BFLPE, I/E model, and/or the combined model	The current review seeks to ascertain whether the combined model is supported by empirical evidence
	At least two distinctive domains of ASC or another motivational construct tested	Less than two domains of ASC or another motivational construct tested	Two domains need to be tested to test for the principal assumptions of the I/E model

Table 2

List of Studies Included in the Review

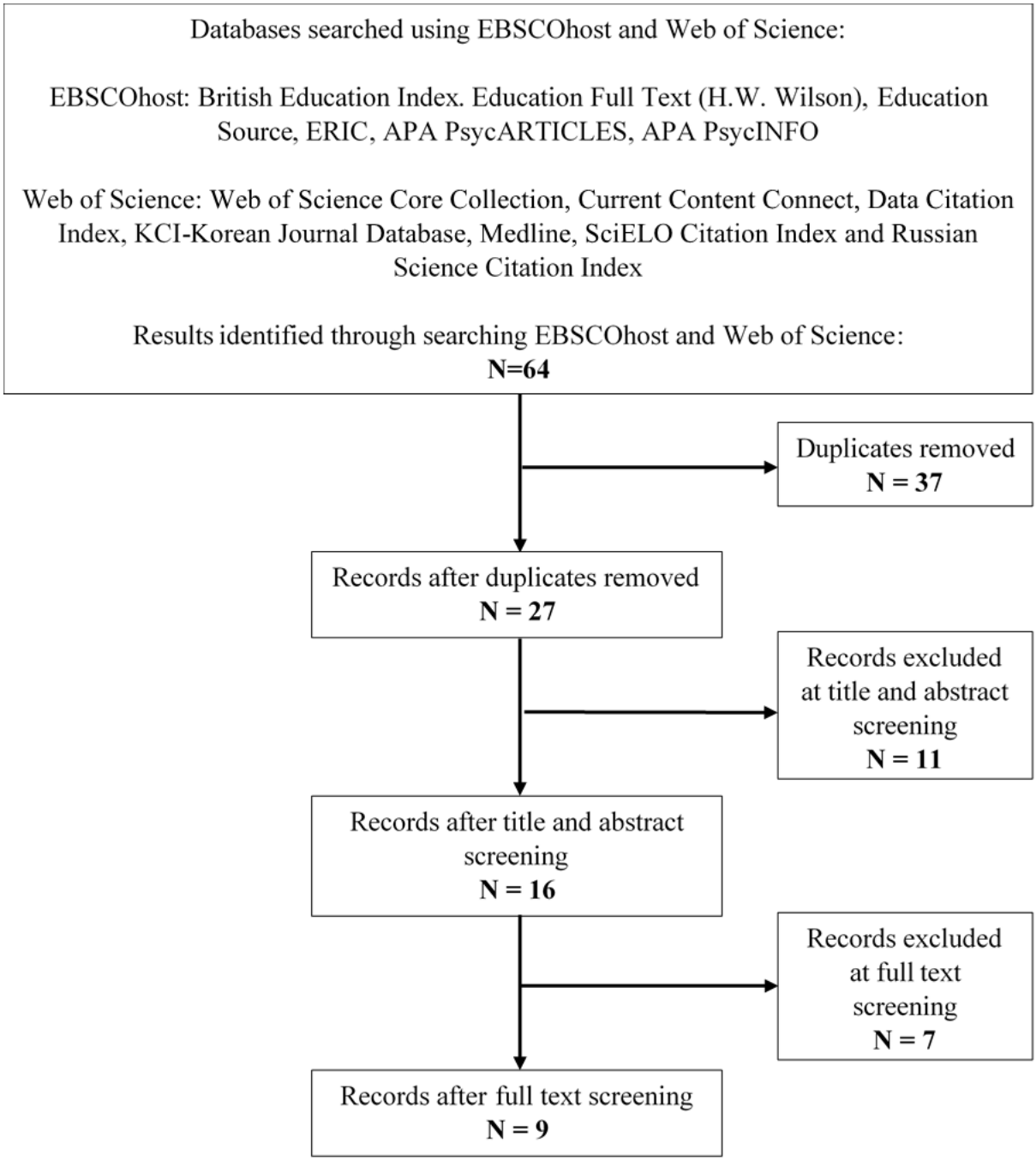
Reference(s)
Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology, 104*(1), 87-107. https://doi.org/10.1037/a0025734
Guo, J., Marsh, H. W., Parker, P. D., & Dicke, T. (2018). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learning & Instruction, 58,* 210-219. https://doi.org/10.1016/j.learninstruc.2018.07.007
Kavanagh, L. (2019). Academic self-concept formation: Testing the internal/external frame of reference model, big-fish-little-pond model, and an integrated model at the end of primary school. *European Journal of Psychology of Education, 35*, 93-109. https://doi.org/10.1007/s10212-019-00416-w
Lohbeck, A., & Möller, J. (2017). Social and dimensional comparison effects on math and reading self-concepts of elementary school children. *Learning and Individual Differences, 54*, 73-81. https://doi.org/10.1016/j.lindif.2017.01.013
Marsh, H. W. (1990c). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology, 82*(1), 107-116. https://doi.org/10.1037/0022-0663.82.1.107
Marsh, H. W. (1994). Using the national longitudinal study of 1988 to evaluate theoretical models of self-concept: The self-description questionnaire. *Journal of Educational Psychology, 86*(3), 439-456. https://doi.org/10.1037/0022-0663.86.3.439
Parker, P. D., Marsh, H. W., Ludtke, O., & Trautwein, U. (2013). Differential school contextual effects for math and English: Integrating the big-fish-little-pond effect and the internal/external frame of reference. *Learning and Instruction, 23*, 78-89. https://doi.org/10.1016/j.learninstruc.2012.07.001
Pinxten, M., Wouters, S., Preckel, F., Niepel, C., De Fraine, B., & Verschueren, K. (2015). The formation of academic self-concept in elementary education: A unifying model

for external and internal comparisons. *Contemporary Educational Psychology, 41*, 124-132. https://doi.org/10.1016/j.cedpsych.2014.12.003

Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning & Instruction, 34*, 32-41. https://doi.org/10.1016/j.learninstruc.2014.08.001

Figure 6

Flow Diagram Documenting the Search Process



2.3.2. Literature Review

A summary of the included studies can be found in Appendix C. This documents the research sample, a brief overview of the study, the measures employed, and the main findings from each study, all of which were considered when reviewing the literature. In reviewing the studies, the studies were evaluated using Gough's (2007) Weight of Evidence (WoE) framework which comprises of three core elements, WoE A, WoE B, and WoE C. A description of each WoE is provided below. An overall weighting, WoE D, is subsequently deciphered for each study, and this is established based on the WoE A, WoE B, and WoE C scores for each study.

WoE A. WoE A appraised the overall methodological quality of the research and how well this was implemented. Each study was evaluated in accordance with the coding protocol that is appended in Appendix D. This protocol was compiled in line with (i) the review question, (ii) relevant studies, and (iii) relevant literature documenting quality standards for correlational research (Howitt & Cramer, 2011; Mertens, 2015). Items for the 'practical and clinical significance' section and the 'data analysis' section of the protocol were taken from Thompson et al.'s (2005) 'Suggested Quality Indicators for Correlational Research'. As documented in Appendix D, studies received a rating of one (low), two (medium), or three (high), depending on their overall score on the coding protocol.

WoE B. WoE B evaluated the appropriateness of the study design for answering the review question. Criteria for judging WOE B were established considering the review question and relevant literature documenting the appropriateness of different types of research methodologies and designs for answering particular types of research questions (Howitt & Cramer, 2011; Mertens, 2015). The criteria are attached in Appendix E. As above for WoE A, studies were awarded a weight of one (low), two (medium), or three (high), depending on their overall score on the WoE B coding protocol. More information on this can be found in Appendix E.

WoE C. WoE C assessed the relevance of the evidence obtained in the study for answering the review question. Consideration was given to factors such as the characteristics of the research sample which may impact the generalisability of study findings and the specific purpose of each research paper, as documented by the authors of each article. The specific criteria which were used to assess this can be found in Appendix F. In line with WoE A and WoE B, each study received a rating of one (low), two (medium), or three (high),

based on the number of criteria which the study adhered to in line with the coding protocol. Further information on this can be found in Appendix F.

Given the WoE A, WoE B, and WoE C ratings assigned to the individual studies, an overall rating was assigned to the study (WoE D). This was ascertained by calculating the mean score of WoE A, WoE B, and WoE C for each study. It is important to note, that the WoE ratings do not necessarily indicate whether or not a study is of high, medium, or low quality, but rather the framework was used to address the overarching review question, to evaluate the studies, and to identify gaps in current literature which may inform areas for future research. In line with WoE A, WoE B, and WoE C, a number of factors were considered when assigning ratings to each study, including the relevance of the studies to the focus of the current review. Table 3 documents the WoE ratings that were allocated to each study.

Table 3

Weight of Evidence Rating for Each Study

	WoE A	WoE B	WoE C	WoE D
Chiu (2012)	3 (high)	1 (low)	3 (high)	2.33 (high)
Guo et al. (2018)	2 (medium)	3 (high)	1 (low)	2 (medium)
Kavanagh (2019)	3 (high)	3 (high)	3 (high)	3 (high)
Lohbeck and Möller (2017)	3 (high)	2 (medium)	2 (medium)	2.33 (high)
Marsh (1990c)	1 (low)	2 (medium)	3 (high)	2 (medium)
Marsh (1994)	2 (medium)	1 (low)	3 (high)	2 (medium)
Parker et al. (2013)	3 (high)	3 (high)	3 (high)	3 (high)
Pinxten et al. (2015)	3 (high)	3 (high)	3 (high)	3 (high)
Schurtz et al. (2014)	1 (low)	1 (low)	2 (medium)	1.33 (low)

2.3.2.1. Participants. Participant sample sizes differed significantly between studies and ranged from 291 participants (Lohbeck & Möller, 2017) to 139174 participants (Chiu, 2012). Eight studies in the current review were secondary analyses which used large-scale datasets (Chiu, 2012; Guo et al., 2018; Kavanagh, 2019; Marsh, 1990c, 1994; Parker et al., 2013; Pinxten et al., 2015; Schurtz et al., 2014), while Lohbeck and Möller (2017) was the only study which reported primary empirical data. Some studies contained data from international assessments, such as Chiu (2012) who evaluated data from 27 countries that participated in the 2003 Trends in International Mathematics and Science Study (TIMSS), and Guo et al. (2018) who evaluated data from 15 countries that completed the 2011 TIMSS and Progress in International Reading Literacy Study (PIRLS). Other studies contained data from one particular country or region, and were typically completed using datasets that were representative of a defined region, country or area. For example, Kavanagh (2019) investigated the relevance of the BFLPE, the basic I/E model, and the combined model to Irish sixth-class pupils' ASCs using data from the 2014 National Assessments of Mathematics and English Reading in Ireland. Appendix C provides further information regarding the specific countries/regions from which sample participants in each study were sourced. For all studies, except Lohbeck and Möller (2017) and Schurtz et al. (2014), there was evidence that probability-based sampling measures were employed during data collection. Implementation of probability-based sampling measures increases the generalisability of research findings (Howitt & Cramer, 2011). It was not documented how Lohbeck and Möller (2017) and Schurtz et al. (2014) recruited sample participants, and therefore their studies obtained a lower WoE C.

Participants in different studies also varied in age/school grade. Samples ranged from participants in grade two, aged between seven to nine years old (Lohbeck & Möller, 2017), to participants who were in their final year of secondary school, who had an average age of 19.76 years (Parker et al., 2013). Accordingly, the current review encompassed data from a diverse range of countries and age/school grade cohorts. In all papers, consideration was also given to the characteristics of the sample population in terms of how this may impact the results, generalisability, and limitations of the findings. This included the age of the pupils in the sample cohort and how this may impact their self-report on motivational construct measures, notably ASC (Lohbeck & Möller, 2017; Pinxten et al., 2015), and factors such as the type of school which sample participants attended. For example, considering the German education system, Parker et al. (2013) noted that results from their analyses of the combined

model differed across themed and traditional university track schools. Themed schools enrol pupils in their programs based on pupils' individual interest, aptitude, and/or achievement in one particular area. The pupils' education within these schools prepares them to pursue future study and/or work in their area of interest/strength. Traditional university track schools are also selective in terms of their enrolment, but they enrol students who demonstrate an overall high standard of competence across many academic areas (e.g. mathematics and English). As proposed by Parker et al. (2013), comparison processes may be accentuated in themed rather than traditional university track schools as differences in pupils' performances in different academic domains are typically more differentiated in themed schools than in traditional university track schools.

2.3.2.2. Measures. Overall, achievement measures employed in the included studies had strong validity and reliability. The tests were typically designed and used to assess national and/or international achievement standards in the relevant subject areas; they contained a varied range of items, and aligned with curriculums and standards that would usually be expected of pupils in the relevant grades. Exceptions to this were Lohbeck and Möller's (2017) measure of mathematics achievement which solely assessed pupils' arithmetic skills, Marsh's (1994) measure of reading achievement which aligned with pupils' self-report of their English self-concept, and Schurtz et al.'s (2014) measure of English as a foreign language competence which consisted of pupils completing a timed four-minute reading test in which they had to identify irrelevant words in sentences. Arguably, these achievement measures may have led to construct underrepresentation of achievement in a given domain compared to the broader self-concept construct measures which they were intended to align with (Howitt & Cramer, 2011). Accordingly, this was considered when assigning WoE A ratings to these studies. Reliability coefficients of all achievement measures were generally above .7. If studies did not report the reliability coefficient of one or more achievement measures, and/or if one or more measure(s) had a reliability coefficient of less than .7, this resulted in the studies receiving a lower score on the WoE A coding protocol (i.e. Guo et al., 2018; Marsh, 1990c; Schurtz et al., 2014).

In all studies, class/school-average achievement measures were computed based on individual pupils' achievement scores. The data was hierarchical such that individual pupils were nested within schools/classes. As outlined by Marsh et al. (2009), such class/school-average achievement measures may be prone to sampling error due to inaccurate sampling of the desired population (i.e. the class/school in this case). For example, if the estimated class-

average achievement measure was computed by calculating the mean of five participants' achievement scores, out of an overall class of 20 pupils, it is likely that this score may not be an accurate representation of the actual whole class mean. Hence, this can impact the validity of the class/school-average measure.

In line with the recommendations of Marsh et al. (2009), many studies in the current review employed sampling procedures which minimised potential sampling error, such as by aiming for full participation among sample schools and/or classes (Chiu, 2012; Guo et al., 2018; Kavanagh, 2019; Pinxten et al., 2015). Another approach outlined by Marsh et al. (2009) which may be employed to minimise the sampling error is to employ a latent aggregation approach to computing class/school-average achievement measures within structural equation models. This approach was implemented by Marsh (1990c), Marsh (1994), and Parker et al. (2013) who only sampled a limited number of pupils from each of their sample schools. Neither Lohbeck & Möller (2017) or Schurtz et al. (2014) acknowledged how their sampling design may have minimised the sampling error in their study when using manifest aggregation to calculate between level variables. Schurtz et al. (2014) used data from the BiKS-8-14 longitudinal study at the University of Bamberg, however when undertaking the current review, information regarding the sampling design, recruitment process and composition of sampled classes could not be found. As the sampling error associated with between level variables (in this case class/school-average achievement) may result in inaccurate study findings, this resulted in both Lohbeck and Möller's (2017) and Schurtz et al.'s (2014) studies receiving a lower WoE B.

ASC measures also demonstrated good validity. Lohbeck and Möller (2017), Marsh (1994), Parker et al. (2013), and Pinxten et al. (2015), explicitly referenced the use of an appropriate version of the 'Self-Description Questionnaire' (Marsh, 1990c), given their sample population. As documented in their studies, this measure is one of the most empirically sound measures of self-concept; it is typically employed when investigating ASCs, and has high validity and reliability (Arens et al., 2013). Concurrently, it should be noted that in Lohbeck and Möller's (2017) study while they used the short German version of the Self-Description Questionnaire I-GS (SDQ I-GS; Arens et al., 2013), their ASC measures contained only three items each as they included items which solely pertained to pupils' beliefs regarding their competencies. They also used a four-point Likert scale, which they shortened from the five-point Likert scale which is typically used and specified on the SDQ I-GS. Lohbeck and Möller (2017) subsequently documented that their results may have been

impacted by ceiling effects due to their ASC measures. Arguably the short three-item questionnaire reduced the potential level of variance which would otherwise have been present in participants' responses.

While other studies did not cite similar sources or literature when describing their ASC measures, the items used in the scales conformed with items that are typically employed to measure ASCs. Hence, only items which pertained to pupils' self-beliefs regarding their competencies and abilities were included in the ASC measurement scales. One exception to this was Kavanagh's (2019) measure of reading self-concept which included items pertaining to pupils' future performance and pupils' report of how their teacher perceived their reading ability. As noted by Kavanagh (2019), such items are not typically used when assessing pupils' ASCs. Although this resulted in Kavanagh's (2019) study receiving a lower mark on the WoE A protocol, overall the study received a high rating due to the sound execution of other methodological features. All reliability coefficients for ASC measures were high and above .7, except for the science self-concept measure used by Chiu (2012), which had a Cronbach's alpha of .69 for the overall world sample.

The only study in the current review that employed motivational construct measures outside of ASC was Schurtz et al. (2014). They investigated the relevance of the model to pupil's academic subject-interests. In conceptualising subject-interest in the context of their study, they aligned the construct with the person-object-conception of interest outlined by Krapp (2002). Essentially, this theory proposes that interest is characterised by the relationship between a person and an object of interest. This relationship embodies both an emotional aspect, such as the enjoyment a person feels when interacting with the object of interest, and a value aspect, such as the personal value which one attributes to the object of interest. Consistent with this theory, Schurtz et al.'s (2014) subject-interest scale items pertained to both pupils' affect towards engaging in subject-specific activities and the value, or importance, which they attribute to the particular subject. The scales also demonstrated good reliability for both English and mathematics subject-interests. For both subjects, and for pupils' responses in both grade five and grade six, Cronbach's alpha was greater than .8. Thus, the subject-interest measures employed by Schurtz et al. (2014) demonstrated good scale reliability.

2.3.2.3. Design. Studies investigating the conventional I/E model were generally executed using two distinctly different academic domains, such as mathematics and verbal

skills. It was rationalised that this would allow for dimensional comparisons between pupils' performances and ASCs in the relevant subject areas to be observed, as dimensional comparisons were likely to be greater between more dissimilar domains (Chiu, 2012; Guo et al, 2018). In the current review, Chiu (2012) was the only study that explored just two subject areas (mathematics and science) which were more similar to each other than those which were habitually examined. Although Chiu (2012) advocated that these domains were conceptually distinct from each other, she stated that the examined areas differed in terms of their relatedness, compared to the relatedness between domains that were traditionally examined. Chiu (2012) noted that this was likely to have impacted the results of her study and should be taken into account when interpreting the findings in line with prior research. This limitation resulted in Chiu's (2012) study receiving a lower WoE B. Concurrently, although Guo et al. (2018) also examined mathematics and science, they also examined reading and thus this allowed for comparisons between the three domains both within the study and with prior literature.

WoE B also considered the time at which data was collected. As all studies were cross-sectional, data was collected during a specific period of time. Academic achievement and ASCs are influenced by factors such as changes in instructors, teaching approaches, and schools, and thus may change over time (Mitchell, 2014; Parker et al., 2013; Roy et al., 2015). Therefore, a higher WoE B was awarded to studies which collected both measures of academic attainment and ASC within the same year. In Marsh (1994), achievement measures were aggregated from data collected in 1988 and 1990, while ASC measures were collected in 1990. In addition, only high schools in which 15 or more pupils who completed the 1988 study attended were included in the evaluation, as high-school average achievement measures were based on data from the pupils who previously completed the 1988 assessments (Marsh, 1994). Arguably the time-lapse between the collection of measures in Marsh's (1994) study may have impacted the validity of the individual and school-average achievement measures, and the comparability between measures of achievement and measures of ASC at a given time. Marsh's study was subsequently awarded a lower WoE B.

2.3.2.4. Data Analysis. Statistical analysis methods that were employed in all studies were justified, appropriate and consistent with prior literature. However, given advancing literature from the past two decades detailing best practice, prior research which investigated the BFLPE and conventional I/E model separately, and the present research question, some studies employed more appropriate analyses than others.

In the current review, studies which employed a multilevel modelling structural equation modelling approach, in which ASC scales and/or motivational construct scales were operationalised as latent constructs, scored higher on the WoE B coding protocol. Given the hierarchical nature of the data, with individual pupils clustered within classes/schools, multilevel modelling was deemed to be the most appropriate method of analysis for investigating the combined model (Marsh et al., 2009). In addition, multilevel modelling is typically employed when investigating the basic BFLPE (Marsh et al., 2009). It was also rationalised that analyses in which motivational constructs, such as ASC, were construed as latent variables would be superior to analyses in which they were operationalised as manifest variables. As outlined by Marsh et al. (2009), manifest variables are variables which can be directly observed and measured, whereas latent variables cannot. Thus, examples of manifest variables include physical attributes such as weight and height. Motivational constructs, such as ASC cannot be objectively measured in this way. They are measured using scales which contain several indicators, and while participants' responses to these indicators can provide insight into their level of ASC, there is likely to be measurement error associated with the scale indicators and participants' responses to them (Marsh et al., 2009). Operationalising unobserved construct scales, such as ASC scales, as latent variables within analyses accounts for the potential measurement error associated with these scales, and implementation of structural equation modelling can facilitate the operationalisation of latent constructs within the analyses (Marsh et al., 2009).

Five studies in the present review obtained a greater WoE B score as they implemented multilevel structural equation models, in which ASC scales were operationalised as latent constructs, when analysing their data (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Parker et al., 2013; Pinxten et al., 2015). Four of these employed latent-manifest multilevel structural equation models, in which ASC variables were construed as latent constructs and achievement variables were manifest (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015). Only one study employed a doubly-latent multilevel structural equation model, in which both school-average achievement and ASC variables were construed as latent constructs (Parker et al., 2013). Marsh (1990c) and Marsh (1994) scored lower on WoE B as they did not employ multilevel modelling when analysing their data. Chiu et al. (2012) and Schurtz et al. (2014) also received a lower WoE B, as they employed a doubly-manifest approach to analysing their

data and thus operationalised all variables, including ASC and/or subject-interest, as manifest variables.

In all studies in the current review, correlation coefficients (*r*) or standardised parameter estimates (β) were provided for all parameters which were relevant to the models which were tested. However, the operationalisation of contextual effects (i.e. for the BFLPE) differed across the studies. As detailed by Marsh et al. (2009), while standardised parameter estimates are commonly used when reporting and interpreting single level analyses, and between level and within level parameters separately, effects sizes should be explicitly operationalised for contextual effects. Marsh et al. (2009) outlined that contextual effects pertain to group level effects or variance, such as class-average achievement, which are observed outside of level one individual effects and cannot be explained by level one effects. Thus, in the current studies, the BFLPE is an example of a contextual effect. While contextual effect sizes for multilevel models may be estimated in a number of different ways, the manner in which effect sizes are operationalised can impact the effect size estimates (Marsh et al., 2009). In line with Marsh et al.'s (2009) recommendations, the most recent studies investigating the combined model have employed a conservative approach to the estimation of effect sizes (ES) using the *ES2* formula: $ES = 2 \times \beta \times SD_{\text{predictor}}/SD_{\text{outcome}}$ (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Parker et al., 2013; Pinxten et al., 2015). This consensus within recent literature is advantageous for reviewing studies and for informing future research, as the contextual effect sizes from these studies may be directly compared.

Most studies in the current review used their data to investigate the conventional I/E model, the BFLPE, and the combined model (Chiu, 2012; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh, 1990c, 1994; Parker et al., 2013; Pinxten et al., 2015; Schurtz et al., 2014). In line with the overall aim of the current review, this facilitated evaluation of the combined model as comparisons could be made between the analysis of each model. Hence, these studies attained a higher WoE C. WoE C also considered the relevance of the individual studies to the review question and the critique of evidence in the paper in line with the review question. This discussion could further aid with interpreting the evidence and provide additional insights which may aid with answering the present review question. As the principal focus of Guo et al. (2018) and Schurtz et al. (2014) was not to investigate the combined model, as outlined by Chiu (2012), and their discussion did not primarily focus on the current review question, they obtained a lower WoE C.

2.3.3. Synthesis of Findings

Overall, the central assumptions of the combined model were largely supported by the studies (Chiu, 2012; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh, 1990c, 1994; Parker et al., 2013; Pinxten et al., 2015; Schurtz et al., 2014). Specifically, class/school-average achievement in one subject was found to be negatively correlated with pupils' self-concept in that subject area. Consistent with the conventional I/E model, individual achievement in one domain was found to have a positive effect on pupils' self-concepts in that domain, but a smaller negative effect on pupils' self-concepts in contrasting domains. The models and effects were also found to be significant when tested in relation to pupils' academic subject-interests (Schurtz et al., 2014). Deviations to these expectations are discussed further below (e.g. Chiu, 2012; Lohbeck & Möller, 2017). Appendix C contains an overview of the key findings from all reviewed studies. Considering the results in their entirety, as the assumptions of both the traditional I/E model and the BFLPE were evident when investigated simultaneously within the combined model, this supports the unique contribution of both social and dimensional comparison processes to the construction of pupils' ASCs. Thus, it enhances our understanding of factors that contribute to pupils' ASCs.

2.3.3.1. Generalisability Across Countries. Consistent with prior research which investigated the conventional I/E model and the BFLPE separately, good support was found for the generalisability of the combined model across countries (Chiu, 2012; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh, 1990c, 1994; Parker et al., 2013; Pinxten et al., 2015; Schurtz et al., 2014). All studies, except Lohbeck and Möller (2017), provided support for the central assumptions of the combined model. Concurrently, the degree to which data from different countries supported the model and effect sizes between relevant variables differed between samples. This was particularly evident in Chiu (2012) and Guo et al. (2018), who evaluated data from 27 countries and 15 countries respectively. For instance, in Chiu's (2012) study, the predictions of the model were fully supported for pupils' mathematics self-concept in 16 countries, and partially supported for pupils' mathematics self-concept in 11 counties. Meanwhile, in relation to pupils' science self-concept the combined model was fully supported in nine countries, partially supported in 15 countries, and not supported in four countries (Chiu, 2012). Regarding the predictions of the model on pupils' reading self-concept and mathematics self-concept in Guo et al's (2018) study, in every country individual mathematics/reading achievement was significantly related to pupils' corresponding domain-specific ASCs, and class-average reading/mathematics achievement was significantly

negatively related to pupils' corresponding domain-specific ASCs (Guo et al., 2018). However, while in all but two countries cross-domain correlations were in the expected direction (i.e. individual mathematics achievement is negatively correlated with pupils' reading self-concept; individual reading achievement is negatively correlated with pupils' mathematics self-concept), these effects were only significant in six countries for mathematics achievement on reading self-concept, and six countries for reading achievement on mathematics self-concept (Guo et al., 2018). These findings support previous advice by Chiu (2012) and Pinxten et al. (2015), who advocated that different countries consider the theory and key assumptions with data from their own country/area to ascertain the relative importance of factors that may be contributing to pupils' ASCs (Chiu, 2012; Pinxten et al., 2015). Subsequently, this data can be used in line with ASC literature to inform educational practices that may be particularly relevant to certain contexts.

2.3.3.2. Generalisability Across Age Groups. The studies indicated that the age of the sample cohort may have influenced the findings. In response to the previous lack of research investigating the BFLPE, the basic I/E model, and/or the Combined model among younger pupils, five studies in the current review analysed data from elementary school pupils (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015; Schurtz et al., 2014). Four of these studies provided support for the combined model among elementary school pupils; the central assumptions of the combined model were found to be significant. Conversely, Lohbeck and Möller (2017) did not find strong support for the Combined model among their sample of grade two pupils. This was the youngest sample cohort in the present review. The authors observed that although all of the expected associations between the variables were in the expected direction, internal cross-domain comparisons were not significant. A small, significant BFLPE was found for mathematics self-concept, but not for reading self-concept. Lohbeck and Möller (2017) postulated that pupils' frames of reference for making comparisons may not yet be as developed as that of older pupils, on which the basic I/E model and BFLPE were traditionally examined. Pinxten et al. (2015) and Guo et al. (2018), who investigated the combined model among fourth-grade pupils, also noted that their effect sizes were smaller compared to those reported in previous investigations of the basic BFLPE and I/E models on older pupils. These samples were the second youngest sampled cohort in the current review. Conversely, consideration needs to be given to the limitations of Lohbeck and Möller's (2017) study. For instance, their measure of mathematics achievement was limited, the authors stated that ceiling effects may have

impacted their ASC measures and subsequent results, and their sample size of just 291 pupils was considerably smaller compared to that of other studies. This limits the degree to which their findings can be generalised among grade two pupils in other contexts.

2.3.3.3. The Relevance of the Combined Model to Different Subject Areas.

Differences in findings also emerged based on the subject areas that were investigated. In particular, the studies underscored the significance of the BFLPE to elementary pupils' mathematics self-concept. In all studies that analysed data from elementary school pupils, effect sizes for the BFLPE were consistently stronger for mathematics self-concept than that of the contrasting domain (Chiu, 2012; Guo, et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015; Schurtz et al., 2014). This indicates that the formation of elementary pupils' mathematics self-concept is particularly susceptible to social comparison processes. Even in Lohbeck and Möller's (2017) study, a significant BFLPE for mathematics self-concept was found among second-grade pupils. Within the Irish context, Kavanagh (2019) highlighted this trend among sixth-class pupils. Kavanagh's (2019) work also highlighted the heightened relevance of dimensional comparisons to pupils' mathematics self-concept compared to their reading self-concept. In Kavanagh's (2019) study, effect sizes for all assumptions of the combined model were stronger for mathematics. Relevant effect sizes for Kavanagh's (2019) study are documented in Appendix C.

In contrast to domains which were habitually examined (i.e. mathematics and English), two studies in the present review investigated science self-concept (Chiu, 2012; Guo et al., 2018). Findings for the effects of the combined model on science self-concept were not as robust as they were for mathematics self-concept or reading self-concept. Compared to mathematics self-concept, Chiu (2012) found that support for the combined model for science self-concept was less generalisable across countries. Moreover, parameter estimates between self-concepts and all relevant social and dimensional comparisons variables were notably higher for mathematics self-concept than for science self-concept. These findings are documented in Appendix C. Likewise, Guo et al. (2018) found that the BFLPE was strongest for mathematics, followed by reading, and then science. The BFLPE was not significant for science in 6 countries. Conversely, while Guo et al.'s (2018) results for mathematics and reading self-concept were consistent with predictions of the combined model, their results for reading and science self-concept were not. Across the examined countries it was found that reading achievement had a positive effect on science self-concept, and science achievement had either a non-significant or positive effect on reading self-

concept. The authors argued that both reading and science incorporate similar skills and approaches to learning at primary level. For instance, pupils may learn a considerable amount of factual science through reading during their elementary years in school. Accordingly, these subjects may be perceived as being similar to each other, and thus achievement in one domain may contribute to a more positive self-concept in the other domain. This is consistent with Chiu (2012), who proposed that the non-significant effect of mathematics achievement on science self-concept, may be due to mathematics and science being more similar to each other than subject areas which were traditionally examined (i.e. mathematics and language skills). This provides further evidence that the influence of the assumptions of the combined model on pupils' ASCs differ across subject domains.

2.3.3.4. The Combined Model and Subject-Interest. Only one study in the current review investigated the relevance of the I/E model, the BFLPE, and the combined model on a motivational construct other than ASC. The study was completed by Schurtz et al. (2014), who investigated the relevance of the models to sixth-grade German pupils' academic subject-interests in mathematics and English. The results of each model found that the parameter estimates were statistically significant. Schurtz et al. (2014) subsequently investigated the combined model on pupils' subject-interests with pupils' grades and ASCs as mediating variables. The model results demonstrated that pupils' domain-specific ASCs had a moderate effect on pupils' domain-specific subject-interests. In addition, the model parameters indicated the following corresponding domain effects: pupils' overall level of competence in a particular subject impacted pupils' grades; in turn, pupils' grades impacted pupils' ASCs. For mathematics, class-average competencies had a moderate effect on pupils' grades and a small effect on pupils' ASCs. For English, class-average competencies only had a small significant effect on pupils' grades. Thus, the model indicated that pupils' overall achievement or level of competence impacted pupils' grades and ASCs, which in turn impacted pupils' subject-interests (Schurtz et al, 2014).

2.3.3.5. Extension of the Central Predictions of the Combined Model. Extending the basic assumptions of the BFLPE and conventional I/E model, seven studies investigated the effect of class/school-average achievement in one domain on pupils' self-concept in a contrasting domain (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh 1990c, 1994; Parker et al., 2013; Pinxten et al, 2015). Both Marsh (1990c) and Pinxten et al. (2015) found small positive effects when investigating the effect of school/class-average achievement on self-concepts in contrasting domains. Additionally, Marsh (1994) found that

school-average English achievement had a small, significant effect on mathematics self-concept, however, school-average mathematics achievement did not have a significant effect on pupils' English self-concepts. Considering the Irish context, Kavanagh (2019) did not find significant cross-domain effects of class-average achievement on mathematics or reading self-concepts. Likewise, Lohbeck and Möller (2017) did not find evidence to support this assumption among second-grade pupils, and Guo et al. (2018) did not find evidence to support the assumption among fourth-grade pupils across 15 countries, one of which was Ireland.

Intriguingly, Parker et al. (2013) found mixed results for the aforementioned extended assumption of the combined model, based on school contexts. They investigated the combined model in themed and traditional university track schools in Germany. Pupils' mathematics and English achievement scores were significantly more differentiated in themed schools than in traditional university track schools. The researchers argued that the heightened discrepancy between pupils' performance in different subject areas in themed schools may lead to increased social and dimensional comparison processes to occur. The results from their study supported this as significant within and cross-domain correlations between school-average achievement and self-concepts were found for both mathematics and English in themed schools. However, in traditional university track schools, only a significant BFLPE for mathematics was found. No significant BFLPE was found for English self-concept and no significant effects were found for cross-domain school-average achievement comparisons. As noted by the authors, this indicates that school context is a key factor in determining how ASCs are constructed. Concurrently, as the overall achievement standard was higher in traditional track schools, it may be fruitful for future research to examine the potential moderating role of pupil ability level on the influence of social and dimensional comparisons on pupils' ASCs.

2.4. Conclusion

2.4.1. Implications for Theory

Overall, findings from the current review support the combined model for ASC formation (Chiu, 2012; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh, 1990c, 1994; Parker et al., 2013; Pinxten et al., 2015). The central assumptions of the basic BFLPE and the I/E model were both evident within the unified model. This signifies that both social and dimensional comparison processes contribute independently to the process of ASC formation.

Thus it advances prior literature which previously investigated the BFLPE and I/E models separately. Combining these two frameworks into a unified model provides a more comprehensive conceptualisation of ASC formation. The current review demonstrates that this model is supported by empirical research (Chiu, 2012; Guo et al., 2018; Kavanagh, 2019; Marsh, 1990c, 1994; Parker et al., 2013 Pinxten et al., 2015).

2.4.1.1. The Combined Model and Contextual Factors. While strong support was found for the generalisability of the combined model across countries some discrepancies in findings between countries and school contexts, such as themed versus university track schools in Germany, were noted (e.g. Chiu, 2012; Guo et al., 2018; Parker et al., 2013). This indicates that contextual factors can impact the relative influence of social and dimensional comparisons on ASC formation. Chiu (2012) suggested that every country may "tell its own story" (p. 100), thus the theoretical relevance of the models to pupils' ASCs should be considered in line with particular countries, cultures, and contexts.

2.4.1.2. The Combined Model and Age. Similarly, the review indicated that the relevance of the theory and associated predictions of the combined model to pupils' ASCs may differ due to the age of the sample cohort. In particular, the combined model was not supported among grade two pupils (Lohbeck & Möller, 2017), and studies which were completed on grade four pupils noted that their effect sizes were typically smaller than those which previously investigated the basic BFLPE and basic I/E model separately among older pupils (Guo et al., 2018; Pinxten et al, 2015). However, while present literature suggests that the impact of social and dimensional comparisons, and accordingly the predictions of the combined model, on young elementary pupils' ASCs are not as pronounced as in older pupils, it has been underscored that elementary school is a critical time during which pupils construe their ASCs; and that social and dimensional comparison processes may be contributing factors to this developmental process (Fang et al., 2018; Guo et al., 2018; Lohbeck & Möller, 2017; Möller et al., 2009; Pinxten et al., 2015).

2.4.1.3. The Relevance of the Combined Model to Different Subject Areas. Differences in findings also emerged based on the subject areas which were investigated. Most notably, the results of the current studies indicated that the predictions of the combined model were most relevant to elementary school pupils' mathematics self-concept (Guo et al., 2018; Kavanagh, 2019; Pinxten et al., 2015). This indicates that social and dimensional comparison processes are likely to be most influential on young pupils' mathematics self-

concepts. Considering this in line with current literature pertaining to children's early childhood learning experiences, it has been noted that parents typically expend more time engaging in early literacy learning activities than mathematical learning activities with their children at home (Ginsberg et al., 2012; Skwarchuk, 2009). Therefore, as proposed by Schurtz et al. (2014) it is possible that motivational constructs pertaining to mathematics, such as pupils' mathematics self-concepts, may be more impacted by social and dimensional comparison processes that may occur within the school environment.

2.4.1.4. The Combined Model and Subject-Interest. In addition to the theory and models being relevant to pupils' ASCs, one study in the present review investigated the relevance of the models to pupils' subject-interests (Schurtz et al., 2014). Consistent with the aforementioned literature pertaining to pupils' ASCs, this study found that the predictions of the models were upheld, and that the findings were particularly pertinent to pupils' mathematics subject-interests. Subsequently, Schurtz et al. (2014) demonstrated that pupils' ASCs may in turn impact their subject-interests. Therefore, extending prior literature which investigated the relevance of the combined model to pupils' ASCs, the findings from Schurtz et al.'s (2014) study suggest that social and dimensional comparison processes may also impact other motivational constructs, notably pupils' subject-interests, which in turn can impact pupils' learning behaviour.

2.4.2. Implications for Practice

The above findings have noteworthy implications for practice. Given the fundamental importance of motivational constructs, such as the development of positive ASCs and interest in learning, for optimising learning, it is important to understand factors which can influence the development of these constructs. Accordingly, this information can be used to further inform policy and practice.

2.4.2.1. Pupils' Attitudes, Self-Beliefs, and Current Curricula. Within the Irish context, as in many international curricula, the development of positive attitudes and dispositions towards learning are now established as key curricula objectives (NCCA, 2018, 2019; Mullis, Martin, Ruddock, et al., 2009; Mullis, Martin, Kennedy, et al., 2009; Mullis et al., 2017, 2019). These include the development of positive self-beliefs regarding pupils' learning abilities, and the development of curiosity and interest towards learning. Indeed, these attitudes and beliefs are particularly pertinent to the aforementioned recent curricula in Ireland, in which constructivist and pupil-led learning methodologies are emphasised.

Positive self-beliefs regarding one's competence can impact pupils' thinking and learning behaviour, such as the persistence which they exert when engaging in learning tasks and the interest which they exhibit towards learning (Deci & Ryan, 1985; Ryan & Deci, 2000b). In turn, these learning behaviours and attitudes result in enhanced learning, and are particularly necessary when pupils are required to engage in constructivist and self-directed learning opportunities. Therefore, as the current findings indicate that social and dimensional comparison processes can influence pupils' ASCs, and potentially additional motivational constructs such as subject-interests, it is important that educational practitioners, such as educational psychologists and teachers, are made aware of the current findings and accordingly incorporate this knowledge into their practice.

2.4.2.2. Disseminating Findings to Enhance Teaching and Learning. Reflecting on the relevance of the current findings to educational psychology within the Irish context, the current model of service which the National Educational Psychological Service (NEPS) provides to schools consists of a three-tiered approach to intervention (NEPS, 2013). The three tiers are support for all pupils in schools, support for some pupils (i.e. pupils in need of additional, and/or more targeted, support), and support for few (i.e. pupils who present with the highest level of individual need; this can typically involve individual casework with a pupil in a school). Support for all can commonly comprise of disseminating psychological knowledge to schools which in turn can inform and enhance practice within schools. For example, best practice guidelines for schools have been published regarding how pupils' wellbeing, behavioural skills, and literacy skills may be promoted (Department of Education and Skills & Department of Health, 2015; NEPS, 2010, 2016, 2019). Meanwhile, guidelines to support mathematics were first published in 2020 (NEPS, 2020). Of note, while the 2020 guidelines provide advice regarding how to alleviate mathematics anxiety which pupils may experience and how to promote pupils' confidence in their mathematical abilities, the relative impact which social and dimensional comparisons may have on pupils' self-beliefs and learning behaviour is not documented. Interestingly, Irish pupils' performance rankings on international tests of fourth-grade pupils have consistently been higher in reading tests, specifically PIRLS, than mathematics assessments, notably the TIMSS study, over the past decade (Clerkin et al., 2016; Eivers & Clerkin, 2012; Eivers, Gilleese, et al., 2017; Perkins & Clerkin, 2020). In addition, the most recent TIMSS study (2019) highlighted that while both fourth-class and second-year pupils in Ireland ranked highly among countries which participated in the 2019 TIMSS, scores among Irelands' highest achieving pupils were lower

than those of similar high achieving countries, who largely ranked lower than Ireland overall (Perkins & Clerkin, 2020). Therefore, it is essential to consider factors which may influence the performance of all pupils, such as factors which may impact the learning attitudes and behaviours of all pupils, which in turn may enhance overall learning and performance. This may be facilitated through dissemination of the findings in relevant best practice guidelines.

2.4.2.3. Practical Strategies. Reflecting upon how the negative impact of social and dimensional comparison processes on pupils' ASCs may be alleviated, it may first be useful to consider strategies which may be particularly useful for encouraging pupils to develop positive self-beliefs and attitudes towards learning. Among those highlighted in the literature and current NEPS guidelines for promoting a culture of confidence within classrooms include normalising mistakes, valuing all pupils' ideas and efforts, and ensuring that teachers do not model or project any anxiety towards mathematics which they themselves may experience to their pupils (NEPS, 2020). In addition, it is critical to differentiate classroom instruction in accordance with pupils' needs. In a novel study investigating the BFLPE on pupils' French self-concept, Roy et al., (2015) demonstrated that the BFLPE was moderated by teachers' use of differentiated classroom instruction. Specifically, Roy et al. (2015) found that the effect of the BFLPE was not significant when teachers reported higher use of differentiated instruction strategies. Thus, current research and literature suggests that the potential negative effect of social and dimensional comparisons on pupils' ASCs may be alleviated by ensuring that classroom learning is differentiated, and by valuing all pupils' contributions and efforts.

2.4.3. Areas for Future Research

2.4.3.1. The Combined Model as a Framework for Future Research. Considering prior literature on the basic I/E model and the BFLPE, and the empirical support for the combined model in the current review, this framework is considered to be valuable for exploring ASC formation, or potentially associated motivational constructs, in future research. In addition, synthesis of the literature and findings from this review resulted in identification of a number of areas for future research. Firstly, as the relative significance of the models to pupils' ASCs has been found to differ across countries, school contexts and the subjects investigated, it is advocated that future research examines the relevance of the models to different educational contexts and academic subjects. For instance, subjects investigated in the current review included pupils' ASC regarding their first language, languages which they learned as an additional language, and science. Therefore it may be

particularly worthwhile to investigate the relevance of the models to other subject areas such as art, music, history, and geography. In addition, Chiu (2012) and Guo et al. (2018) demonstrated that the statistical significance and effect sizes relating to the models differed across the countries and subject domains which were investigated. Accordingly, by investigating the relative significance of the models to pupils' ASCs across different countries, settings, cultures and subjects, the interpretation and application of the findings may be considered in line with the relevant educational contexts and thus facilitate optimal implementation of the findings.

2.4.3.2. The Relevance of the Combined Model to Younger Pupils. Another area which emerged from the review was the need to investigate the relevance of the models to younger pupils' ASCs. As previously noted, only one study in the present review investigated the model on pupils' in second-grade (Lohbeck & Möller, 2017). This was the youngest cohort of pupils out of all of the studies in the present review, and the only study in which the predictions of the combined model were not supported, despite the combined model demonstrating a significant BFLPE for mathematics and a significant relationship between individual reading achievement and pupils' reading self-concept. However, the sample size in this study was considerably smaller than other studies in the current review, and the ASC measures arguably lacked sufficient variance due to adaptations which the authors employed which may have resulted in ceiling effects. Thus, the generalisability of these findings are questionable, and research should continue to investigate the relevance of the models to younger cohorts of pupils.

Given that elementary school is a formative stage for young pupils' ASCs (Guo, et al., 2018), future research may explore when pupils are likely to commence engaging in social and dimensional comparative processes which may adversely impact their ASCs. Other questions to consider might be: Are some pupils more likely to be influenced by these processes at a younger age than others? Do factors such as pupil ability level or learning difficulties accentuate or reduce the relative impact of social and dimensional comparisons on the development of pupils' ASCs? Such insights would be advantageous for advancing future knowledge in the area of ASC formation and be beneficial for informing future educational practices which take a developmental approach to fostering positive ASCs among all pupils. Furthermore, no study in the current review considered ability level as a potential moderator of social and/or dimensional comparison effects, although in relation to the potential impact of social comparisons on pupils' ASCs it has been advocated that this be investigated in

future research (Fang et al., 2018; Marsh et al., 2009). These insights could further inform future educational practices, such as when the optimum time might be to implement proactive approaches to combat the potential negative influence of social and dimensional comparison processes on the formation of pupils' ASCs.

2.4.3.3. Implementing Different Study Designs. The current research sought solely to investigate and synthesise findings from the combined model, as proposed by Chiu (2012). Therefore, research which incorporated additional theories into their analyses, such as the reciprocal effects model, were excluded from the current analysis (e.g. Marsh, 2018). Following the current review, it is advocated that future research also consider such extensions to the combined model to further our conceptualisation of ASC formation. For example, it was noted that future research may investigate the combined model and the reciprocal effects model simultaneously using longitudinal data (Guo et al., 2018; Pinxten et al., 2015; Kavanagh et al., 2019). Although this was considered to be outside the scope of the current review, Marsh et al.'s (2018) longitudinal study which investigated comparison processes on pupils' mathematics self-concept only may also be fruitful for gaining insight into ASC formation. In addition, literature has underscored the need for future research to consider investigating the models through implementing different types of study designs, such as qualitative research and/or experimental studies which employ trivial self-concept and achievement measures (Chiu, 2012).

2.4.3.4. The Combined Model and other Motivational Constructs. Finally, in line with the recommendation of Guo et al. (2018), it is advocated that the relevance of the combined model to other motivational constructs be explored. Only one study in the present review investigated the relevance of the basic I/E model, the BFLPE, and the combined model on a construct outside of ASC (Schurtz et al., 2014). This is perhaps unsurprising as the basic I/E and BFLPE models were initially posited to account for pupils' ASCs, and accordingly research has largely focused on the relevance of the models to pupils' ASCs. Schurtz et al. (2014) investigated the relevance of the three models to pupils' subject-interests. However, they employed a doubly-manifest approach to their data analysis, whereas recent studies have employed latent-manifest, or doubly-latent modelling approaches. Operationalising motivational constructs, such as ASC or subject-interest, as latent variables accounts for the measurement error of relevant scales (Marsh et al., 2009). In addition, while they controlled for individual level achievement (level one) when calculating effect sizes for class-average achievement (level two), their method of calculating the BFLPE was different

to the aforementioned studies which employed latent modelling approaches. Marsh et al. (2009) demonstrated that differences in the operationalisation of variables and computation of effect sizes within analyses can result in differences in the results. For example, in Marsh et al. (2009), it was found that effect sizes for the BFLPE were smaller for an analysis which employed a doubly-manifest approach, and larger for analyses of the same data which employed latent modelling approaches. Hence, in light of the limited research investigating the relevance of the combined model to pupils' subject-interests, and advancements in the typical types of analyses which are employed to investigate the model, it is advocated that future research investigate the combined model on other motivational constructs. Such studies should be computed in line with contemporary approaches to the analysis of the combined model as this will facilitate comparison of results and effect sizes between studies.

Chapter Three: Empirical Paper

3.1. Introduction

3.1.1. Motivation and Learning.

Pupils' self-beliefs and attitudes towards learning can have a critical impact on their learning behaviour and attainment (Baten et al., 2019; Dweck, 1999, 2006, 2014; Lee & Sue, 2015; Mitchell, 2014). Indeed, international research has demonstrated that pupils who report higher confidence in their learning abilities, higher interest in learning, and/or attribute greater value or usefulness to learning, typically perform higher on standardised performance tests (Mullis et al., 2017; Mullis et al., 2020; Mullis, 2007; Organisation for Economic Co-operation and Development, 2004, 2007, 2014). Therefore, it is positive to note that the development of positive attitudes and self-beliefs towards learning are now underscored as key objectives in both national and international curricula (National Council for Curriculum and Assessment [NCCA], 2018, 2019; Mullis, Martin, Ruddock, et al., 2009; Mullis, Martin, Kennedy, et al., 2009; Mullis & Martin, 2017, 2019).

3.1.2. Social and Dimensional Comparisons

Given the importance of fostering positive attitudes and dispositions towards learning, ample literature has focused on identifying factors which may be influential in shaping pupils' learning attitudes and beliefs. Two such factors or processes which have been widely discussed within the literature are that of comparisons which pupils make between their performance and that of their peers, also known as social comparisons, and comparisons which pupils make between their own individual performance in different academic domains, also known as dimensional comparisons. In particular, two prominent theoretical models emerged in the 1980s to account for how these processes may influence pupils' domain-specific academic self-concepts (ASCs), such as their reading self-concept, mathematics self-concept, etc. These are the big-fish-little-pond effect (BFLPE; Marsh, 1987; Marsh & Parker, 1984) and the internal/external frame of reference model (I/E model; Marsh, 1986). Within the past decade these two theories have been combined into a single framework, and for the purpose of the present paper this framework will be referred to as the combined model (Chiu, 2012). An overview of each of these models is provided below.

3.1.3. The Big-Fish-Little-Pond Effect

The BFLPE posits that pupils' ASCs are shaped by comparisons which they make between their performance and that of their peers (Marsh, 1987; Marsh & Parker, 1984; Thijs et al., 2010). Essentially, the theory advocates that the performance of peers is used as a reference point by the pupils when they construct their own ASCs. Consistent with this theory, ample research has demonstrated that while pupils' individual achievement in a particular subject area is positively correlated with their ASC in that area, class or school average achievement is negatively correlated with pupils' ASC in that area (Fang et al., 2018; Marsh & Hau, 2003).

3.1.4. The Internal/External Frame of Reference Model

First proposed by Marsh (1986), the I/E model theorizes that pupils' ASCs are shaped by comparisons which pupils make between their own, individual performances in different academic domains. Thus, achievement in one domain is used a reference point by the pupil when they construe their ASCs in both that domain and contrasting domains. Examples of contrasting domains which have been typically explored include mathematics and reading/English. Consistent with this theoretical background, the model predicts that achievement in one academic domain will be positively correlated with pupils' ASC in that domain, but to a lesser extent will be negatively correlated with pupils' ASC in a contrasting domain (see Figure 4). A robust research base has also investigated and supported the predictions of this model (Chiu, 2008; Marsh & Hau, 2004; Möller et al., 2009). Interestingly, it has been noted that although previous theory regarding the I/E model may have posited that the I/E model also accounts for external comparisons, such as grades and the performance of peers, this aspect of the model was overlooked in previous research which solely tested the aforementioned predictions of the basic I/E model (Chiu, 2012; Parker et al., 2013; Pinxten et al., 2015).

3.1.5. The Combined Model

Following three decades of literature which separately tested the BFLPE and the I/E model, Chiu (2012) proposed the combined model which, as the name suggests, integrated both of these models into a unified framework. Chiu (2012) reasoned that if both the predictions of the I/E model and the BFLPE were upheld within the unified model, this would provide evidence for the unique contribution of both social and dimensional comparison

processes on pupils' ASCs. Accordingly, it would also further our conceptualisation and understanding of factors which influence the development of pupils' ASCs. The predictions of this model can be seen in Figure 5. As can be seen in Figure 5, in addition to the central predictions of the BFLPE and I/E model, it has also been posited that the predictions of the combined model may be extended, such that class-average achievement in one domain may have a positive effect on pupils' ASC in a contrasting domain; however, this has been largely unsupported within research which has investigated the combined model to date (e.g. Guo et al., 2018; Kavanagh, 2019; Pinxten et al., 2015).

Considering the central predictions of the model (i.e. those which align with the basic BFLPE and I/E models), present literature has provided empirical evidence in support of the predictions of the combined model across a range of countries and age groups, from pupils in fourth-grade to pupils in their final year of secondary school (Chiu, 2012; Guo et al., 2018; Kavanagh, 2019; Marsh 1990c, 1994; Parker et al., 2013 Pinxten et al., 2015). However, as noted by Chiu (2012), the relative significance of the model to pupils' ASCs can tend to differ across countries. Consequently, the relevance of the models to particular countries or contexts should be considered when interpreting and using the findings to inform practice (Chiu, 2012). In addition, the strength of the correlations between the variables, and accordingly the relative significance of the model to pupils' ASCs, can differ based on the subject domains which are investigated (Guo et al., 2018). Most notably, research has consistently highlighted that the negative correlation between class-average achievement and pupils' ASC is most significant for mathematics (Chiu, 2012; Guo, et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015; Schurtz et al., 2014). Considering each of the aforementioned papers individually, the contextual effect (i.e. BFLPE) within the combined model was more pronounced for mathematics than for the other subject domains which were investigated.

3.1.6. The Combined Model and Subject-Interest

As the basic I/E model, the BFLPE, and the combined model were initially posited to account for factors which may impact pupils' ASCs, research to date has largely focused on the relevance of the models to pupils' ASCs. Concurrently, given the robust literature base which supports the combined model for pupils' ASCs, the question emerges as to whether the predictions of the combined model would also be pertinent to other motivational constructs. One such study was completed by Schurtz et al. (2014). Schurtz et al. (2014) investigated the

I/E model, the BFLPE, and the combined model on sixth-grade pupils' English and mathematics subject-interests. In their study, their conceptualisation of subject-interest was informed by the person-object-conception of interest (Krapp, 2002). This theory posits that interest is characterised by the relationship between a person and the object of interest, and this relationship is symbolised by the emotion or affect which an individual feels towards the object of interest and the personal value which they attribute to the object of interest. Accordingly, Schurtz et al.'s (2014) English and mathematics subject-interest measures pertained to pupils' intrinsic interest or affect towards engaging in the subjects, and the importance which they placed on learning the particular subject. Their results indicated that the central predictions of each model were significant. However, considering the practical significance of the findings in line with the effect size thresholds set out by Ferguson (2009), in which the smallest effect size threshold of .2 indicates that the findings have practical significance, none of the predictions which related to pupils' English subject-interest reached this threshold. The effect sizes for the correlations between individual achievement and class-average achievement on pupils' mathematics subject-interest were consistently small (i.e. > .2, < .5). In addition, Schurtz et al. (2014) investigated a model in which pupils' grades and ASCs were operationalised as mediating variables within the combined model, and this was also supported by their analysis. On the basis of a literature review on the combined model which was completed in advance of the present investigation, outside of Schurtz et al. (2014) no other study which investigated the relevance of the combined model to a motivational construct other than ASC was found.

3.1.7. ASC and Intrinsic Motivation

ASC refers to an individual's beliefs regarding their ability to learn, perform and achieve in particular academic domains (Marsh, 1986, 1990a, 1990b; Marsh et al., 1988; Marsh et al., 2006; Marsh and Shavelson, 1985). In earlier studies and measures examining ASC, the measurement scales commonly included affective items, such as 'I like [name of specific subject]'. However, the results of confirmatory factor analyses have since indicated that pupils' beliefs regarding their abilities and competencies are distinctive from affective components on ASC scales (e.g. Arens et al., 2013; Marsh et al., 1999). Accordingly, recent research, and in particular research pertaining to the theoretical models which are pertinent to the present research, has advocated that ASC scales pertain solely to pupils' beliefs regarding

their abilities and competencies (e.g. Guo et al., 2018; Lohbeck & Möller; Marsh et al., 2009; Pinxten et al., 2015).

Theoretically, literature has also distinguished between pupils' beliefs regarding their competencies, pupils' intrinsic interest or positive affect towards engaging in particular activities, and pupils' extrinsic motivation or beliefs regarding the importance of engaging in particular activities (Wigfield & Eccles, 2000). This understanding both aligns with and has informed the development of scales in international studies such as the Trends in International Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS; Mullis & Martin, 2013, 2015, 2017, 2019; Mullis, Martin, Foy, & Arora, 2012). In these studies, intrinsic motivation aligns with the conceptualisation of intrinsic motivation outlined by Deci and Ryan (1985). Thus, it is a force which drives pupils' behaviour, and pupils who are intrinsically motivated to learn demonstrate greater enjoyment and interest towards learning. Accordingly, measurement scale indicators pertain to pupils' interest and enjoyment towards engaging in domain-specific academic pursuits. In turn, it is posited that this disposition towards learning motivates pupils to learn more. Interestingly, results from the most recent TIMSS study found that intrinsic motivation was more predictive of academic achievement than extrinsic motivation, as indicated by the value or usefulness which pupils reportedly attributed to learning mathematics (Mullis, Martin, Foy, et al., 2020).

Consistent with the above literature, for the purpose of the current study intrinsic motivation refers to pupils' positive affect and interest towards engaging in academic pursuits in the relevant subject area; pupils' perceptions regarding the importance or utility of engaging in academic pursuits are not included. ASC, as outlined above, refers to pupils' self-beliefs regarding their academic competencies. While these constructs are clearly distinctive of each other, they are also understood to be related to each other within self-determination theory (SDT; Deci & Ryan, 1985; Ryan & Deci, 2017).

As outlined by Ryan and Deci (2017), a fundamental assumption underlying SDT is that people are innately motivated to grow, to learn, and to realise their full potential; however, this disposition may be facilitated or thwarted by the fulfilment or absence of three basic psychological needs which are autonomy, competence, and relatedness. In establishing different types of motivation, SDT recognises intrinsic motivation (i.e. whereby an individual is motivated to learn as they find learning to be inherently enjoyable, satisfying and

interesting) is the most advantageous for learning and results in enhanced pupil engagement and learning outcomes (Ryan & Deci, 2017, 2020). Consistent with SDT, pupils' intrinsic motivation may be enhanced through the fulfilment of the three basic psychological needs. The psychological need of competence, which as documented by Ryan and Deci (2017) refers to a pupil's belief in their ability to learn and succeed in their environment, may be regarded as closely related to the conceptualisation of pupils' ASCs. Thus, if pupils' ASCs are impacted by social and dimensional comparison processes, it follows that their intrinsic motivation may be similarly impacted by social and dimensional comparison processes. Moreover, any potential impact of social and dimensional comparison processes on pupils' intrinsic motivation may be mediated by pupils' ASCs. Indeed, this theoretical underpinning informed the research questions and models presented in the current paper, including the mediation effects model (section 3.1.9.5.; see Figure 7).

3.1.8. Operationalisation of the Combined Model.

To investigate the combined model, the predictions of the model are usually tested using a multilevel modelling approach to the data analysis (e.g. Kavanagh, 2019; Lohbeck & Möller, 2017; Parker et al., 2013; Pinxten et al., 2015). This approach accounts for the hierarchical nature of the class or school average achievement variables. It is worth noting that prior studies have differed in terms of how they operationalised the variables within the multilevel models. While earlier studies of the combined model, such as Chiu (2012) and Schurtz et al. (2014), employed a doubly-manifest approach, recently latent-manifest modelling approaches in which ASCs were construed as latent variables have been employed (Guo et al., 2018; Kavanagh, 2019; Lohbeck and Möller, 2017; Pinxten et al., 2015). As detailed by Marsh et al. (2009), operationalising unobserved variables, such as ASC, as latent constructs accounts for the measurement error of latent variable scale indicators.

Parker et al. (2013) employed a doubly-latent approach to analysing their data as they noted that their measure for school-average achievement, which contained 30 pupils per school, was not representative of the whole-school average achievement. Thus, in line with the recommendations of Marsh et al. (2009), latent aggregation of achievement variables was employed to account for the sampling error of the level two variables, in their case school-average achievement. Equally, it should be noted that the sampling error may also be minimised through implementation of a robust sampling design, such as by aiming for full

participation among sampled schools or classes (Guo et al., 2018; Kavanagh, 2019; Pinxten et al., 2015).

The operationalisation of contextual effect sizes (i.e. the BFLPE) has also varied between studies. However, recent studies have tended to implement a conservative approach to the estimation of contextual effects using the ES2 formula outlined by Marsh et al. (2009), $ES = 2 \times \beta \times SD_{predictor}/SD_{outcome}$ (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Parker et al., 2013; Pinxten et al., 2015). Use of a consistent approach to the operationalisation of contextual effect sizes across the aforementioned studies is advantageous for comparing contextual effect sizes between studies. Conversely, while earlier studies such as Chiu (2012) and Schurtz et al. (2014) controlled for level one effects when calculating level two effect sizes, they used a different approach to the aforementioned studies. As demonstrated by Marsh et al.'s (2009) research, differences in the operationalisation of effect sizes can result in differences in the observed effect sizes from analyses. Thus, this should be considered when comparing the contextual effect sizes from different studies.

3.1.9. The Present Study

Given the paucity of research investigating the combined model to motivational constructs other than ASC, the present study sought to investigate the relevance of the BFLPE, the I/E model, and the combined model to pupils' intrinsic motivation. The study also sought to investigate pupils' ASCs as potential mediating variables within the combined model. In contrast to Schurtz et al.'s (2014) study, the intrinsic motivation measure in the current analyses solely pertained to pupils' positive affect and interest towards engaging in the subject-specific academic pursuits; interest-related indicators pertaining to the importance or value which pupils may associate with the subject domain were not included. The present study also investigated the relevance of the models to Irish fourth-class primary school pupils' intrinsic motivation, as opposed to fifth-grade and sixth-grade pupils who just commenced secondary school in Germany, as in Schurtz et al.'s (2014) study. Furthermore, in line with advancing statistical analysis procedures for operationalising the combined model, a latent-manifest approach was employed to analyse the data and contextual effect sizes were calculated in line with recent literature (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh et al., 2009; Parker et al., 2013; Pinxten et al., 2015).

Consistent with the overall aims of the current research, a number of statistical analyses were completed to investigate the following research questions:

3.1.9.1. Research Question One. Are the predictions of the basic I/E model upheld when investigated in relation to fourth-class pupils' mathematics and reading achievement, and their intrinsic motivation towards these academic domains?

3.1.9.2. Research Question Two. Will the predictions of the BFLPE be upheld when tested in relation to pupils' mathematics achievement and their intrinsic motivation towards engaging in mathematics?

3.1.9.3. Research Question Three. Will the predictions of the BFLPE be upheld when tested in relation to pupils' reading achievement and their intrinsic motivation towards engaging in reading?

3.1.9.4. Research Question Four. Will the central predictions of the combined model be upheld when tested in relation to fourth-class pupils' mathematics and reading achievement, and their intrinsic motivation towards these academic domains?

3.1.9.5. Research Question Five. Within the combined model, will potential social and dimensional comparison effects on pupils' intrinsic motivation be mediated by pupils' ASCs?

3.2. Method

3.2.1. Participants

Data were sourced from the TIMSS and PIRLS 2011 Fourth Grade Combined International Database (Foy, 2013; TIMSS and PIRLS International Study Centre, 2019). This database was chosen as typically the PIRLS study occurs every five years while the TIMSS study occurs every four years. However in 2011, the year in which the studies were completed coincided which provided a unique opportunity for the results of both studies to be analysed and compared simultaneously (Mullis, Martin, Kennedy, et al., 2009). For countries in which the 2011 TIMSS and 2011 PIRLS were administered to the same cohort of pupils, the combined database was established, and it contains data for pupils who completed both of these assessments.

A total of 4383 Irish fourth-class pupils from 150 different schools and 220 different classes were included in this database. 157 participants were subsequently deleted as they

contained missing values for all items on at least one full subject-specific intrinsic motivation or self-concept scale. This left a total of 4226 participants from 150 different schools and 220 different classes in the sample which was analysed in the present study. As the analysis progressed, some additional classes of participants were deleted as it emerged that there was insufficient variance in participants' responses to the intrinsic motivation and/or self-concept scales within individual classes, which made data from those classes unsuitable for linear modelling/comparison of between-level and within level variance. These additional alterations to the sample size are outlined in the applicable sections below.

3.2.2. Design

3.2.2.1. Overall Study Design. Consistent with prior research, and the design of the 2011 TIMSS and PIRLS studies, the current study consisted of a secondary analysis of a cross-sectional dataset (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller., 2017; Parker et al., 2013; Pinxten et al., 2015; Schurtz et al., 2014).

3.2.2.2. Sample Recruitment in TIMSS and PIRLS. As documented by Joncas and Foy (2013) and Eivers and Clerkin (2012), for the 2011 TIMSS and PIRLS studies a stratified two-stage cluster sampling design was employed for the purpose of recruiting schools and participants. In the first stage, schools were selected from the national target population using a systematic random sampling approach, and the probability of schools being selected from the national sample was proportional to the size of the school. The sampling approach was also configured such that schools in areas with a high level of socioeconomic disadvantage, schools who provided education to a specific gender, and schools who taught through the Irish language would be selected and represented within the sample population in line with their representation within the national target population. The study design also stipulated that no more than five percent of potential participants, including schools, classes, and individuals, could be excluded from the target population for the purpose of selecting the nationally representative sample. Concurrently some exclusions were permitted due to factors such as the school solely providing education to pupils with intellectual, physical and/or language difficulties which may impair the pupils' ability to follow the fundamental instructions and process required to complete the test(s).

In the second stage of the process, classes from within selected schools were chosen at random, and with equal probabilities of being selected, using the WinW3S within-school

sampling software which was developed by the International Association for the Evaluation of Educational Achievement (IEA) and Statistics Canada. In addition, the sampling design aimed for a 100% participation rate among participating schools, classes, and pupils. To achieve this, high standards for participation rates were set that national samples were to adhere to for the sample to be accepted in the TIMSS and/or PIRLS.

3.2.2.3. Assessment Design. Matrix sampling designs, whereby assessments were divided into subtests and participants were assigned to complete particular subtests, were employed for both the mathematics achievement test in TIMSS and the reading achievement test in PIRLS (Foy, Brossman, et al., 2013). Therefore, while both the TIMSS mathematics assessment and the PIRLS reading assessment assessed pupils' abilities across a broad range of subject-specific competencies, individual participants completed only a subtest of assessment items from the overall item pool. As outlined by Foy, Brossman, et al. (2013), the rationale for this was to ensure coverage of a comprehensive range of subject-specific knowledge and skills among the target population, despite administration time constraints. Specifically, for participants to complete every item in each item pool, it would have required six hours for the reading assessment and eight and a half hours for the mathematics assessment (Mullis, Martin, Kennedy, et al., 2009; Mullis, Martin, Ruddock, et al., 2009). Therefore, as noted by the aforementioned authors, it would not have been appropriate to expect young pupils to concentrate and complete all assessment items simultaneously, and the design also had to be cognisant of the overall onus on participating schools and pupils.

Considering the matrix sampling design, when scaling, analysing, and obtaining overall achievement scores for participating countries and individual pupils, statistical procedures involving item response theory and plausible value theory were employed (Foy, Brossman, et al., 2013). Accordingly, in the TIMSS and PIRLS 2011 Fourth Grade Combined International Database five plausible, or estimated, values are provided for each participant's individual mathematics achievement score and for their individual reading achievement score. The plausible values are based on pupils' responses to the subtest of items which they completed, and on probability regarding how they may have performed on the overall assessment given their performance on the subtest that they completed and other background characteristics. In line with the achievement measures outlined below, participants' plausible values were used as the individual achievement measures in the current study. Further information regarding the sampling design and procedure involved in calculating individual achievement scores may be found in Foy, Brossman, et al. (2013).

3.2.3. Measures

3.2.3.1. Individual Mathematics Achievement. As outlined by Mullis, Martin, Ruddock, et al. (2009), the TIMSS mathematics assessment examined pupils' abilities in three content areas, which were number, geometric shapes and measures, and data display; and three cognitive areas which were knowing, applying, and reasoning. In total, the mathematics assessment item pool contained 14 blocks. The blocks were distributed between 14 test booklets, with each booklet containing two blocks, and each participant in TIMSS completed one of these booklets. As outlined above, five plausible values were computed for each pupil and these were used as the measure for individual mathematics achievement.

In addition to assessing a broad range of mathematical competencies, the TIMSS mathematics assessment is specifically designed for pupils in fourth-grade, considering pupils' age and international curricula expectations for pupils at this academic level (Mullis, Martin, Ruddock, et al., 2009). Mullis Martin, Ruddock, et al. also documented that the assessment was constructed in line with former cycles of the TIMSS mathematics assessment, which previously investigated international mathematics achievement among fourth-grade pupils. Indeed, some of the assessment items in the 2011 TIMSS mathematics assessment were taken directly from previous TIMSS assessment cycles. Good internal reliability was also reported for the TIMSS mathematics achievement measure. As noted by Foy, Martin, et al. (2013), the Cronbach's alpha reliability coefficient for the test was .82 among the international sample of fourth-grade pupils, and .84 among the Irish sample of fourth-class pupils.

3.2.3.2. Individual Reading Achievement. Pupils' overall achievement scores from the PIRLS assessment, as documented in the TIMSS and PIRLS 2011 Fourth Grade Combined International Database, were used as the individual reading achievement measure. As outlined in the PIRLS 2011 Assessment Framework (Mullis, Martin, Kennedy, et al., 2009), out of the complete item pool, five of the assessment items were literacy passages and another five were information texts. Pupils had 40 minutes to complete each item or 'testing block'. The ten testing blocks were distributed among 13 booklets, with each booklet containing two testing blocks and each pupil completing one booklet.

Consistent with the operationalisation of the TIMSS assessment, the PIRLS reading assessment was compiled in line with previous cycles of the PIRLS study, with some items in the 2011 assessment items being taken directly from previous PIRLS assessment cycles

(Mullis, Martin, Kennedy, et al., 2009). As outlined by Mullis, Martin, Kennedy, et al., the content was also specifically designed considering international curricula and the age of the target population. In addition, when choosing texts for the assessment, input was sought from educational professionals in participating countries and consideration was given to cultural factors which may impact pupils' performance. Accordingly, items which may require participants to unduly rely on culture-specific knowledge were omitted from the assessment. Good reliability was also reported for the PIRLS reading achievement measure. As reported by Foy, Martin, et al. (2013), Cronbach's alpha for the assessment was .88 among the international sample and .89 among the Irish sample.

3.2.3.3. Class-Average Mathematics Achievement. Class-average mathematics achievement was computed by aggregating pupils' individual mathematics achievement scores at the class-level (manifest aggregation).

3.2.3.4. Class-Average Reading Achievement. Class-average reading achievement was computed by aggregating pupils' individual reading achievement scores at the class-level (manifest aggregation).

3.2.3.5. Intrinsic Motivation for Mathematics. Indicators of intrinsic motivation for mathematics were taken from pupils' responses to the fourth-grade 'TIMSS 2011 Students Like Learning Mathematics Scale' (TIMSS and PIRLS International Study Centre, 2013a). This scale consisted of the following items: (1) I enjoy learning mathematics (R), (2) I wish I did not have to study mathematics, (3) mathematics is boring, (4) I learn many interesting things in mathematics (R), and (5) I like mathematics (R). Items with '(R)' after them indicate that the item was reverse coded. Pupils responded to each item on a four-point Likert scale: agree a lot, agree a little, disagree a little, disagree a lot. As with all scales in TIMSS and PIRLS, the scale is based on an underlying latent construct, in this case intrinsic motivation as it captures pupils' feeling of enjoyment and desire towards engaging in the activity of interest (Mullis, Martin, Foy, & Arora, 2012, 2013). Among the sample in the current study, the scale demonstrates good internal reliability ($\alpha = .89$).

3.2.3.6. Intrinsic Motivation for Reading. Indicators of intrinsic motivation for reading were taken from pupils' responses to six items on the 'PIRLS 2011 Students Like Reading Scale' (TIMSS and PIRLS International Study Centre, 2013b). The six items were (1) I read only if I have to, (2) I like talking about what I read with other people (R), (3) I would be happy if someone gave me a book as a present (R), (4) I think reading is boring, (5)

I would like to have more time for reading (R), and (6) I enjoy reading (R). Items with '(R)' after them indicate that the item was reverse coded. Pupils also indicated their responses to these scale items on a four-point Likert scale: agree a lot, agree a little, disagree a little, disagree a lot. Consistent with the intrinsic motivation for mathematics scale, the intrinsic motivation for reading scale consisted of indicators which related to pupils' enjoyment, interest and positive feelings towards engaging in reading activities (Mullis, Martin, Foy, & Drucker, 2012). The scale demonstrates good internal reliability among the current sample (α = .81). Further examination of the reliability of the scale can be found in the data analysis section below, which details a confirmatory factor analysis which was completed for all intrinsic motivation and self-concept scales in the present study.

3.2.3.7. Mathematics Self-Concept. Consistent with Guo et al. (2018), the self-concept measure consisted of five indicators from the 'TIMSS 2011 Students Confident in Mathematics Scale' (TIMSS and PIRLS International Study Centre, 2013c). The five items were (1) I usually do well in mathematics (R), (2) mathematics is harder for me than for many of my classmates, (3) I am just not good at mathematics, (4) I learn things quickly in mathematics (R), and (5) mathematics is harder for me than any other subject. Items with '(R)' after them indicate that the item was reverse coded. Pupils indicated their responses to these items on a four-point Likert scale: agree a lot, agree a little, disagree a little, disagree a lot. As with the reading self-concept items below, these items conform with items which are typically employed in contemporary research investigating ASCs, as they fundamentally relate to pupils' beliefs regarding their ability and competency in the specific academic domain (e.g. Kavanagh, 2019; Lohbeck & Möller, 2017). Based on their analysis of this scale among 15 countries from the 2011 TIMSS and PIRLS Combined International database, Guo et al. (2018) also reported good internal reliability and convergent validity for this scale. They reported that Cronbach's alpha was .83 (*SD* = .03), and that all factor indicators from a confirmatory factor analysis were greater than .5 which, as documented by Hair et al. (2014), demonstrates good convergent validity. Good internal reliability was also found for this scale among the current sample (α = .85).

3.2.3.8. Reading Self-Concept. The reading self-concept measure was also consistent with the items used by Guo et al. (2018), and contained four items from the 'PIRLS 2011 Students Confident in Reading Scale' (TIMSS and PIRLS International Study Centre, 2013d). The four items were: (1) I usually do well in reading (R), (2) reading is easy for me (R), (3) reading is harder for me than for many of my classmates, and (4) reading is harder for

me than any other subject. Items with '(R)' after them indicate that the item was reverse coded. Consistent with the format of the above scales, pupils responded to each item on a four-point Likert scale: agree a lot, agree a little, disagree a little, disagree a lot. As noted above, in line with prior research and current literature, the items are deemed to be valid ASC indicators. In Guo et al.'s (2018) research which used data from 15 countries from the TIMSS and PIRLS Combined International database, a confirmatory factor analysis found that all factor indicators on this scale were above .5, demonstrating good convergent validity. In addition, they reported that the scale demonstrated good internal consistency among their sample (α = .76, *SD* = .03). The scale also demonstrated good internal reliability among the current sample (α = .75).

3.2.4. Procedure

3.2.4.1. TIMSS Assessment. As outlined in the TIMSS 2011 Assessment Frameworks (Mullis, Martin, Ruddock, et al., 2009), and in line with the matrix sampling design, for the TIMSS assessment pupils completed one of 14 assessment booklets. One section of the booklet assessed pupils' mathematics performance, while the other section assessed pupils' science performance. Booklets alternated in the order in which science and mathematics assessments were presented, with seven booklets commencing with the mathematics assessment, and seven commencing with the science assessment. Pupils completed the assessment in the order in which the assessments were presented in the booklets. In total, pupils had 72 minutes to complete the TIMSS achievement assessments, 36 minutes for the mathematics assessment and 36 minutes for the science assessment. After 36 minutes, pupils were provided with a break before completing the second section of the booklet. Pupils were also provided with a break after they completed section two of the assessment, and subsequently they were provided with 30 minutes to complete the pupil questionnaire. The questionnaire contained the item indicators for the intrinsic motivation for mathematics and the mathematics self-concept measures noted above.

3.2.4.2. PIRLS Assessment. As documented in the measures section above, and consistent with the matrix sampling design, for the PIRLS assessment pupils completed one of 13 booklets (Mullis, Martin, Kennedy, et al., 2009). Each booklet contained two 40-minute testing blocks, each of which consisted of a literacy passage and 12-17 accompanying questions for participants to complete (Mullis, Drucker, et al., 2013). Following the assessment, pupils completed the pupil questionnaire, which required 15-30 minutes to

complete. In line with the TIMSS study, the pupil questionnaire contained the intrinsic motivation for reading and reading self-concept item indicators. Additional information on the TIMSS and PIRLS testing procedure can be found in the relevant assessment framework manuals (Mullis, Martin, Kennedy, et al., 2009; Mullis, Martin, Ruddock, et al., 2009).

3.3. Data Analysis and Results

3.3.1. Overview of Data Analysis

3.3.1.1. Data Preparation. Prior to completing the analyses, relevant scale items were reverse scored in SPSS. In the international database, Likert scale statements were coded as follows: agree a lot = 1, agree a little = 2, disagree a little = 3, disagree a lot = 4. Therefore, items that were positively phrased were reverse coded, such that positive scores resulted in higher intrinsic motivation and/or self-concepts among both positively and negatively worded items. Reverse coded items are followed by a (R) after the item statement in the intrinsic motivation and self-concept measures sections above. An exploration of missing data within the dataset, which was also completed in SPSS, revealed that no more than 3.7% of data was missing per variable, which given the large sample size was deemed to be relatively small and not of significant concern (Kline, 2011). Missing data were accounted for using full-information maximum likelihood, the default estimator in MPLUS (Muthén & Muthén, 2017).

3.3.1.2. Operationalisation of Analyses. A series of analyses, as outlined in the subsequent sections below, were completed in MPLUS software, version 8.4 (Muthén & Muthén, 2017). MPLUS syntax for each model can be found in Appendix G and Appendix H. Individual achievement was operationalised as a level one manifest variable. All intrinsic motivation and self-concept measures were operationalised as level one latent variables. The factor structure of these latent variables can be seen in the confirmatory factor analysis section below. In line with recent research, the level two predictor in multilevel models was class-average achievement, rather than school-average achievement, and this was operationalised as a manifest variable (e.g. Guo et al., 2018; Lohbeck & Möller, 2017; Kavanagh, 2019). It was reasoned that class-level effects may be most appropriate to investigate within the Irish context as within Ireland pupils in fourth-class are typically taught by a single teacher in a single classroom for the entirety of the school day. Accordingly, school-average achievement was not included in the multilevel models. The multilevel models comprised of two levels only, level one was individual achievement and level two

was class-average achievement. As some sampled classes were from the same school, the TYPE = COMPLEX TWO-LEVEL command was used in multilevel analyses to account for classes nested within schools.

The TYPE = COMPLEX and the TYPE = COMPLEX TWO-LEVEL commands were also employed for single-level and multilevel analyses respectively to account for the complex sampling design. For both single and multilevel analyses, the HOUWGT weighting variable was employed. This is suitable for use in pupil-level analyses, and has also been employed in multilevel models using the TIMSS and PIRLS 2011 Fourth Grade Combined International Database (Foy, 2013; Guo et al., 2018; Martin & Mullis, 2013). In addition, to account for violation of the assumptions of univariate and multivariate normality in intrinsic motivation and self-concept measures, the MLR estimator was used in all analyses; for all constructs in the measurement model, analyses of the factors and factor indicators indicated that the data was negatively skewed.

Considering the achievement variables, for multilevel models, pupils' individual achievement scores were grand mean centred within MPLUS. The TYPE = IMPUTATION function in MPLUS was used to account for the five plausible values provided for each participant for each achievement measure. Accordingly, for each model that was investigated five analyses were computed, one for each set of plausible values, and the average of the outcomes from the five analyses was then estimated. When running each model that was investigated, the analyses for each set of plausible values along with the derivation of the average model all occurred within MPLUS. As the TYPE = IMPUTATION function was employed, each model that was investigated only had to be run once for this to occur.

3.3.1.3. Effect Sizes. Consistent with contemporary literature, contextual effect sizes were estimated using the formula outlined by Marsh et al. (2009), using the MODEL CONSTRAINT command in MPLUS (e.g. Guo et al., 2018; Lohbeck & Möller, 2017). Specifically, this formula is $ES = 2 \times \beta \times SD_{\text{predictor}}/SD_{\text{outcome}}$. Effect sizes for all analyses were interpreted in line with Ferguson (2009), with the cut off thresholds of .2, .5, and .8 denoting small, moderate and strong effect sizes respectively. Ferguson (2009) also recommended that the minimum practically significant effect size for social science research is .2 for associations. This was considered to be most appropriate for interpreting the results from the current analysis as educational psychology research is typically used to inform

policy and practice, and the large sample size undermines the validity of significance thresholds (e.g. Ferguson, 2009).

3.3.2. Confirmatory Factor Analyses

Prior to completing the structural path analyses, a number of confirmatory factor analyses were computed to investigate the model fit and factor structure of the latent constructs in the measurement model, notably intrinsic motivation for reading, intrinsic motivation for mathematics, reading self-concept and mathematics self-concept. In the first model, a four-factor model, with intrinsic motivation for reading, intrinsic motivation for mathematics, reading self-concept and mathematics self-concept as the four factors, was computed. Comparison of the size of the correlations between positively and negatively worded factor indicators, both within and between scales, indicated that overall positively worded items were more correlated with positively worded items and negatively worded items were more correlated with negatively worded items. Correlations between factor indicators can be found in the correlation matrix in Appendix I. Examination of the modification indices also indicated that it may be beneficial to correlate error terms of positively and negatively worded items, an approach which was also employed by Guo et al. (2018) when they performed a similar secondary analysis using 2011 TIMSS and PIRLS data. For example, the largest modification index value for correlating error terms of factor indicators was for the item pair ‘I wish I did not have to study mathematics’ and ‘Mathematics is boring’. The modification index value indicated that correlating the error terms for these items would result in a change of 454.392 in the chi-squared test of model fit. Likewise, the modification index value for correlating the error terms for the positively phrased items ‘I usually do well in reading’ and ‘I usually do well in mathematics’ indicated that this specification within the model would result in a change of 199.950 to the chi-square test of model fit. In fact, the highest seven modification index values, all of which were greater than 125.00, followed this trend. Therefore, in the second model correlations between error terms of positively phrased indicators with positively phrased indicators, and negatively phrased indicators with negatively phrased indicators, were added to the model specification. Following this, model three built on model two and was computed to account for clustering at the class-level, the complex survey design, and accordingly the use of sample weights within the model.

In all models, factor variances were set to one to establish a scale for the measurement model, while all factor loadings of the indicator variables were estimated. As set out in Table 4 below, the model fit of each model was assessed in line with a number of model fit indices, namely the Standardized Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), Tucker–Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA). For comparison of fit across models, the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were also considered. Drawing on prominent literature, for CFI and TLI, values closer to or above .95 indicate good model fit, with values closer to one indicating better fit; for SRMR and RMSEA smaller values indicate better fit, with recommended values of close to or below.08 for SRMR and .06 for RMSEA indicating good model fit (Brown, 2015; Hu & Bentler, 1999; Kline, 2016). As noted by Kline (2016), when comparing models with different factor structures, smaller AIC and BIC values indicate better model fit. While chi-squared statistics are also reported along with p-values for the sake of completeness and to align with reporting standards from previous studies, for both the confirmatory factor analyses and subsequent models involving structural path analyses the p-values are largely irrelevant given the size of the sample under investigation.

Table 4

Model Fit Indices for Models One to Three

Model	χ^2 *df*	SRMR	CFI	TLI	RMSEA (90% CI)	AIC	BIC
1	3443.066 *df* = 164	.055	.884	.866	.069 [.067, .071]	193773.853	194192.887
2	719.084 *df* = 94	.033	.978	.955	.040 [.037, .042]	190341.626	191205.091
3	518.867 *df* = 94	.034	.979	.957	.033 [.030, .035]	188214.656	189078.122

Note. χ^2 = Chi-Square Test of Model Fit, in all three models $p < .001$; df = degrees of freedom; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

As can be seen in Table 4, consideration of model fit indices indicated that the final model provided the best fit for the data. In model three, SRMR, CFI, TLI, and RMSEA values all indicated good model fit (Brown, 2015; Hu & Bentler, 1999; Kline, 2016). Item fit indices in model three also demonstrated a good factor structure; factor loadings between each factor and its indicators can be seen in Table 5. All factor loadings were significant, in all cases $p < .001$. Hair et al. (2014) advocate that for convergent validity, all factor indicators should be above .5, and preferably above .7. As documented in Table 5, most factors loadings were in line with Hair et al.'s (2014) recommendations, and while four were below the .5 threshold, they did not fall below .4 and were included in subsequent analyses. Considering the size of the parameter estimates in line with Ferguson's (2009) criteria, all factor indicators were practically significant.

In addition, correlations between the four factors in model three were all below .8, demonstrating good discriminant validity (Brown, 2015). Correlations between the four

factors can be found in Table 6. As model three demonstrated the best model fit for the four factors, this model was specified in the measurement part of subsequent analyses.

Table 5

Estimated Factor Loadings for Indicators of Each Latent Construct, as per CFA Model Three

Indicator	Scale			
	Intrinsic Motivation Mathematics	Mathematics Self-Concept	Intrinsic Motivation Reading	Reading Self-Concept
1	0.888	0.705	0.432	0.755
2	0.643	0.604	0.431	0.795
3	0.743	0.707	0.676	0.495
4	0.683	0.711	0.764	0.492
5	0.929	0.683	0.722	
6			0.878	

Note. The number of each scale indicator corresponds with the number in the measures section of this paper. For all estimated factor loadings, $p < .001$.

Table 6

Model Three: Estimated Correlation Matrix for the Latent Variables

Latent Construct	Latent Construct			
	IMR	IMM	RSC	MSC
IMR	1.000			
IMM	0.223	1.000		
RSC	0.430	0.020	1.000	
MSC	0.118	0.675	0.215	1.000

Note. IMR = Intrinsic Motivation for Reading; IMM = Intrinsic Motivation for Mathematics; RSC = Reading Self-Concept; MSC = Mathematics Self-Concept.

3.3.3. Research Question One: The I/E Model

To investigate the basic I/E model, intrinsic motivation for mathematics was regressed onto individual mathematics achievement and individual reading achievement, and intrinsic motivation for reading was regressed onto individual reading achievement and individual mathematics achievement. Model fit statistics indicated that this model fit the data well (χ^2 (34) = 422.106, $p > .001$; SRMR = 0.043; RMSEA = 0.052; CFI = 0.968; TLI = 0.927). Correlations between the manifest and latent constructs indicated that mathematics achievement and reading achievement were strongly correlated with each other ($r = .787$), while there was only a small correlation between intrinsic motivation for mathematics and intrinsic motivation for reading ($r = .222$).

As can be seen in Table 7, the results showed that the pathway from reading achievement to intrinsic motivation for reading was both statistically and practically significant, and the size of this association was small. However, the pathway from mathematics achievement to intrinsic motivation for mathematics was neither statistically or practically significant. Cross-domain parameters between achievement and intrinsic motivation variables were also non-significant.

3.3.4. Research Question Two: The BFLPE for Mathematics

Preliminary analyses of multilevel models within MPLUS highlighted cluster invariance in participants' responses to intrinsic motivation measures within some classes. Therefore, at this stage in the analysis these classes were omitted from analyses of multilevel models. In total, 15 classes were deleted. This left a total of 141 schools, 205 classes, and 4118 pupils, 2023 of which were girls and 2095 of which were boys in the dataset. Participants ranged in age from 8.33 years to 12.17 years (M = 10.34, SD = 0.41). The SPSS syntax which was used to remove these classes from the data set, and which includes the identification numbers of classes which were removed, can be found in Appendix J.

To investigate the BFLPE for pupils' intrinsic motivation for mathematics, intrinsic motivation for mathematics was regressed onto level one individual mathematics achievement and level two class-average mathematics achievement. Model fit statistics demonstrated that this model fit the data very well (χ^2 (14) = 60.115, $p < .05$; CFI = .995; TLI = .989; RMSEA = .028). Results from the model, as documented in Table 7, were statistically significant and consistent with the predictions of the BFLPE. However, in line with Ferguson's guidelines, only the relationship between class-average mathematics achievement and intrinsic motivation for mathematics was practically significant (β = -.326). The relationship between individual mathematics achievement and intrinsic motivation was not (β = .133). Consistent with this interpretation, a small contextual effect was found for the BFLPE for mathematics (ES = -.234, p = .004).

3.3.5. Research Question Three: The BFLPE for Reading

Model fit statistics indicated that the BFLPE for reading did not fit the model as well (χ^2 (21) = 786.334, $p < .05$; CFI = .907; TLI = . 814; RMSEA = .094). In addition, no contextual effect for the BFLPE was found in the reading domain (ES = -0.002, p = .98). As can be seen in Table 7, while the beta coefficient for individual reading achievement on intrinsic motivation for reading was significant, the beta coefficient for class-average reading achievement on intrinsic motivation for reading was non-significant. SRMR model fit statistics also indicated a good model fit for the within part of this model (SRMR = .046), but not for the between part of the model (SRMR = .121).

3.3.6. Research Question Four: The Combined Model.

At this stage of the analysis, the value of achievement variables, which ranged from 204 to 829, were divided by 100. The rationale for this was to bring the variance between one and ten as large variances within the achievement measures and/or between the two scale types (i.e. survey and standardised tests) were resulting in convergence issues. Prior to this transformation, mathematics achievement measures were approximated to two decimal places within the dataset, while reading achievement measures were approximated to six decimal places. After the measures were divided by 100, mathematics achievement measures were approximated to four decimal places, while reading achievement measures were approximated to six decimal places.

Model fit statistics are summarised in Table 8. While CFI, TLI, and RMSEA indicated that the combined model may fit the data well (χ^2 (98) = 1285.199, $p < .05$; CFI = .953; TLI = .925; RMSEA = .054), AIC and BIC were higher for the combined model than for the I/E and BFLPE models. It is possible that the less accurate fit could be due to the between part of the model as SRMR was .046 for the within part of the model and .178 for the between part of the model. Results for the model can be seen in Table 7. While a small contextual effect was found for mathematics (*ES* = -.231, $p < .01$), no contextual effect was found for reading (*ES* = -.014, $p = .772$).

3.3.7. Research Question Five: Mediation Analysis

This model built on the combined model, as self-concept was included as a mediating variable. Level one indirect within-domain and cross-domain pathways were also included and tested within the model using the MODEL CONSTRAINT command in MPLUS. A diagram of this model can be seen in Figure 7. This diagram contains results for corresponding domain effects from the analysis. Arrows in Figure 7 also demonstrate cross-domain effects which were estimated, however, these effect sizes are not reported in Figure 7 as no cross-domain effects were found to be of practical significance.

Preliminary analysis for the mediation effects model highlighted additional cluster invariance in participants' responses, this time in relation to self-concept measures. Therefore, these classes were omitted from the analysis. This left a total of 140 schools, 200 classes and 4040 pupils in the analysis sample. A list of the class identification numbers of classes that were removed at this stage can also be found in Appendix J.

All model fit indices indicated that this model did not fit the data well (χ^2 (334) = 7895.568, $p < .05$; CFI = .826; TLI = .759; RMSEA = .075). Contextual effects were found for mathematics achievement on intrinsic motivation for mathematics (*ES* = -219, $p < .05$), mathematics achievement on mathematics self-concept (*ES* = -.362, $p < .05$,), reading achievement on intrinsic motivation for reading (*ES* = .736, $p < .05$), and reading achievement on reading self-concept (*ES* = -.275, $p < .05$). A small significant effect was also found for the mediation pathway between mathematics achievement, mathematics self-concept and intrinsic motivation for mathematics (*ES* = .317, $p < .05$). All other mediation pathways indicated a negligible effect based on effect size. All cross-domain direct pathways within the model were also non-significant on the basis of their effect size.

Following the mediation effects analysis, previous models were replicated on this dataset, which consisted of 4040 participants and in which achievement variables were divided by 100. The results from this can be seen in Appendix K; of note, the results from each model did not differ significantly from those which are reported in the current paper. MPLUS syntax that was employed for the I/E model, the BFLPE for mathematics, the BFLPE for reading, the combined model, and the mediation analysis can be found in Appendix H.

Table 7

Standardised Parameter Estimates (β) and Standard Errors (SE) for Each Model

Model	Intrinsic Motivation Reading β *(SE)*	intrinsic motivation Mathematics β *(SE)*
Basic I/E Model		
Reading achievement	.374 (.041)***	-.003 (.077)
Mathematics achievement	-.032 (.035)	.077 (.086)
BFLPE Mathematics		
Individual achievement		.133 (.023)***
Class-average achievement		-.326 (.106)**
BFLPE Reading		
Individual achievement	.365 (.036)***	
Class-average achievement	-.010 (.357)	
Combined Model		
Individual reading achievement	.411 (.039)***	-.103 (.085)
Individual mathematics achievement	-.036 (.034)	.222 (.088)*
Class-average reading achievement	-.069 (.303)	
Class-average mathematics achievement		-.328 (.110)**

Note. *$p < .05$, **$p < .01$, ***$p < .001$

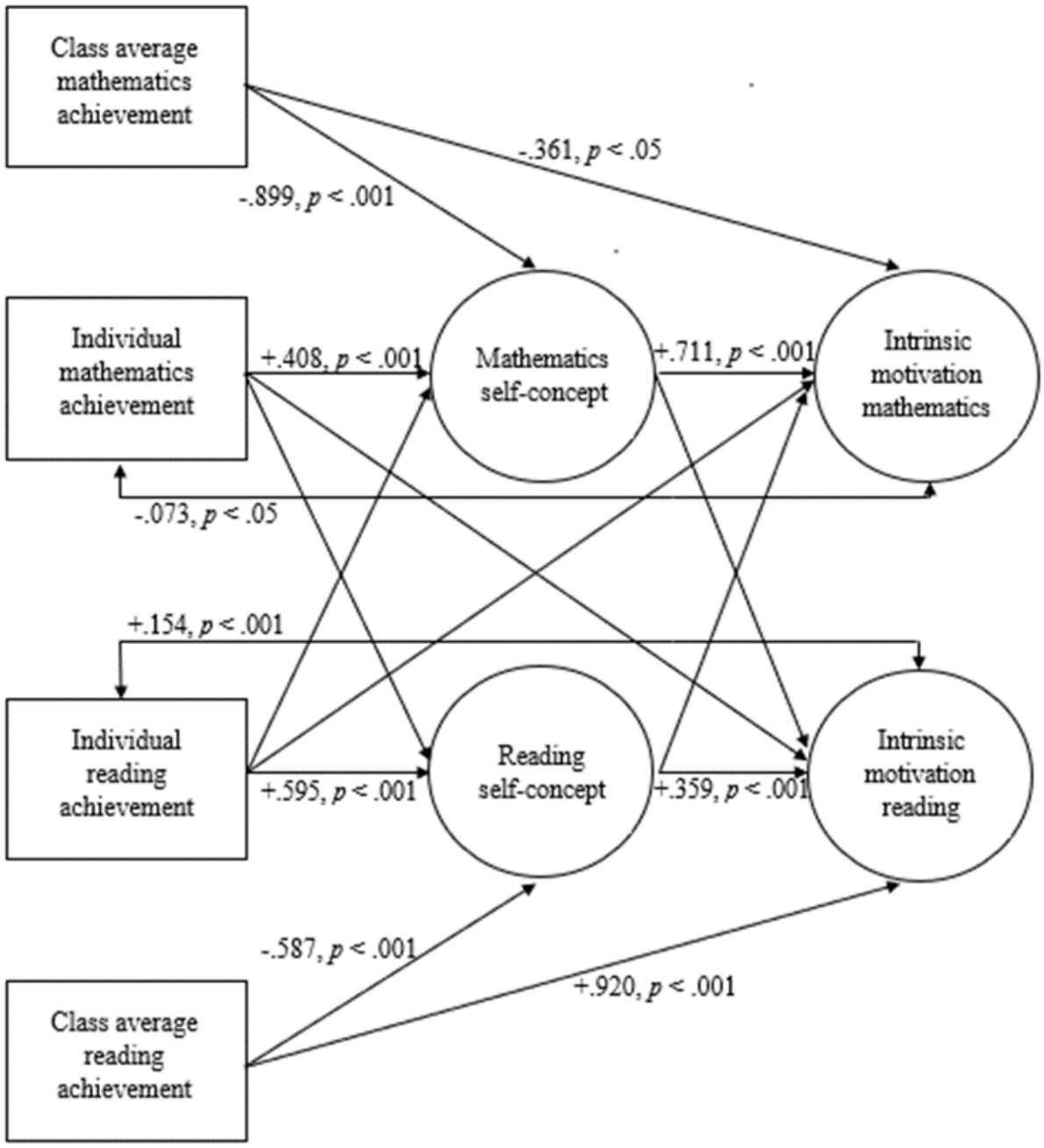
Table 8

Model Fit Indices for Models One to Five

Model	SRMR (*SD*)	CFI (*SD*)	TLI (*SD*)	RMSEA (*SD*)	AIC (*SD*)	BIC (*SD*)
Basic I/E	.043 (.001)	.968 (.003)	.927 (.006)	.052 (.002)	108187.410 (26.848)	108600.096 (26.848)
BFLPE M	.015 (<.001) [within] .047 (.002) [between]	.995 (<.001)	.989 (.001)	.028 (.001)	49019.166 (11.484)	49227.829 (11.484)
BFLPE R	.046 (.001) [within] .121 (.010) [between]	.907 (.004)	.814 (.814)	.094 (.002)	62048.533 (24.967)	62307.781 (24.967)
Combined	.046 (.001) [within] .178 (.003) [between]	.953 (.001)	.925 (.002)	.054 (.001)	122923.811 (56.245)	123549.800 (56.245)
Mediation	.115 (.003) [within] .354 (.002) [between]	.826 (.001)	.759 (.002)	.075 (<.001)	203256.233 (67.751)	204491.817 (67.751)

Figure 7

Mediation Effects Model



Note. Arrows in this diagram demonstrate both within-domain and cross-domain parameters which were estimated in the mediation effects model (research question 5). Standardised beta coefficients, along with *p* values for the sake of completion, are reported for all within-domain parameters. Standardised beta coefficients are not reported for cross-domain effects as all cross-domain effects were practically non-significant (i.e. $\beta < .2$).

3.4. Discussion

Overall, results from the present study demonstrate that the predictions of the combined model are not unanimously upheld when tested on Irish fourth-class pupils' mathematics and reading achievements and intrinsic motivation. However, considering each of the research questions and the associated analyses, noteworthy findings emerged which may inform our understanding of the relationship between individual achievement and class-average achievement on pupils' intrinsic motivation for mathematics and reading.

3.4.1. The Findings and Previous Research

3.4.1.1. The I/E Model. In the first analysis, the I/E model was tested in relation to pupils' intrinsic motivation. In contrast to the predictions of this model, only the relationship between individual reading achievement and pupils' intrinsic motivation for reading was practically significant, and the size of this effect was small. All cross-domain effects were both statistically and practically non-significant. Thus, these results do not coincide with previous research which traditionally investigated basic I/E model on pupils' ASCs (Marsh & Hau, 2004; Möller et al., 2009). The results also deviate from Schurtz et al.'s (2014) study who investigated the relevance of the I/E model to sixth-grade German pupils' subject-interests. Specifically, Schurtz et al. (2014) found that all parameter estimates within the I/E model were statistically significant, however, only the within and cross-domain effects on mathematics subject-interest reached a level which indicated that the findings had practical significance and these effect sizes were small (i.e. > .2, < .5).

3.4.1.2. The BFLPE for Mathematics. The second structural path analysis model investigated the BFLPE on pupils' intrinsic motivation for mathematics. Examination of model fit statistics indicated that this model fit the data very well, and that this model was the best fit to the data of all of the models which were computed. Both the parameters between individual achievement and intrinsic motivation, and class-average achievement and intrinsic motivation were statistically significant within the model. However, only the parameter between class-average mathematics achievement and intrinsic motivation was practically significant, and this effect size was small. In addition, a small contextual effect was found for class-average mathematics achievement on pupils' intrinsic motivation for mathematics (*ES* = -.234, p = .004). Therefore, the findings from this model coincide with a robust literature base pertaining to the BFLPE which has advocated, on the basis of similar BFLPE studies and theoretical underpinnings, that class or school average achievement has a negative impact on

pupils' ASCs (Fang et al., 2018). It also coincides with Schurtz et al. (2014) who demonstrated that class-average achievement is negatively correlated with pupils' subject-interest in mathematics. Thus, in line with prior theoretical and empirical literature, the results indicate that social comparison processes influence not only pupils' mathematics self-concept, but also their intrinsic motivation for mathematics.

3.4.1.3. The BFLPE for Reading. In contrast to the results which were found for the BFLPE for mathematics, the BFLPE model which investigated the relationship between individual reading achievement and class-average reading achievement on pupils' intrinsic motivation for reading produced a different pattern of results. The model did not fit the data as well, and model fit statistics indicated that this was due to between part of the model, as SRMR was .46 for the within part of the model but .119 for the between part of the model. No contextual effect was found for class-average reading achievement on pupils' intrinsic motivation for reading, and the pathway between class-average reading achievement and intrinsic motivation for reading was also non-significant. Therefore, as the predictions of the BFLPE model were not upheld when tested on pupils' intrinsic motivation for reading, this suggests that social comparisons which pupils may make between their performance and that of their classmates do not significantly impact pupils' intrinsic motivation for reading. Considering this in line with previous research, while Schurtz et al. (2014) found that both the parameters between individual English achievement and class-average English achievement on pupils' English subject-interest were statistically significant, the effect sizes were unlikely to be practically significant according to Ferguson's (2009) guidelines. Thus, their study likewise indicated that while class-average mathematics achievement had practical significance for pupils' mathematics subject-interest, class-average English achievement did not have practical significance for pupils' English subject-interest. In interpreting the findings from this model, it is also important to note that results from the within part of the current BFLPE for reading model are consistent with findings from the basic I/E model in this study. In both models, a small significant effect was found for the relationship between pupils' individual achievement and pupils' intrinsic motivation for reading.

3.4.1.4. The Combined Model. Consistent with the above findings from the I/E and BFLPE models, the predictions of the combined model in relation to pupils' intrinsic motivation for mathematics and readings were not fully supported in the current study. While CFI, TLI, and RMSEA model fit indices indicated that the model fit the data well, SRMR indicated a good model fit for the within part of the model (SRMR = .46) but not for the

between part of the model (SRMR = .178). However it is important to note that findings from the combined model were largely consistent with the aforementioned models. Specifically, small effects were found for individual reading achievement on intrinsic motivation for reading, and for class-average mathematics achievement on intrinsic motivation for mathematics. All cross-domain parameters were both statistically and practically non-significant. Considering contextual effects, a small contextual effect was found for class-average mathematics achievement on intrinsic motivation for mathematics, but no contextual effect was found for class-average reading achievement on pupils' intrinsic motivation for reading. In addition, within the combined model a small statistically and practically significant effect was found for individual mathematics achievement on intrinsic motivation for mathematics. However, this should be considered along with the previous analyses in this paper. Notably, this relationship was not found to be practically or statistically significant in the I/E model, or to be practically significant in the BFLPE model for mathematics which demonstrated the best model fit to the data. Thus, considering the results from all three models, there is limited support for the practical significance of this pathway.

3.4.1.5. The Combined Model with Pupils' ASCs as Mediating Variables. In the final model, pupils' ASCs were input as potential mediating variables within the combined model. However, all model fit indices indicated that this model did not fit the data well. In addition, as noted in the results section the predictions of this model were not fully upheld. All cross-domain effects lacked practical significance. Schurtz. et al. (2014) computed a similar model which indicated that the effects of class-average achievement and individual achievement on pupils' subject-interests were mediated by pupils' grades and pupils' ASCs. However, their model only fit the data when non-significant pathways were removed.

3.4.1.6. Synthesis of Findings from Each Model. Synthesising the results from each of the models in this study, the results demonstrate that while class-average mathematics achievement is predictive of pupils' intrinsic motivation for mathematics, class-average reading achievement is not predictive of pupils' intrinsic motivation for reading. Therefore, the BFLPE model may be considered relevant to fourth-class pupils' intrinsic motivation for mathematics but not for reading. Hence, in line with the theoretical underpinning of the BFLPE model, the results indicate that pupils' intrinsic motivation for mathematics is influenced by social comparisons which pupils make between their standard of achievement and that of their classmates. Concurrently, while the results suggest that similar social comparisons may not be influential on pupils' intrinsic motivation for reading, the results

highlighted that fourth-class pupils' individual reading achievement is predictive of pupils' intrinsic motivation for reading. Conversely, considered together the models provided limited evidence in support of a similar relationship between pupils' individual mathematics achievement and intrinsic motivation for mathematics. Thus, considering the theoretical underpinnings of the I/E model, the BFLPE, and the combined model, the results suggest that while social comparisons significantly influence pupils' intrinsic motivation for mathematics, pupils' individual achievement in reading is predictive of their intrinsic motivation for reading.

3.4.2. Limitations

3.4.2.1. The Time of Data Collection. Data for the 2011 TIMSS and PIRLS were collected between March and June of 2011 (IEA, 2013a, 2013b). Accordingly, this data reflects the responses of pupils who completed the assessments 10 years ago. Since 2011 new guidelines, curricula, and policies have been published in Ireland with the aim of furthering pupils' learning in mathematics and English and promoting pupils' wellbeing. For example, in 2011 a national strategy was published in Ireland which outlined provisions to increase Irish primary school pupils' standard of numeracy and literacy, and this included increasing the amount of time which pupils spend learning numeracy and literacy in school (Department of Education and Skills [DES], 2011). Following this, NEPS (2016, 2019) produced guidelines to promote the development of pupils' literacy skills, and more recently they produced guidelines on how to promote the development of pupils' mathematics skills (NEPS, 2020). The DES and the Department of Health (DH) have also produced guidelines for schools on the importance of promoting wellbeing and how this may be achieved (DES & DH, 2015; Government of Ireland, 2019). In addition, within this timeframe a new English language curriculum has been rolled out in Irish primary schools, and a draft version of a new primary school mathematics curriculum has been published (NCCA, 2018, 2019). Of particular note, each of these documents underscores the need to develop positive attitudes and dispositions towards learning, and the emphasis on these needs is increasingly evident in more recent publications. Therefore, although the robust sampling and implementation design indicate that this is a nationally representative sample, the generalisability of the findings should be considered in line with the time during which the data was collected.

3.4.2.2. Ceiling Effects. It is also possible that the results from the current analyses are impacted by ceiling effects. Preliminary analyses of all intrinsic motivation and self-

concept scales indicated that the data for these scales, and individual item indicators on the scales, were negatively skewed. In many cases, this resulted in invariance in participants' responses within individual classes, and as noted previously these classes were subsequently excluded from the analysis. Prior literature has suggested that such ceiling effects for self-concept measures may be due to younger pupils having higher and less differentiated ASCs than older pupils (Lohbeck & Möller, 2017). Although, noting that the predictions of the combined model were not upheld among their sample of second-grade pupils, on second-grade pupils' ASCs, Lohbeck and Möller (2017) also reported that their measure of ASC was an adapted version of the German version of the Self-Description Questionnaire (SDQ I-GS; Arens et al., 2013). Specifically, their measure pertained solely to pupils' beliefs regarding their competencies and therefore contained three scale items only, and the authors also reduced the Likert scale from five points to four points. Subsequently, they acknowledged that ceiling effects may have impacted their findings. Equally, it is possible that ceiling effects may be attributed to bias in participants' responses. For example, participants may have felt that they had to answer in a particular way in school or during the assessments, such that positive answers were more favourable than negative answers. Mertens (2015) outlined the need to consider such biases when completing educational research.

3.4.3. Scientific and Practical Implications

3.4.3.1. Scientific Implications. The findings from this research have important scientific implications. In the first instance, the results demonstrate that the predictions of the I/E model, the BFLPE for reading, and the combined model are not unanimously supported by the data. This suggests these models have limited scope in terms of providing a comprehensive framework which may inform our understanding of factors which influence Irish fourth-class pupils' intrinsic motivation. Conversely, the results further our understanding of factors which may influence pupils' intrinsic motivation. Of note, the current results add to a robust literature base which has highlighted a practically significant positive relationship between reading achievement and intrinsic motivation for reading (Logan et al., 2011; Wang & Guthrie, 2004). In addition, the results highlighted a contextual effect for class-average mathematics achievement on pupils' intrinsic motivation for mathematics. Specifically, as class-average achievement increases, pupils' intrinsic motivation for mathematics decreases. This highlights the relevance of the BFLPE to other motivational constructs for mathematics, outside of ASC.

Interestingly, previous literature which investigated the combined model on elementary school pupils has continuously found the contextual effect of class-average mathematics achievement on pupils' mathematics self-concept to be larger in size than the contextual effect for the other domain that was examined within the study (Chiu, 2012; Guo, et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015). In addition, although sparse research has investigated the relevance of the BFLPE or the combined model to motivational constructs outside of ASC, Schurtz et al. (2014) demonstrated that class-average achievement had a small, practically significant effect on pupils' subject-interest for mathematics, but not for English. Likewise Trautwein et al. (2006), who investigated the relevance of the BFLPE to pupils' subject-interest for mathematics, found that class-average achievement was negatively associated with pupils' interest for mathematics and that this relationship was mediated by pupils' ASCs. Similar to Schurtz et al. (2014), and unlike the current study, Trautwein et al.'s (2006) interest measure pertained to both pupils' feelings towards mathematics, and the personal importance which they placed on learning mathematics. Therefore, given the current findings in line with prior literature, it appears that the BFLPE may be particularly influential in understanding factors which may shape pupils' self-beliefs and attitudes towards mathematics.

3.4.3.2. Practical Implications. The results of the present analyses also have important implications for practice. Most notably, as the present study has underscored the significance of the BFLPE for mathematics, it is important that practices be employed to reduce the potential negative impact of social comparisons on pupils' intrinsic motivation for mathematics. For example, some practices which have been highlighted in prior literature for ameliorating the potential negative impact of social comparison effects in school include increased use of differentiated instruction and focusing on individual performance, or frames of reference, when teaching (Lüdtke et al., 2005; Roy et al., 2015). It is also possible that pupils' self-beliefs and attitudes towards mathematics may be more likely to be shaped and formed by their school experiences, such as the performance of peers in school, as pupils primarily experience learning mathematics in the school context (Schurtz et al., 2014). Conversely, pupils may be more likely to experience and engage in literacy activities in a range of contexts outside of school, such as when reading for pleasure or engaging in early literacy activities at home (Ginsberg et al., 2012; Schurtz et al., 2014; Skwarchuk, 2009). Indeed, given that intrinsic motivation can result in deeper and enhanced learning, it is important that all pupils are supported in developing a positive attitude and disposition

towards mathematics. Therefore, educational practitioners need to be aware of factors which may impact pupils' attitudes and of how such factors may influence pupils' attitudes.

Chapter Four: Critical Review and impact Statement

This chapter will critically evaluate the research process as a whole. This will commence with an overview of the epistemological position in which the research was situated. Next, ethical concerns which were encountered and addressed during the research process will be reflected upon. Following this, the research that is reported in the empirical paper will be critically appraised, and this will include critique of the design, methodology, and statistical analyses which were employed. Subsequently, implications of the findings from the current study for advancing psychological knowledge, for informing practice, and for informing future areas for research will be highlighted. Finally, the chapter will end with a critical impact statement which will underscore the relevance of the current research for public benefit, both inside and outside of academia.

4.1. Epistemological Position

The epistemological position in which this research is underpinned is the postpositivist paradigm. As outlined by Mertens (2015), the postpositivist paradigm arose in light of critique regarding the positivist paradigm. Essentially, the positivist paradigm advocates that human behaviour and social processes can be observed and studied objectively, akin to how subjects such as chemistry, physics, and geography which explore the natural world may be understood and studied (Mertens, 2015). However, while the postpositivist paradigm values objectivity, it recognises that the social world encompasses unobservable and subjective features, such as people's thoughts and feelings (Mertens, 2015). It also acknowledges that those who complete the research are influenced by their own theories, values, and hypotheses (Robson & McCartan, 2016). Accordingly, unlike the positivist paradigm which proposes that the world may be studied and understood with certainty, the postpositivist paradigm advises that human behaviour may only be understood probabilistically, and the researcher must both minimise and acknowledge potential bias when undertaking the research (Mertens, 2015; Robson & McCartan, 2016).

The postpositive paradigm was deemed to be most appropriate for the current research as in line with the ontology of this paradigm, the research was seeking to prove or disprove a particular psychological theory (Mertens, 2015). Specifically, the study sought to investigate if the predictions of the internal/external frame of reference model (I/E model; Marsh, 1986), the big-fish-little-pond effect (BFLPE; Marsh, 1987; Marsh & Parker, 1984), and the

combined model (Chiu, 2012) were upheld when tested in relation to Irish fourth-class pupils' intrinsic motivation for mathematics and reading. Therefore, to prevent my own prior knowledge, understanding, values, and theoretical disposition from impacting the results, it was important to maintain an objective stance when gathering and interpreting the data. This was achieved by using an international database, the data of which was collected by a school staff member other than the pupils' class teacher and by following standardised administration procedures (Eivers & Clerkin, 2012).

The quantitative methodology and ethical disposition that was employed throughout this research also aligned with the postpositivist paradigm. It has been noted that cross-sectional correlational research can provide information to inform reasonable causal inferences and accordingly, provide evidence for proving and/or disproving particular theories (Mertens, 2015; Thompson et al., 2005). In addition, Mertens (2015) documented that cross-sectional research is appropriate for investigating relationships between variables, such as academic achievement and intrinsic motivation, that would be unfeasible and/or unethical to manipulate in an experimental design. Further information regarding how ethics informed the current research is documented in section 4.2. below.

Consideration of the paradigm in which this research is underpinned is critical for facilitating optimum interpretation and application of the findings. The current research questions arose from examining prior research pertaining to the BFLPE, the I/E model, and the combined model. Therefore, it is acknowledged that the interpretation of the current findings is influenced by these models and the theoretical frameworks in which they are situated.

4.2. Ethical Considerations

Consistent with the axiology of the postpositivist paradigm, ethical considerations and practices were pivotal to informing the research process. In particular, the postpositivist paradigm emphasises the principles of beneficence, respect, and justice for all participants (Mertens, 2015). These principles guided the original research study and my subsequent actions when the COVID-19 pandemic disrupted provision of education for pupils, and consequently disrupted data collection for the original study (see section 1.4).

In the first instance, a challenge arose when a decision was made by the Department of Education and Skills (DES; 2020) to cancel all standardised tests in primary schools for the 2019/2020 academic year. Thus, participants in the original study would no longer

complete their second-class standardised tests (which were necessary as part of the original data collection procedure). Consequently, it was hoped that pupils' first-class test scores could potentially be used instead. However, participating pupils, their parents, teachers, and schools had not previously agreed to these scores being accessed for the purpose of the research. Therefore, in line with the first principle outlined in the PSI Code of Professional Ethics, 'Respect for the Rights and Dignity of the Person', informed consent to access these scores had to be sought (Psychological Society of Ireland, 2019). It was envisaged that this may be challenging due to the school closures, as school communities could no longer be physically accessed. In particular, it may be challenging to gain informed consent from pupils. Collaboration and resubmission to the Mary Immaculate Research Ethics Committee clarified that in light of the pandemic and the fact that previously consent had been sought from the individual pupils in the study, consent via email from parents and the school principal would suffice. After this clarification was gained, I still had concern that I would not be able to gain consent directly from the pupils themselves and contemplated whether or not seeking the results was ethically sound. As set out in Article 12 of the United Nations Convention on the Rights of the Child, children have a right to give their opinion regarding decisions which may impact them (United Nations, 2010). Hence, to capture the child's voice, I stipulated in the parent consent form that the use of first-class test scores be discussed with the child prior to parent(s)/guardian(s) providing consent (see Appendix L). After reflecting on the situation, I also concluded that accessing the scores would be highly unlikely to impact the pupils involved in any particular harmful way, as assessment scores and personal details would remain completely anonymous.

Unfortunately, an insufficient number of consent forms to access first-class test scores, and accordingly data, were gathered at the end of June 2020 to complete the intended analysis. Therefore, it was necessary to consider how the research may proceed. Given the impact of the pandemic on society, including disruption to school life and pupils' education, it was even more pertinent to consider the ethics regarding asking school communities to participate in the research at this time. Specifically, following the school closures, schools had to focus on reopening safely and meeting the emotional and learning needs of the children in their care post-lockdown (Government of Ireland, 2020). Thus, it was arguably unethical to ask schools to spend time engaging in the study at this time, and when contemplating how the project may continue this was a key consideration.

Consistent with the above concerns, use of the 2011 TIMSS and PIRLS combined international database, which was already established and accessible, was deemed to be a more viable option. Through use of this database the principle of beneficence was upheld. Specifically, this principle advises that harm to participants be minimised and the results of research maximised (Mertens, 2015). Thus, this principle was adhered to by using data which is readily available to education researchers, as no additional time was required from schools during the pandemic and data which was already available would be capitalised on. Participants' anonymity was also assured as all data which was obtained from the 2011 TIMSS and PIRLS database was anonymous, such that personal or identifying details regarding participating pupils, teachers, principals, parents, classes, and schools were not available. The studies were also reviewed and adapted appropriately by the Educational Research Centre for use within the Irish context, and schools were invited to participate prior to data collection (Eivers & Clerkin, 2012). Confirmation of consent to use this data was also sought via email from the International Association for the Evaluation of Educational Achievement (IEA) in June 2020 (see Appendix M).

4.3. Critical Appraisal of the Empirical paper

4.3.1. Design

As outlined in the empirical paper, a stratified two-stage cluster sampling design was employed to recruit participants in the 2011 TIMSS and 2011 PIRLS, and a matrix sampling design was employed for gathering assessment data (Eivers & Clerkin, 2012; Foy, Brossman, et al., 2013; Joncas & Foy, 2013). Hence, due to the aforementioned methods for data collection that were employed in the TIMSS and PIRLS 2011 studies, a noteworthy strength of the current research is that the dataset characterises a nationally representative sample. Thus, the findings may be considered as largely generalisable to fourth-class pupils in mainstream classes in Ireland.

The sampling design also stipulated criteria which ensured close to maximum participation rates among sampled schools, classes, and individual participants (Eivers & Clerkin, 2012; Joncas & Foy, 2013). Therefore, the sampling error for class-average achievement measures was minimized (Marsh et al., 2009). This was particularly advantageous for the present study, as in line with prior research and theory regarding the BFLPE and combined model, these level two measures were used as achievement indicators

or reference points which pupils may compare their performance to (Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015; Schurtz et al., 2014).

As documented by Eivers and Clerkin (2012) and Foy, Brossman, et al. (2013), the target population from which schools and participants were drawn excluded special schools, special classes, pupils attending private primary schools, and pupils in mainstream schools who may have been unable to complete the assessment. Pupils in mainstream schools who may have been excluded were pupils with physical disabilities, intellectual disabilities, or pupils for whom English was an additional language, whose corresponding level of need(s) would mean that they would be unable to follow the general directions or instructions required to complete the assessment. These excluded pupils were identified by school principals or other competent members of staff, and accounted for less than five percent of the overall target population in Ireland. Pupils excluded from mainstream classes amounted to less than 0.1 percent of the sampled classes in Ireland, and only two pupils were excluded from the studies as their parents refused to consent to the pupils' participation in the studies. Accordingly, while the current results may be generalisable among fourth-class pupils in mainstream classes in Ireland, caution should be exercised to the generalisation of the results to pupils who fall into categories which are outside of the sample population.

The consequences of implementing a matrix sampling design should also be considered. As discussed by Foy, Brossman, et al. (2013), matrix sampling designs, such as those employed in TIMSS and PIRLS, are advantageous for such international studies which seek to gain an indication of the overall national achievement standards in sampled countries. This is because the design can facilitate assessing a wide array of content among the sample population when the material is shared among participants and subsequently aggregated to obtain an overall national sample score. Thus, it is useful for gaining an accurate estimate of overall national achievement scores. However, considering individual achievement scores within these designs, each participant only completes a selection of items from the overall assessment item pool. Therefore, individual scores are not necessarily representative of the actual individual achievement scores which participants may have obtained had they completed the assessment in full themselves. To account for this in the 2011 TIMSS and PIRLS, plausible value methodology and item response theory were employed to provide estimates of pupils' individual achievement scores, and five plausible values were provided for each participant in the dataset. Thus, the five plausible values are based on pupils' individual performance on the achievement assessments, and other factors such as how their

performance and patterns of responses on the items they completed compared with the performance of other pupils with similar background characteristics. Thus, the individual achievement scores or plausible values used in the present study are estimated achievement scores, rather than pupils' actual attainment scores on the achievement assessments.

The current study was also cross-sectional, which has noteworthy implications regarding interpretation of the findings. As previously outlined, cross-sectional studies can be beneficial for investigating relationships between variables, and thus providing evidence which may conform or deviate from the predictions of particular theories. Accordingly, they provide evidence in support or against particular theories (Howitt & Cramer, 2011; Thompson et al., 2005). Building on the strong theoretical base which informed the combined model, much research to date which has investigated the combined model to pupils' ASCs has also been cross-sectional. Therefore, this design was advantageous for investigating the research questions in the empirical paper. As the design aligned with previous research in the area (e.g. Kavanagh, 2019; Lohbeck & Möller, 2017; Pinxten et al., 2015), this facilitated interpretation of the findings with previous research which investigated the relevance of the models to pupils' ASCs and subject-interests.

However, caution should be exercised when interpreting findings from cross-sectional studies (Marsh et al., 2009). Thompson et al. (2005) explained that findings from cross-sectional research may be inferred as effects if they conform with the predictions of particular theories. Conversely, it is important to remember that cross-sectional research solely provides evidence in support of correlations between variables. Causation between variables cannot be confirmed.

4.3.2. Measures

4.3.2.1. Achievement Measures. As detailed in the measures section of the empirical paper, achievement measures in the current study demonstrated good validity and reliability, both among the Irish sample and internationally. Comparing the TIMSS and PIRLS achievement assessment measures to those which were intended to be used in the original study, the original study aimed to use the following standardised assessments: For mathematics, the Level 2 Standardized Irish Graded Mathematics Attainment Test (SIGMA-T; Wall & Burke, 2007) or the Level 2 Drumcondra Primary Mathematics Test – Revised (Educational Research Centre, 2005), and for reading, the Level 2 Mary Immaculate Reading Attainment Test (MICRA-T; Wall & Burke, 2003) or the Level 2 Drumcondra Primary

Reading Test – Revised (Educational Research Centre, 2006). Each of these assessments was specifically designed to assess the achievement of Irish primary school pupils' achievement in the relative domains, and were normed on the Irish population. Arguably, the current achievement measures which were used instead are more advantageous as they were internationally valid and reliable and aligned with international standards for fourth-grade pupils. As noted previously, given that the assessments were to be used internationally, specific consideration was given to ensuring that the measures were free from potential cultural biases. In addition, prior to the assessments being administered to Irish pupils they were reviewed and adapted by the Education Research Centre in Ireland (Eivers & Clerkin, 2012). These minor adaptations to the assessments included changing some words and spellings so that they adhered to Irish norms, such as changing the spelling of 'color' to 'colour', and changing the word 'sidewalk' to 'footpath'.

Moreover, an additional advantage to using the achievement measures from an international database, as opposed to the standardized tests for primary schools in Ireland, is that only one achievement test for each subject was administered among the Irish population. In spite of the matrix sampling design, achievement scores were calculated using the same overarching assessment and the same statistical procedures. Thus, the results may be considered comparable to each other. Conversely, the use of the two different aforementioned standardized assessments per subject in Irish primary schools, which were normed separately on the Irish population, has been critiqued in Ireland (DES, 2016). The assessments were not designed to facilitate comparison between each other. As documented by the DES (2016), while the results between the two assessments per subject are approximately similar across schools, some variation between the results obtained by schools who administered different tests has been reported. Although this may be attributed to other factors such as higher-achieving or lower-achieving schools administering particular tests, the DES (2016) advocated that one standardized test be administered and used nationally in future. Therefore, using the international dataset measures in the present study overcame this concern which would have been a limitation in the original study.

4.3.2.2. ASC Measures. As noted in the methodology section of the empirical paper, ASC measures in the current study also demonstrated good validity and reliability. Concurrently, it was noted that items and data from ASC scales were negatively skewed, and thus the current findings may have been impacted by ceiling effects. This is something that was not widely noted in prior research, such as Guo et al. (2018) who used the same ASC

measures from the same international dataset but across 15 countries. However, considering the study which was originally proposed, the ASC measures consisted of six items and a five-point Likert scale (see Appendix N), as opposed to the four and five item ASC measures with four-point Likert scales which were used in the present investigation. Indeed, the scale in the original study was configured due to previous literature detailing that ASCs were less differentiated among younger pupils, and thus more sensitive ASC measures should be established when investigating younger pupils' ASCs (Lohbeck & Möller, 2017; Marsh et al., 1999). Due to using a pre-established database, these measures could not be operationalised within the design. It is important that the potential impact of ceiling effects in ASC measures be borne in mind when interpreting the findings from the current study.

It is also possible that ceiling effects may be attributed to bias which may have occurred due to factors such as how pupils perceived and/or responded to the task of completing the survey questionnaire. Leong and Austin (2006) and Howitt (2020) caution that researchers need to be aware of social desirability bias, that is bias which occurs due to participants wanting to provide socially acceptable and/or desirable responses, when designing and interpreting self-report measures such as questionnaires. In the context of the present study, it is possible that the pupils felt the need to respond in a socially desirable and favourable way in school. For example, they may have perceived the questionnaire to be similar to the attainment test and hence they may have wanted to provide the 'correct' response, which in school may be being a competent student who likes learning, to achieve a higher score. Accordingly, they may have responded 'agree a lot' to items such as 'reading is easy for me' or 'I learn things quickly in mathematics', even if they believed that these responses were not personally true for them. Similarly, pupils may have felt compelled to provide socially desirable responses to questionnaire items to please adults who may view or correct their tests/responses. In school, pupils' work and performance on standardised tests are normally evaluated and viewed by key adults in their lives such as their teachers, parents, and the school principal. Therefore, in the interest of depicting a positive self-image and being viewed positively by significant adults, pupils may have reported more socially desirable responses. Alternatively, potential bias in pupils' responses may even be unintentional if they were encouraged not to voice negative thoughts and feelings about their abilities previously. Thus, is it possible that ceiling effects in the current study are associated with bias in pupils' responses. Further implications for the consideration of such ceiling effects in future practice and research are documented in sections 4.4.2. and 4.4.3.4.

4.3.2.3. Intrinsic Motivation Measures. Intrinsic motivation measures were taken from the 2011 TIMSS and PIRLS context questionnaire scales. Fundamentally, motivational constructs which were assessed in these questionnaires align with expectancy-value theory, which posits that pupils' motivation to learn comprises of three aspects: pupils' self-beliefs regarding their abilities, pupils' intrinsic value towards the subject domain, and the usefulness or value which pupils place on learning the subject (Mullis et al., 2012; Wigfield & Eccles, 2000). As outlined by Wigfield and Eccles (2000), the intrinsic value component of their theory is consistent with Deci and Ryan's (1985) conceptualisation of intrinsic motivation, such that it pertains to the interest and enjoyment which people attribute to learning or engaging in a particular activity/subject. Indeed, the TIMSS and PIRLS 'students like reading' and 'students like mathematics' measures, updated versions of which have been employed in subsequent TIMSS and PIRLS studies, specifically align with Deci and Ryan's conceptualisation of intrinsic motivation (Mullis et al., 2012; Mullis & Martin, 2013, 2015, 2017, 2019).

Consistent with the above literature, the theoretical underpinning for the intrinsic motivation measure aligned with the theoretical base underpinning the TIMSS and PIRLS. Initially, it was considered that the alternative motivation measures align with pupils' subject-interests as conceptualised in the person-object-conception of interest theory outlined by Krapp (2002), as this theoretical framework was also used in previous research investigating the predictions of the combined model (Schurtz et al., 2014). On reflection, as such a measure would encompass both pupils' affect towards a particular subject and the personal importance which pupils attribute to the subject, it did not align with the scales which were employed in the TIMSS and PIRLS studies. In addition, exploring the relevance of the combined model to pupils' subject-interests may not have been as valuable to explore as it was investigated in previous research. Therefore, it was decided that the research investigate the predictions of the combined model primarily to the affect which pupils attributed to engaging in particular subjects (i.e. their intrinsic motivation).

One deviation of the intrinsic motivation measure defined in the current study from that in the PIRLS 2011 study, is that two items relating to how often pupils engaged in reading activities were omitted from the present study (Mullis, Martin, et al., 2012). The rationale for this was so that both the 'students like mathematics' scale and the 'students like reading' scale would solely pertain to pupils' affect towards engaging in particular subjects. The TIMSS measure did not contain similar items which related to how often pupils engaged

in a particular activity (Mullis, Martin, Foy, & Arora, 2012). Moreover, it is possible that pupils may frequently engage in reading activities outside of school regularly for reasons other than that they like reading. For example, they may be encouraged to read daily by their parents or they may read as they think that it is important to practice reading regularly. Thus, these two items were left out of the intrinsic motivation scale that was used in the present study.

It should be noted that intrinsic motivation measures in the current study were also negatively skewed. As critiqued in section 4.3.2.2. above in relation to ASC measures, this may be due to bias which may have occurred due to factors such as how pupils perceived and/or responded to the task of completing the survey questionnaire. Akin to the critique of the ASC measures, students may have felt compelled to provide socially desirable and favourable responses. For example, they may have felt obliged to please adults, and/or they may have been encouraged to withhold from voicing negative thoughts and feelings regarding their affect and interest towards mathematics or reading. This has implications for future practice and research, as documented in sections 4.4.2. and 4.4.3.4. It is important that the potential impact of ceiling effects associated with intrinsic motivation measures be taken into account when interpreting the current findings.

4.3.2.4. Data Analysis. The data analysis methods which were employed were a strength of the current research. The use of MPLUS software to complete the analysis for the combined model aligns with recent research and recommendations regarding the operationalisation of multilevel structural equation models (e.g. Guo et al., 2018; Kavanagh, 2019; Lohbeck & Möller, 2017; Marsh et al., 2009; Pinxten et al., 2015). Perhaps most notably, a latent-manifest approach was employed to analysing the data. This was preferable to a doubly-manifest approach to analysing the data, as it accounted for the measurement error in the intrinsic motivation and ASC scales (Marsh et al., 2009). A doubly-latent approach to analysing the data, in which class-average achievement was also construed as a latent variable, was also considered. However, it was rationalised that a latent-manifest modelling approach was advantageous as it was consistent with recent research, and due to the high participation rates among sampled schools and classes the sampling error was minimal.

Accurately accounting for the use of plausible values within the analyses was also given due consideration. In line with recent research (e.g. Guo et al., 2018) and the

recommendations of Rubin (1987), each of the analyses were first completed for each set of plausible values, such that five analyses were completed for each analysis. Thus, the first analysis was completed using pupils' first plausible value for mathematics and their first plausible value for reading, the second analysis was completed using pupils' second plausible value for mathematics and their second plausible value for reading, etc. This process was completed for each set of plausible values, and subsequently the results from each of the five analyses were combined to establish the overall results for each model. This methodology was implemented over a different approach which was employed by Chiu (2012), in which the mean of the five plausible values for each participant was computed prior to completing each analysis. Hence, only one analysis was completed for each analysis in Chiu's study. However, such an approach to analysing the data is not recommended by the IEA (2017), as it does not accurately account for the plausible values within the dataset.

Regarding the sample weights, in collaboration with my supervisors it was decided that the HOUWGT weighting variable be used for both single and multilevel analyses. Indeed, this approach was reasoned to be appropriate as it aligned with prior research which completed multilevel analyses using the 2011 TIMSS and PIRLS combined international dataset (Guo et al., 2018; Martin & Mullis, 2013). In addition, in January 2021 contact was made with the IEA's Co-Head of Research and Analysis Unit and Head of the Sampling Unit, Dr Sabine Meinck, for the purpose of gaining further insight into how the sample weights from TIMSS and PIRLS may be employed effectively within multilevel structural equation models. As outlined in the referenced conference notes provided by Dr Meinck, guidelines regarding optimum use of sample weights within multilevel models using data from TIMSS and PIRLS have yet to be established (Meinck, 2019). Accordingly, Meinck identified this topic as an area for future research. However, it was also noted that to date research involving multilevel analyses using the 2011 TIMSS and PIRLS dataset has employed the HOUWGT weighting variable. Thus, as the current analysis aligns with prior research and best practice literature for investigating the combined model, and ensured to account for specific features of large-scale complex survey data within the analyses, the data analysis was a strength of the current research.

4.4. Implications

4.4.1. Understanding and Knowledge in Psychology

The current research has important implications regarding advancing understanding and knowledge in psychology. In the first instance, the research analyses completed for the empirical paper sought to investigate if the predictions of the combined model were upheld when tested in relation to Irish fourth-class pupils' intrinsic motivation for mathematics and reading. Investigating the predictions of the combined model to pupils' intrinsic motivation was not explored in previous research. However, although prior literature found that the predictions of the model were upheld when tested in relation to pupils' ASCs in a number of different subject areas, the predictions were not unanimously upheld for fourth-class pupils' intrinsic motivation for mathematics and reading. Of particular note, there was a complete lack of cross-domain dimensional effects within the model. Considering this in line with the theoretical foundation of the combined model, the study advances current literature as while cross-domain dimensional comparison processes may impact pupils' ASCs, the current research indicates that cross-domain dimensional comparisons do not impact fourth-class pupils' intrinsic motivation. Hence, the current research indicates that the combined model is not an appropriate framework to consider when conceptualising how social and dimensional comparison processes may impact pupils' intrinsic motivation for mathematics and reading. Concurrently, considering the results from each analysis, important findings emerged. These findings in and of themselves also add to current understanding and knowledge in psychology and have important implications for practice.

As previously documented, the initial literature review highlighted that in all studies that investigated the combined model, the pathway between class-average achievement and pupils' ASC was consistently stronger for mathematics than for the other subject that was investigated in the model. Furthering this research base, the findings from the current empirical paper found that class-average achievement was significantly correlated with pupils' intrinsic motivation for mathematics, but not for reading. Hence, the current research highlights a significant negative relationship between class-average mathematics achievement and motivational constructs pertaining to mathematics, specifically pupils' mathematics self-concept, subject-interest, and intrinsic motivation. Consistent with the theoretical underpinning of the combined model, this implies that social comparison processes, which relate to comparisons which pupils make between their performance and that of peers,

influence pupils' self-beliefs and attitudes towards mathematics. Evidently, this understanding and awareness is valuable given the importance of such factors on pupils' learning, wellbeing, and performance. For instance, consistent with a robust theoretical and empirical evidence base, it is widely acknowledged that positive self-beliefs regarding one's abilities and engaging in activities and/or areas of interest are associated with better mental health (Gray, 2015; Kennerly et al., 2017; Stallard, 2005; 2019). Engaging in activities of interest is in itself, intrinsically rewarding and is associated with positive feelings (Gray, 2015). Additionally, when pupils feel confident in their abilities, they are more likely to engage in and to persist with new learning tasks and challenges which they encounter (Mitchell, 2014). Moreover, when people are intrinsically motivated to learn, such that they are interested in and enjoy learning, this can result in a deeper focus on, and engagement with, the learning material and accordingly deeper learning (Gray, 2015; Mitchell, 2014). Therefore, it is important that this new knowledge and understanding be used to inform future practice.

Furthermore, the current findings underscored a significant positive relationship between individual reading achievement and pupils' intrinsic motivation for reading. While this finding may have been expected in line with prior literature, considered alongside other information from the current analyses it provides insight into factors which influence pupils' intrinsic motivation in different subject areas. Specifically, the results indicate that pupils' intrinsic motivation for reading is significantly influenced by pupils' individual reading achievement, but it is not significantly influenced by class-average achievement or social comparison processes. Conversely, pupils' intrinsic motivation for mathematics is significantly negatively influenced by class-average mathematics achievement or social comparisons which pupils make between their performance and that of classmates, and intrinsic motivation for mathematics is not significantly influenced by pupils' individual mathematics achievement. The latter interpretation regarding the significance of individual mathematics achievement on pupils' intrinsic motivation for mathematics was taken from the BFLPE for mathematics model, as this model was the best fit of all models to the data. Clearly, these novel insights regarding factors which influence pupils' intrinsic motivation for mathematics and reading also have important implications for policy and practice in Ireland.

4.4.2. Implications for Practice

Intrinsic motivation is linked to advantageous learning outcomes, including increased engagement in the subject of interest and enhanced acquisition of knowledge and/or skills. Peoples' intrinsic motivation towards a particular area can also influence their choices and aspirations regarding their future careers (Boekeloo et al., 2015; Marusic, 2014). In addition, it is widely recognised that there is an increased need for mathematical skills within the workplace, and that higher levels of mathematics skills within a workforce can result in better outcomes for economies (Fritz et al., 2019; Lerner & Johns, 2015). Therefore, it is important that the impact of class-average mathematics achievement not only on pupils' mathematics ASCs, but also their intrinsic motivation is shared with educational practitioners so that these insights can inform practice.

Firstly, practitioners should be informed about the potential negative impact which attending a higher-achieving school may have on pupils' mathematics self-concept and intrinsic motivation for mathematics. Indeed, provision should be made to ensure that pupils are afforded opportunities that may enhance their beliefs regarding their mathematical skills and their intrinsic motivation towards mathematics. Concurrently, in contrast to the negative impact of class-average mathematics achievement on pupils' ASCs and intrinsic motivation, it has been proposed that attending a higher-achieving or prestigious school may actually enhance pupils' ASCs. This concept is also known as 'reflected glory', and was reviewed by Parker et al. (2013). However, Parker et al. (2013) concluded that although attending a higher-achieving school may initially have a positive impact on pupils' ASCs, it does not mitigate against the overall negative impact of social comparison processes on pupils' ASCs (i.e. the BFLPE). Therefore, it is necessary to consider specific processes and practices which may alleviate the negative impact of school-based social comparison processes on pupils' mathematics self-concepts and intrinsic motivation for mathematics.

Considering how pupils' mathematics self-concept may be positively fostered in school, recent best practice guidelines provided by NEPS (2020) for schools documented that a culture of confidence should be created to alleviate anxiety which pupils may experience towards mathematics. They noted strategies which may be employed to achieve this such as ensuring that pupils feel safe to engage in learning and to make mistakes, fostering positive adult-pupil relationships, and ensuring that potential anxiety which teachers may feel towards mathematics is not projected onto pupils. However, while these strategies primarily aim to

alleviate anxiety and promote pupils' confidence, the explicit need to consider the impact of social comparison processes on pupils' ASCs or motivation, and the potential value of promoting a culture of interest and enjoyment towards mathematics was not detailed in the guidelines. Conversely, within the NEPS (2019) guidelines for struggling readers, the importance of developing interest in reading both informally within the home environment and in school is emphasised. This is interesting to consider alongside the current findings in which class-average reading achievement was not significantly correlated with pupils' intrinsic motivation for reading. Although the NEPS reading (2019) and mathematics (2020) guidelines were published after data for the current analysis was completed, previous literature has also documented that there is more of a tendency towards and emphasis placed on engaging in reading activities for pleasure both inside and outside of school, than for mathematics. Consistent with Schutz et al.'s (2014) interpretation of their findings, this may explain why pupils' intrinsic motivation for mathematics may be more amenable to social comparison processes within the school context; as school is the principal place in which pupils' complete mathematics, it is the reference point which they use to construe their mathematics self-concepts and interests. Therefore, the importance of promoting pupils' intrinsic motivation for mathematics and engaging in mathematics activities for pleasure both inside and outside of school should be emphasised among educational and psychology practitioners, and among other key stakeholders in education such as parents and policymakers.

In practice, education practitioners (e.g. teachers) and/or psychology practitioners (e.g. educational psychologists) may seek to assess, monitor, and evaluate pupils' ASCs and/or intrinsic motivation, with the aim of supporting pupils to develop more positive ASCs and to experience higher levels of intrinsic motivation towards learning. They may choose to employ measurement scales which are typically used in research studies, such as the ASC and intrinsic motivation measures that were employed in the present study, to achieve this. Employing these measures may seem particularly advantageous in terms of their use as a pre- and post-measure. However, in light of the possibility of ceiling effects which were associated with these measures in the present study, caution should be highlighted to education and/or psychology practitioners who consider using such measures, and the benefit of employing a range of methods to gain a more comprehensive insight into pupils' ASCs and/or intrinsic motivation should be outlined (e.g. interviews with the pupil, behavioural observations, interviews with parents and teachers, etc.).

The current study also highlights noteworthy considerations regarding how schools may be structured. For example, within the German education system, pupils commence second-level schooling in fifth-grade, and attend one of three second-level schools (Schurtz et al., 2014). The second-level school which they attend largely depends on their academic ability level. Within such a stratified model, it may be argued that the impact of social comparison effects may adversely impact higher-achieving pupils. While such a system does not exist in Ireland, the standard of achievement within individual schools can differ based on the socioeconomic status of the families whose children attend the school. Concurrently, within Ireland there has been a continuous progression towards a more inclusive education system whereby all pupils', including pupils who historically may have attended a special school due to their learning differences, are educated alongside their peers in mainstream schools within their community (Griffin & Shevlin, 2011). As the current results indicate that higher class-average mathematics achievement is negatively correlated with pupils' intrinsic motivation for mathematics, it is possible that this more inclusive setting may actually have an overall positive effect on pupils' intrinsic motivation for mathematics in mainstream classrooms. This may be reasoned on the basis that including pupils with learning differences (e.g. pupils with diagnosed learning difficulties) within mainstream classes may lower the overall standard of mathematics achievement within the class. Thus, pupils' may compare their mathematics abilities to an overall lower general standard within their class, or alternatively to a broader continuum of performance which is more representative of all children within their community. However, it is also possible that this more inclusive setting may adversely impact the intrinsic motivation of pupils with learning differences who are educated within mainstream classrooms. As the overall standard of performance is likely to be higher in a mainstream classroom than in a special school/class, pupils with learning differences in mainstream classrooms may now be comparing their performance to an overall higher standard than they would have if they were in a special school/class. In turn, this may impact their intrinsic motivation towards learning mathematics. It is imperative that these points are considered in the ongoing inclusion debate within Ireland and when configuring and facilitating provision for lower-achieving pupils within mainstream schools.

4.4.3. Implications for Future Research

In addition to the areas for future research which were reported in Chapter 2, reflecting upon the research documented in the empirical paper also highlighted a number of

topics which would be pertinent for future research to explore. These topics were identified in light of the findings, strengths, and limitations of the current research.

4.4.3.1. The BFLPE and Mathematics. The current analyses highlighted that class-average mathematics achievement had a small effect on pupils' intrinsic motivation for mathematics. The same relationship was not evident between class-average reading achievement and pupils' intrinsic motivation for reading. Indeed, the second analysis in the empirical paper (i.e. 'Research Question Two: The BFLPE for Mathematics'), was the best fit to the data out of all of the analyses which were completed for the current research study. Prior research investigating the combined model has also shown that social comparison processes are particularly influential on pupils' mathematics ASCs. Moreover, research investigating the BFLPE and the combined model on pupils' subject interests has likewise indicated that social comparison processes can adversely impact pupils' mathematics subject-interests, and this effect may be mediated by pupils' ASCs. Given these findings, it is advocated that the BFLPE may be a promising framework to employ in future research when exploring how other motivational constructs pertaining to pupils' motivation and achievement in mathematics may be construed.

4.4.3.2. The Sample Cohort. As data from the chosen sample (i.e. the 2011 TIMSS and PIRLS) were collected in 2011, it is advised that the current research be replicated among a more recent sample of fourth-class pupils in Ireland. For example, researchers may consider examining the effect of the BFLPE on pupils' motivation for mathematics using the 2019 TIMSS data. In terms of investigating the I/E model and/or the combined model to pupils' intrinsic motivation, such an opportunity may arise among Irish fourth-class pupils when the year in which the TIMSS and PIRLS studies are completed coincide again.

Future research may also consider exploring the relevance of social and dimensional comparison effects on lower-performing pupils, such as pupils who attend special schools, special classes, or who present with significant learning differences in mainstream schools. This may be particularly pertinent to explore as these cohorts of pupils were excluded from the 2011 TIMSS and PIRLS studies. For further information on groups which were excluded from this sample, please see section 4.3.1.

4.4.3.3. Potential Developmental Effects. The current analysis was completed on a sample of fourth-class pupils. Fourth-class pupils were chosen as research had already found the BFLPE, the I/E model, and the combined model to be relevant to primary school-aged

pupils' ASCs, such as fourth-grade/class pupils in Flanders (Pinxten et al., 2015) and sixth-class pupils in Ireland (Kavanagh, 2019). Therefore, it was reasoned that if research demonstrated that the models were relevant to primary school-aged children's ASCs, they may also be relevant to their intrinsic motivation. The availability of data also had to be considered, and relevant data from fourth-class pupils only was available from the chosen dataset.

The possibility of completing a comparative study, whereby the model predictions that were examined in the current study be undertaken taken with two age cohorts (e.g. fourth-class and fifth-class pupils) was also contemplated when deciding on the focus for the current research. This would have been advantageous for exploring whether age or developmental effects may impact the findings. It is possible that findings may differ based on the age of the sample cohort as while previous research found social and dimensional comparison processes to be influential on primary school-aged pupils' ASCs, the researchers noted that their findings were typically smaller than findings from similar research which was completed with older pupils (e.g. Pinxten et al., 2015; Guo et al., 2018). A comparative study could not be completed for the purpose of the present analysis as it was not possible to source a dataset with the necessary data. For example, noting Kavanagh's analysis on a sample of sixth-class pupils from the National Assessments of Mathematics and English Reading study, it was proposed that the current research use that dataset as the study was also completed with second-class pupils. However, relevant intrinsic motivation and/or ASC measures were not available for sixth-class or second-class pupils. Similarly, a comparative study could not be completed using data from the 2011 TIMSS and PIRLS, as PIRLS is only completed with fourth-class pupils. It is recommended that future research consider undertaking a comparative study, as this has the potential to further our understanding of how age and/or developmental factors may impact pupils' intrinsic motivation. For instance, research may investigate the current models in relation to pupils' intrinsic motivation for mathematics and science using data from any year of TIMSS, as every cycle of TIMSS investigates both fourth-grade pupils and eighth-grade pupils.

4.4.3.4. Implementation of Alternative Measures. As critiqued in sections 4.3.2.2. and 4.3.2.3., ASC and intrinsic motivation measures were negatively skewed. Ceiling effects associated with these measures may have impacted the current findings. Consequently, it is recommended that future research ensure to establish and implement more sensitive ASC and intrinsic motivation measures. Potential examples of such ASC measures, which were

intended to be used in the original study with second-class pupils, may be found in Appendix N. These measures contain six item indicators and a five-point Likert scale, as opposed to the four/five item indicators and four-point Likert scales that were used to measure pupils' ASCs in the current analyses. Interestingly, more recent TIMSS and PIRLS studies have included more items on the relevant 'students like reading' and 'students like mathematics' scales (Mullis & Martin, 2019; Mullis, Martin, et al., 2020). Hence, it is possible that these updated scales may be more sensitive to capturing a more accurate measure of pupils' intrinsic motivation.

It is also advised that future research explore if pupils' responses to ASC and intrinsic motivation scales are likely to be subject to bias, and specifically how and what factors or types of biases may impact pupils' responses. In addition, such research may consider how potential biases in pupils' responses may be overcome. Implementation of alternative types of studies and measures may be helpful in this regard. For example, qualitative or mixed method studies may provide further insight into pupils' ASCs and/or intrinsic motivation via interviews with pupils, and/or pupils' parents and teachers. These could be considered alongside more sensitive quantitative measures, such as quantitative survey scales that are similar to the ASC and intrinsic motivation measures in the current study but contain additional scale items and a broader Likert scale. Overall, it is important that future research is cognisant of ceiling effects and the potential limitations of quantitative scale measures. Moreover, future research may seek to establish and implement more sound measures to overcome potential limitations, such as ceiling effects and response bias, that may be associated with current scale measures.

In addition to employing alternative ASC and intrinsic motivation measures, future research may also contemplate implementing alternative achievement measures. For example, school-based teacher assigned grades or evaluations. This may prove to be valuable as research investigating the combined model to date has primarily used standardized tests/ achievement measures.

4.4.3.5. Factors That Influence the BFLPE. A plethora of quantitative research has supported the BFLPE and theoretically it is posited that the BFLPE occurs due to comparisons which pupils make between their performance and that of their peers in their learning context (i.e., class/school; e.g. Fang et al., 2018; Marsh, 1987). However, as critiqued by Dai and Rinn (2008), as research which has examined the BFLPE is largely

based on cross-sectional, correlational studies, the actual cause of the BFLPE (i.e. contextual effect) may be due to factors other than social comparison processes. For example, the contextual effect may be attributed to the learning content and performance expectations that are set by teachers being higher within higher-achieving classes. Thus, pupils in higher-achieving classes may be comparing their performance to an overall higher standard, as established by the learning content that is presented to them and/or by the expectations set by their teachers, rather than comparing their performance to that of their peers. Drawing on theoretical and empirical evidence, Marsh et al. (2008) and Huguet et al. (2009) refuted Dai and Rinn's (2008) critique, advocating that the BFLPE primarily occurs due to social comparison processes, and this understanding continues to underpin the BFLPE. Although, as acknowledged by Marsh et al. (2008) and Huguet et al. (2009), it is plausible that other contextual factors and/or processes may also influence the BFLPE.

Considering what factors or processes may impact the BFLPE, it seems necessary to reflect on how pupils may 'tune in' to the average achievement levels within their learning context/class and their relative position. For instance, it has been speculated that competitive learning environments and/or discourse and feedback which focuses on comparisons between pupils may be likely to heighten pupils' attention towards their relative performance and that of their peers, and thus accentuate the BFLPE (Cheng & Lam, 2007; Marsh et al., 2008). Conversely less competitive environments which focus on individualised learning targets and gains, and in which discourse and feedback to students reflects this, may be likely to curtail the BFLPE (Marsh et al., 2008). This may be reasoned on the basis that if pupils are encouraged to tune into and compare their performance to their own personal level of attainment and progress, they may refer to this when evaluating their learning ability, rather than evaluating their performance/ability in comparison to that of their peers. Consistent with these propositions, it may be fruitful for future research to investigate class-level differences in competitiveness-discourse in a multi-level structural equation model to see if it influences the impact of the BFLPE.

Likewise, other school-based or classroom-based structural factors, such as ability-based grouping, may impact the BFLPE. Ability-based grouping is recommended for some instructional approaches recommended by NEPS for reading (2019), and arguably ability-based grouping can facilitate with differentiating learning content and ensuring that tasks are appropriately challenging in line with pupils' needs. This may promote the development of more positive ASCs as pupils experience a sense of accomplishment through completing

developmentally appropriate learning tasks. Alternatively, it seems plausible to speculate that ability-based grouping within classes/schools may reinforce particular perceptions and beliefs regarding pupils' relative standing within their learning environment and in turn their level of academic ability. For example, pupils may engage in processes such as noting who is in each group and comparing performances and learning between groups. Pupils may identify with their group and view their skills accordingly (e.g. 'the slower-paced group'). Teacher expectations and discourse regarding the groups may also have the potential to influence these perceptions. Consequently, these processes may enhance the size of the BFLPE.

As postulated above, there are several potential moderators which may influence the BFLPE, such as a competitive learning environment, discourse which focuses on comparing pupils' performances, and ability-based grouping. It is important that future research investigate the potential relevance of these factors to accentuating and/or thwarting the BFLPE. Moreover, future research should explore if these factors or processes are more pertinent to mathematics instruction than to other subject areas, as the current research underscored the heightened significance of the BFLPE to pupils' ASCs and intrinsic motivation for mathematics.

4.5. Impact Statement

Findings from the current study have several potential benefits for future research and practice. From a policy and practice perspective, the importance of fostering positive self-beliefs and attitudes towards learning is increasingly evident in new Irish curricula for mathematics and English (NCCA, 2018, 2019). It is also underscored in recent wellbeing guidelines for schools (Government of Ireland, 2019). Therefore, it is important that factors which may influence the development of pupils' learning attitudes and self-beliefs are identified. In turn, this can inform relevant guidelines and best practice literature regarding how to promote the development of positive learning attitudes and self-beliefs.

Of particular note, this research highlighted that social comparison processes may adversely impact pupils' intrinsic motivation for mathematics. This finding may be specifically relevant to understanding and informing future practice regarding how pupils' achievement, interest, and motivation to engage in mathematics may be promoted. These considerations are of importance for society at present given the increased demand for, and advantage of, a high standard of mathematical skills within the workforce (Fritz et al., 2019; Lerner & Johns, 2015).

This research study extends prior literature as it was the first to investigate the predictions of the combined model on pupils' intrinsic motivation. Although a similar study was completed by Schurtz et al. (2014) regarding the relevance of the model to pupils' subject-interests, the present study was completed on a sample of Irish pupils and using data analysis techniques which align with the most recent and recommended methods for analysing the combined model. While the predictions were not universally upheld, it may prove advantageous to consider the results in line with recent literature for the purpose of informing understanding and future research on the combined model.

The current study has also underscored a number of areas for future research. This includes investigating the relevance of social and dimensional comparison processes on younger pupils' subject-specific attitudes and self-beliefs, and implementation of alternative designs and measures. Thus, the research provides direction for future researchers who may be interested in pursuing research relating to the present field of study.

To realise the potential impact of the current research, it is intended that the insights and findings which emerged will be publicised in a number of ways. To date, an oral presentation which provided an overview of the proposed research in the empirical paper was presented at the 2020 Psychological Society of Ireland Annual Conference. It is envisioned that the findings, along with implications for future research and practice will continue to be disseminated at relevant conferences and/or platforms, and in my future professional practice when liaising with relevant partners such as teachers and parents. It is also envisaged that the empirical paper will be published in a relevant peer-reviewed journal article. Through communicating and discussing the insights of the current research in a variety of different ways and with a variety of different stakeholders, it is anticipated that the potential benefit of the present study both inside and outside of academia will be maximised.

References

Alexander, P. A., & Jetton, T. L. (1996). The role of importance and interest in the processing of text. *Educational Psychology Review, 8*(1), 89-121. https://doi.org/10.1007/BF01761832

Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology, 103*(4), 970-981. https://doi.org/10.1037/a0025047

Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2013). A short German version of the self-description questionnaire I: Theoretical and empirical comparability. *International Journal of Research & Method in Education, 36,* 415–438. http://dx.doi.org/10.1080/1743727X.2012.710503

Baten, E., Pixner, S., & Desoete, A. (2019). Motivational and math anxiety perspective for mathematical learning and learning difficulties. In A. Fritz, V. G. Haase, & P. Räsänen (Eds.), *International Handbook of Mathematical Learning Difficulties* (pp. 457-467). Springer International Publishing.

Boekeloo, B. O., Jones, C., Bhagat, K., Siddiqui, J., & Wang, M. Q. (2015). The role of intrinsic motivation in the pursuit of health science-related careers among youth from underrepresented low socioeconomic populations. *Journal of Urban Health, 92*(5), 980-994. https://doi.org/10.1007/s11524-015-9987-7

Brown, T. A. (2015). *Confirmatory factor analysis in applied research* (2nd ed.). The Guilford Press.

Carvalho, E., & Skipper, Y. (2020). A two-component growth mindset intervention for young people with SEND. *Journal of Research in Special Educational Needs, 20*(3), 195-205. https://doi.org/10.1111/1471-3802.12472

Center for Self-Determination Theory [@centerforSDT]. (2017 November 29). *We're giving the motivation continuum an updated look (aka the organismic integration theory taxonomy of regulatory styles) #scholarslikewords #wednesdaywork thoughts?* Twitter. https://twitter.com/centerforsdt/status/935914261746016256

Cheng, R. W., & Lam, S. (2007). Self-construal and social comparison effects. *British Journal of Educational Psychology, 77*(1), 197-211. https://doi.org/10.1348/000709905X72795

Chiu, M. (2008). Achievements and self-concepts in a comparison of math and science: Exploring the internal/external frame of reference model across 28 countries. *Educational Research and Evaluation, 14*(3), 235-254. http://dx.doi.org/10.1080/13803610802048858

Chiu, M. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology, 104*(1), 87-107. http://dx.doi.org/10.1037/a0025734

Clerkin, A., Perkins, R., & Cunningham, R. (2016). *TIMSS 2015 in Ireland: Mathematics and science in primary and post-primary schools*. Educational Research Centre. https://www.erc.ie/wp-content/uploads/2016/11/TIMSS-initial-report-FINAL.pdf

Dai, D. Y., & Rinn, A. N. (2008). The big-fish-little-pond effect: What do we know and where do we go from here? *Educational Psychology Review, 20*(3), 283-317. https://doi.org/10.1007/s10648-008-9071-x

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer.

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 2011-2021*. https://www.education.ie/en/publications/policy-reports/lit_num_strategy_full.pdf

Department of Education and Skills. (2016). *Standardised achievement tests: An analysis of the results at primary school level for 2011-12 and 2012-13.* https://assets.gov.ie/24671/373257000c4d4c15a36567827997a3ad.pdf

Department of Education and Skills. (2020). *Standardised testing in primary schools 2019/2020.* https://www.education.ie/en/Schools-Colleges/Services/Returns/Supporting-Assessment/standardised-testing-in-primary-schools-2019-2020.pdf

Department of Education and Skills., & Department of Health. (2015). *Guidelines for mental health promotion: Well-being in primary schools.* https://www.education.ie/en/Publications/Education-Reports/Well-Being-in-Primary-Schools-Guidelines-for-Mental-Health-Promotion.pdf

Durik, A. M. (2017). Academic self-concept. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopaedia of Personality and Individual Differences*. Springer. https://doi.org/10.1007/978-3-319-28099-8_527-1

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success.* Random House.

Dweck, C. S. (2014, November). *The power of believing that you can improve* [Video]. TED. https://www.ted.com/talks/carol_dweck_the_power_of_believing_that_you_can_improve#t-506975

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78-89. http://dx.doi.org/10.1080/00461520902832368

Educational Research Centre. (2005). *The Drumcondra primary mathematics test – revised* (2nd ed.). Educational Research Centre.

Educational Research Centre. (2006). *The Drumcondra primary reading test – revised* (2nd ed.). Educational Research Centre.

Eivers, E., & Clerkin, A. (2011). PIRLS & TIMSS 2011: *Reading, mathematics and science outcomes for Ireland*. Educational Research Centre. https://www.erc.ie/documents/pt_2011_main_report.pdf

Eivers, E., & Clerkin, A. (2012). *TIMSS and PIRLS 2011: Technical report for Ireland.* Educational Research Centre. https://www.erc.ie/documents/pt_2011_technical_report.pdf

Eivers, E., Gilleece, L., & Delaney, E. (2017). *Reading achievement in PIRLS 2016: Initial report for Ireland*. Educational Research Centre. https://www.erc.ie/wp-content/uploads/2017/12/PIRLS-2016_inital-report-IRL.pdf

Fang, J., Huang, X., Zhang, M., Huang, F., Li, Z., & Yuan, Q. (2018). The big-fish-little-pond effect on academic self-concept: A meta-analysis. *Frontiers in Psychology, 9*. http://dx.doi.org/10.3389/fpsyg.2018.01569

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology, Research and Practice, 40*(5), 532-538. http://dx.doi.org/10.1037/a0015808

Field, A. P. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE.

Field, A. (2000, 05 24). SEM. Discovering statistics. https://www.discoveringstatistics.com/repository/sem.pdf

Foy, P. (2013). *TIMSS and PIRLS 2011 user guide for the fourth grade combined international database*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/timsspirls2011/downloads/TP11_UserGuide.pdf

Foy, P., Brossman, B., & Galia, J. (2013). *Scaling the TIMSS and PIRLS 2011 achievement data* (M. Martin & I. V. A. Eds.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/methods/pdf/TP11_Scaling_Achievement.pdf

Foy, P., Martin, M. O., Mullis, I. V. S., & Stanco, G. (2013). *Reviewing the TIMSS and PIRLS 2011 achievement item statistics* (M. Martin & I. V. A. Eds.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/methods/pdf/TP11_Reviewing_Achievement.pdf

Fritz, A., Haase, V. G., & Räsänen, P. (2019). *International Handbook of Mathematical Learning Difficulties: From the Laboratory to the Classroom*. Springer International Publishing. Kindle Edition.

Froiland, J. M., & Worrell, F. C. (2016). Intrinsic motivation, learning goals, engagement, and achievement in a diverse high school. *Psychology in the Schools, 53*(3), 321-336. https://doi.org/10.1002/pits.21901

Ginsburg, H. P., Duch, H., Ertle, B., & Noble, K. G. (2012). How can parents help their children learn math? In B. H. Wasik (Ed.), *Handbook of family literacy* (pp. 51–65). Routledge/Taylor & Francis Group.

Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education, 22*(2), 213-228. http://dx.doi.org/10.1080/02671520701296189

Government of Ireland. (2019). *Wellbeing policy statement and framework for practice 2018–2023.* Department of Education and Skills. https://www.education.ie/en/Publications/Policy-Reports/wellbeing-policy-statement-and-framework-for-practice-2018%E2%80%932023.pdf

Government of Ireland. (2020). *Reopening our primary and special schools.* https://www.gov.ie/en/publication/a0bff-reopening-our-primary-and-special-schools/

Gray, P. (2015). *Free to learn: Why unleashing the instinct to play will make our children happier, more self-reliant, and better students for life*. Basic Books.

Griffin, S. & Shevlin, M. (2011). *Responding to special educational needs: An Irish perspective* (2nd ed.). Gill & Macmillan.

Guo, J., Marsh, H. W., Parker, P. D., & Dicke, T. (2018). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learning and Instruction, 58*, 210-219. http://dx.doi.org/10.1016/j.learninstruc.2018.07.007

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology, 51*(8), 1163-1176. http://dx.doi.org/10.1037/a0039440

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th Ed.). Pearson.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*(4), 549-571. https://doi.org/10.3102/00346543060004549

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4

Howitt, D., & Cramer, D. (2011). *Introduction to research methods in psychology* (3rd ed.). Prentice Hall.

Howitt, D. (2020). *Research methods in psychology* (6th ed.). Pearson.

Hox, J. J., & Bechger, T. M. An introduction to structural equation modeling. *Family Science Review, 11*, 354-373.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. http://dx.doi.org/10.1080/10705519909540118

Huguet, P., Dumas, F., Marsh, H., Régner, I., Wheeler, L., Suls, J., Seaton, M., & Nezlek, J. (2009). Clarifying the role of social comparison in the big-fish-little-pond effect (BFLPE): An integrative study. *Journal of Personality and Social Psychology, 97*(1), 156-170. https://doi.org/10.1037/a0015558

IEA. (2013a). *PIRLS 2011: Progress in international reading literacy study*. IEA TIMSS and PIRLS international study centre. https://timssandpirls.bc.edu/pirls2011/schedule.html

IEA. (2013b). *TIMSS 2011:Trends in international mathematics and science study*. IEA TIMSS and PIRLS international study centre. https://timssandpirls.bc.edu/timss2011/schedule.html

IEA. [IEA Education]. (2017, April 26). *The IEA's IDB Analyzer: Introduction* [Video]. Youtube. https://www.youtube.com/watch?v=cI_wkadWiYs&feature=emb_logo

Joncas, M., & Foy, P. (2013). *Sample design in TIMSS and PIRLS* (M. Martin & I. V. A. Eds.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/methods/pdf/TP_Sampling_Design.pdf

Kadir, M.S., & Yeung, A.S. (2016). Academic self-concept. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopaedia of Personality and Individual Differences*. Springer. https://doi.org/10.1007/978-3-319-28099-8_1118-1

Kavanagh, L. (2019). Academic self-concept formation: Testing the internal/external frame of reference model, big-fish-little-pond model, and an integrated model at the end of primary school. *European Journal of Psychology of Education, 35,* 93–109. http://dx.doi.org/10.1007/s10212-019-00416-w

Kennerly, H., Kirk, J., Westbrook, D. (2017). *An introduction to cognitive behaviour therapy: Skills & applications* (3rd ed.). SAGE Publications.

Kline, E. T. (2011). *Principles and practices of structural equation modeling* (3rd ed.). The Guilford Press.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Krapp, A. (1993). The construct of interest: Characteristics of individual interests and interest-delated actions from the perspective of a person-object-theory. *Studies in Educational Psychology, 4.* Universitat der Bundeswehr.

Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education, 14*(1), 23-40. https://doi.org/10.1007/BF03173109

Krapp, A. (2002). An educational-psychological theory of interest and its relation to self-determination theory. In E. L. Deci, & R. M. Ryan (Eds.), *The handbook of self-determination research* (pp. 405 - 427). University of Rochester Press.

Lee, C., & Sue, J. W. (2015). Mathematical resilience: What is it and why is it important? In S. Chinn (Ed.), *The routledge international handbook of dyscalculia and mathematical learning difficulties* (pp. 337-345). Routledge.

Leong, F. T. L., & Austin, J. T. (2006). *The psychology research handbook: A guide for graduate students and research assistants* (2nd ed.). Sage Publications.

Lerner, J. W., & Johns, B. H. (2015). *Learning disabilities and related mild disabilities: Characteristics, teaching strategies, and new directions* (13th ed.). Wadsworth Cengage Learning.

Liou, P.-Y. (2014). Investigation of the big-fish-little-pond effect on students' self-concept of learning mathematics and science in Taiwan: Results from TIMSS 2011. *The Asia-Pacific Education Researcher, 23*(3), 769-778. http://dx.doi.org/10.1007/s40299-013-0152-3

Logan, S., Medford, E., & Hughes, N. (2011). The importance of intrinsic motivation for high and low ability readers' reading comprehension performance. *Learning and Individual Differences, 21*(1), 124-128. https://doi.org/10.1016/j.lindif.2010.09.011

Lohbeck, A., & Möller, J. (2017). Social and dimensional comparison effects on math and reading self-concepts of elementary school children. *Learning and Individual Differences, 54*, 73-81. http://dx.doi.org/10.1016/j.lindif.2017.01.013

Lüdtke, O., Köller, O., Marsh, H. W., & Trautwein, U. (2005). Teacher frame of reference and the big-fish-little-pond effect. *Contemporary Educational Psychology, 30*(3), 263-285. https://doi.org/10.1016/j.cedpsych.2004.10.002

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*(1), 129-149. http://dx.doi.org/10.3102/00028312023001129

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology, 79*(3), 280-295. http://dx.doi.org/10.1037/0022-0663.79.3.280

Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2*(2), 77-172. http://dx.doi.org/10.1007/BF01322177

Marsh, H. W. (1990b). The structure of academic self-concept: The marsh/shavelson model. *Journal of Educational Psychology, 82*(4), 623-636. http://dx.doi.org/10.1037/0022-0663.82.4.623

Marsh, H. W. (1990c). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology, 82*(1), 107-116. https://doi.org/10.1037//0022-0663.82.1.107

Marsh, H. W. (1994). Using the National Longitudinal Study of 1988 to evaluate theoretical models of self-concept: The self-description questionnaire. *Journal of Educational Psychology, 86*(3), 439-456. http://dx.doi.org/10.1037/0022-0663.86.3.439

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology, 80*(3), 366-380. http://dx.doi.org/10.1037/0022-0663.80.3.366

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*(2), 133-163. http://dx.doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., Craven, R., & Debus, R. (1999). Separation of competency and affect components of multiple dimensions of academic self-concept: A developmental perspective. *Merrill-Palmer Quarterly, 45*(4), 567-601.

Marsh, H. W., & Hau, K.-T. (2003). Big-fish-little-pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist, 58*(5), 364-376. http://dx.doi.org/10.1037/0003-066X.58.5.364

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96*(1), 56-67. http://dx.doi.org/10.1037/0022-0663.96.1.56

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: Counterbalancing contrast and reflected-glory effects in Hong Kong schools. *Journal of Personality and Social Psychology, 78*(2), 337-349. http://dx.doi.org/10.1037/0022-3514.78.2.337

Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behavioral Research, 44*(6), 764-802. http://dx.doi.org/10.1080/00273170903333665

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*(1), 213-231. http://dx.doi.org/10.1037/0022-3514.47.1.213

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology, 54*(2), 263-280. http://dx.doi.org/10.1037/dev0000393

Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Craven, R. G. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review, 20*(3), 319-350. https://doi.org/10.1007/s10648-008-9075-6

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist, 20*(3), 107-123. https://doi.org/10.1207/s15326985ep2003_1

Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89*(1), 41-54. https://doi.org/10.1037/0022-0663.89.1.41

Martin, M. O., & Mullis, I.V.S. (Eds.). (2013).*TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade - implications for early learning.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/timsspirls2011/downloads/TP11_Relationship_Report.pdf

Marusic, M. (2014). Is the intrinsic motivation of career choice important for the later professional development of teachers? *Psihologija*, *47*(4), 449-464. https://doi.org/10.2298/PSI1404449M

McDaniel, M. A., Waddill, P. J., Finstad, K., & Bourg, T. (2000). The effects of text-based interest on attention and recall. *Journal of Educational Psychology*, *92*(3), 492-502. https://doi.org/10.1037/0022-0663.92.3.492

Meinck, S. (2019). *Multilevel modelling with large-scale assessment data: Uncovering sampling-related challenges* [Conference proposal]. Proposal for the 8th International Research Conference of IEA, Copenhagen

Mertens, D. M. (2015). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (4th ed.). SAGE Publications.

Mitchell, D. R. (2014). *What really works in special and inclusive education: Using evidence-based teaching strategies* (2nd ed.). Routledge.

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research, 79*(3), 1129-1167. http://dx.doi.org/10.3102/0034654309337522

Mullis, I. V. S. (2007). *PIRLS 2006 international report: IEA's progress in international reading literacy study in primary schools in 40 countries*. TIMSS & PIRLS. https://timss.bc.edu/pirls2006/intl_rpt.html

Mullis, I.V.S., Drucker, K. T., Preuschoff, C., Arora, A., & Stanco, G.M. (2013). *Assessment framework and instrument development* (M. Martin & I. V. A. Eds.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/methods/pdf/TP_Instrument_Devel.pdf

Mullis, I.V.S. & Martin, M.O. (Eds.) (2013). *TIMSS 2015 assessment frameworks*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/timss2015/frameworks.html

Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/pirls2016/framework.html

Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 assessment frameworks.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/timss2019/frameworks/

Mullis, I. V. S., & Martin, O. M. (Eds.). (2019). *PIRLS 2021 assessment frameworks.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://www.iea.nl/publications/assessment-framework/pirls-2021-assessment-frameworks

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/timss2011/international-results-mathematics.html

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2013). *Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales* (M. Martin & I. V. A. Eds.). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/methods/pdf/TP11_Context_Q_Scales.pdf

Mullis, I.V.S., Martin, M.O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 international results in reading.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/pirls2011/international-results-pirls.html

Mullis, I.V.S., Martin, M.O., Foy, P., Hooper, M. (2017). *PIRLS 2016 international results in reading*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational

Achievement. http://timssandpirls.bc.edu/pirls2016/international-results/pirls/student-achievement/pirls-achievement-results/

Mullis, I.V.S., Martin, M.O., Foy, P., Kelly, D. L., Fishbein, B. (2020). *TIMSS 2019 international results in mathematics and science.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement. https://timssandpirls.bc.edu/timss2019/international-results/

Mullis, I. V. S., Martin, O. M., Kennedy, A. M., Trong, K. L., & Sainsbury, M. (2009). *PIRLS 2011 assessment framework.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. https://timssandpirls.bc.edu/pirls2011/framework.html

Mullis, I. V. S., Martin, O. M., Ruddock, G. J., O'Sullivan, C. Y., & Preuschof, C. (2009). *TIMSS 2011 assessment frameworks.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. https://timss.bc.edu/timss2011/frameworks.html

Muthén, L. K., & Muthén, B. O. (2017). *MPLUS statistical analysis with latent variables: User's guide* (8th ed.). Muthén & Muthén.

NCCA. (2018*). Primary mathematics curriculum: Draft specification: Junior infants to second class.* https://ncca.ie/media/3148/primary_mathsspec_en.pdf

NCCA. (2019). *Primary language curriculum.* Department of Education and Skills.

Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education, 7*(2), 133-144. https://doi.org/10.1177/1477878509104318

NEPS. (2010). *Behavioural, emotional and social difficulties: A continuum of support: Guidelines for teachers*. https://www.education.ie/en/Schools-Colleges/Services/National-Educational-Psychological-Service-NEPS-/neps_besd_continuum_teacher_guide.pdf

NEPS. (2013). *The National Educational Psychological Service (NEPS): Revised model of service statement*. NEPS

NEPS. (2016). *A balanced approach to literacy development in the early years*. https://www.education.ie/en/Schools-Colleges/Services/National-Educational-Psychological-Service-NEPS-/Balanced-Approach-to-Literacy-Development-in-the-Early-Years-.pdf

NEPS. (2019). *Effective interventions for struggling readers* (2 ed.). https://www.education.ie/en/Education-Staff/Information/NEPS-Literacy-Resource/neps_literacy_good_practice_guide.pdf

NEPS. (2020). *Maths support: A good practice guide for teachers 2020.* https://www.education.ie/en/Schools-Colleges/Services/National-Educational-Psychological-Service-NEPS-/maths-support-guide-2020.pdf

Ommundsen, Y., Haugen, R., & Lund, T. (2005). Academic self-concept, implicit theories of ability, and self-regulation strategies. *Scandinavian Journal of Educational Research, 49*(5), 461-474. https://doi.org/10.1080/00313830500267838

Organisation for Economic Co-operation and Development. (2004). *Learning for tomorrow's world: First results from PISA 2003.* OECD Publishing. https://www.oecd-ilibrary.org/education/learning-for-tomorrow-s-world_9789264006416-en

Organisation for Economic Co-operation and Development. (2007). *PISA 2006: Science competencies for tomorrow's world.* OECD Publishing. http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/pisa2006results.htm#how_to_obtain

Organisation for Economic Co-operation and Development. (2014). *PISA 2012 results: Ready to learn: Students' engagement, drive and self-beliefs*. OECD Publishing. http://www.deslibris.ca/ID/243495

Parker, P. D., Marsh, H. W., Luedtke, O., & Trautwein, U. (2013). Differential school contextual effects for math and English: Integrating the big-fish-little-pond effect and the internal/external frame of reference. *Learning and Instruction, 23*, 78-89. http://dx.doi.org/10.1016/j.learninstruc.2012.07.001

Perkins, R., & Clerkin, A. (2020). *TIMSS 2019: Ireland's results in mathematics and science.* Educational Research Centre. https://www.erc.ie/wp-content/uploads/2021/01/03-ERC-TIMSS-2019-Report_A4_Online.pdf

Pinxten, M., Wouters, S., Preckel, F., Niepel, C., De Fraine, B., & Verschueren, K. (2015). The formation of academic self-concept in elementary education: A unifying model for external and internal comparisons. *Contemporary Educational Psychology, 41*, 124-132. http://dx.doi.org/10.1016/j.cedpsych.2014.12.003

Psychological Society of Ireland. (2019). *Code of Professional Ethics* (4th ed.). https://www.psychologicalsociety.ie/footer/Code-of-Ethics

Reeve, J. (2012). A self-determination theory perspective on student engagement. in S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 149-172). Springer.

Renninger, K. A., Ewen, L., & Lasher, A. K. (2002). Individual interest as context in expository text and mathematical word problems. *Learning and Instruction, 12*(4), 467–491. https://doi.org/10.1016/S0959-4752(01)00012-3

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 173–195). Academic.

Robson, C., & McCartan, K. (2016). *Real world research: A resource for users of social research methods in applied settings* (4th ed.). John Wiley & Sons.

Roy, A., Guay, F., & Valois, P. (2015). The big-fish–little-pond effect on academic self-concept: The moderating role of differentiated instruction and individual achievement. *Learning and Individual Differences, 42*, 110-116. http://dx.doi.org/10.1016/j.lindif.2015.07.009

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* John Wiley & Sons.

Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher, 39*(2), 142-151. https://doi.org/10.3102/0013189X10363170

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*, 54-67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist, 55*(1), 68-78. http://dx.doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. The Guilford Press.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning & Instruction, 34*, 32-41. http://dx.doi.org/10.1016/j.learninstruc.2014.08.001

Seaton, M., Marsh, H. W., Parker, P. D., Craven, R. G., & Yeung, A. S. (2015). The reciprocal effects model revisited: Extending its reach to gifted students attending academically selective schools. *The Gifted Child Quarterly, 59*(3), 143-156. https://doi.org/10.1177/0016986215583870

Shavelson R., Hubner J., & Stanton G. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research, 46*, 407-441. http://dx.doi.org/10.3102/00346543046003407

Skwarchuk, S. L. (2009). How do parents support preschoolers' numeracy learning experiences at home? *Early Childhood Education Journal, 37*(3), 189-197. http://dx.doi.org/10.1007/s10643-009-0340-1

Stallard, P. (2005). *A clinician's guide to think good-feel good: Using CBT with children and young people*. John Wiley & Sons.

Stallard, P. (2019). *Think good, feel good: A cognitive behavioural therapy workbook for children and young people*. John Wiley & Sons.

Sung, Y.-T., Huang, L.-Y., Tseng, F.-L., & Chang, K.-E. (2014). The aspects and ability groups in which little fish perform worse than big fish: Examining the big-fish-little-

pond effect in the context of school tracking. *Contemporary Educational Psychology, 39*(3), 220-232. http://dx.doi.org/10.1016/j.cedpsych.2014.05.002

Taylor, G., Jungert, T., Mageau, G. A., Schattke, K., Dedic, H., Rosenfield, S., & Koestner, R. (2014). A self-determination theory approach to predicting school achievement over time: The unique role of intrinsic motivation. *Contemporary Educational Psychology, 39*, 4, 342-358. https://doi.org/10.1016/j.cedpsych.2014.08.002

Thijs, J., Verkuyten, M., & Helmond, P. (2010). A further examination of the big-fish-little-pond effect: Perceived position in class, class size, and gender comparisons. *Sociology of Education, 83*(4), 333-345. http://dx.doi.org/10.1177/0038040710383521

Thompson, B., Diamond, K. E., McWilliam, R., Snyder, P., & Snyder, S. W. (2005). Evaluating the quality of evidence from correlational research for evidence-based practice. *Exceptional Children, 71*(2), 181-194. http://dx.doi.org/10.1177/001440290507100204

TIMSS and PIRLS International Study Centre. (2013a). *TIMSS 2011: The TIMSS 2011 students like learning mathematics scale, fourth grade* (M. Martin & I. V. A. Eds.). Methods and procedures in TIMSS and PIRLS 2011. https://timssandpirls.bc.edu/methods/pdf/T11_G4_M_Scales_SLM.pdf

TIMSS and PIRLS International Study Centre. (2013b). *PIRLS 2011: The PIRLS 2011 students like reading scale* (M. Martin & I. V. A. Eds.). Methods and procedures in TIMSS and PIRLS 2011. https://timssandpirls.bc.edu/methods/pdf/P11_R_Scales_SLR.pdf

TIMSS and PIRLS International Study Centre. (2013c). *TIMSS 2011: The TIMSS 2011 students confident in mathematics scale, fourth grade* (M. Martin & I. V. A. Eds.). Methods and procedures in TIMSS and PIRLS 2011. https://timssandpirls.bc.edu/methods/pdf/T11_G4_M_Scales_SCM.pdf

TIMSS and PIRLS International Study Centre. (2013d). *PIRLS 2011: PIRLS 2011 students confident in reading scale* (M. Martin & I. V. A. Eds.). Methods and procedures in TIMSS and PIRLS 2011. https://timssandpirls.bc.edu/methods/pdf/P11_R_Scales_SCR.pdf

TIMSS and PIRLS International Study Centre. (2019). *Relationships among reading, mathematics, and science achievement at the fourth grade*. TIMSS and PIRLS International Study Centre. https://timssandpirls.bc.edu/timsspirls2011/international-database.html

United Nations. (2010). *The United Nations convention on the rights of the child*. Children's Rights Alliance. https://www.childrensrights.ie/sites/default/files/submissions_reports/files/UNCRCEnglish_0.pdf

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist, 39*(2), 111-133. https://doi.org/10.1207/s15326985ep3902_3

Wall, E., & Burke, K. (2003). *Mary Immaculate College reading attainment test: Level 2*. CJFallon.

Wall, E., & Burke, K. (2007). *Standardised Irish graded mathematics test: Level 2*. CJFallon.

Wang, J. H., & Guthrie, J. T. (2004). Modelling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between United States and Chinese students. *Reading Research Quarterly, 39*(2), 162. https://doi.org/10.1598/RRQ.39.2.2

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Yeung, A. S. (2011). Student self-concept and effort: Gender and grade differences. *Educational Psychology, 31*(6), 749–772. https://doi.org/10.1080/01443410.2011.608487.

Yeung, A. S., Kuppan, L., Foong, S. K., Wong, D., Kadir, M. S., Lee, P., & Yau, C. M. (2010). Domain specificity of self-concept and parent expectation influences on short-term and long-term learning of physics. *New Horizons in Education, 58*(2), 54–72.

Yeung, A. S., Kuppan, L., Kadir, M. S., & Foong, S. K. (2010). Boys' and girls' self-beliefs, engagement, and aspirations in physics. *International Journal of Learning, 17*(10), 397–417.

Appendices

Appendix A

Confirmation of Ethical Approval for the Original Study

Confirmation of Ethical Approval for the Original Study (Prior to the 2020 COVID-19 Related School Closures)

Mary Immaculate College Research Ethics Committee

MIREC-4: MIREC Chair Decision Form

APPLICATION NO. A20-003

1. PROJECT TITLE

The Relevance of the Integrated Internal/External Frame of Reference Model to Second Class Pupils' Mathematics and Reading Self-Concepts

2. APPLICANT

Name:	Aoife Cassidy
Department / Centre / Other:	Educational Psychology, Inclusive & Special Education
Position:	Postgraduate Researcher

3. DECISION OF MIREC CHAIR

☐	Ethical clearance through MIREC is required.
☐	Ethical clearance through MIREC is not required and therefore the researcher need take no further action in this regard.
✓	Ethical clearance is required and granted. Referral to MIREC is not necessary.
☐	Ethical clearance is required but the full MIREC process is not. Ethical clearance is therefore granted if required for external funding applications and the researcher need take no further action in this regard.
☐	Insufficient information provided by applicant / Amendments required.

4. REASON(S) FOR DECISION

A20-003 - Aoife Cassidy - ***The Relevance of the Integrated Internal/External Frame of Reference Model to Second Class Pupils' Mathematics and Reading Self-Concepts***

I have reviewed this application and I believe it satisfies MIREC requirements. Therefore, it is approved in full.

5. DECLARATION (MIREC CHAIR)

Name (Print):	Dr Áine Lawlor
Signature:	Áine Lawlor
Date:	21st January 2020

MIREC-4 Rev 3 Page 1 of 1

Appendix B

List of Excluded Studies from the Literature Review

Table B 1

Articles Excluded During Title and Abstract Screening

Article	Exclusion criteria
Cambria, J., Brandt, H., Nagengast, B., & Trautwein, U. (2017). Frame of reference effects on values in mathematics: evidence from German secondary school students. *Zdm-Mathematics Education, 49*(3), 435-447. http://dx.doi.org/10.1007/s11858-017-0841	3
Fang, J., Huang, X., Zhang, M., Huang, F., Li, Z., & Yuan, Q. (2018). The big-fish-little-pond effect on academic self-concept: A meta-analysis. *Frontiers in Psychology, 9*, 1569-1569. http://dx.doi.org/10.3389/fpsyg.2018.01569	3
Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2012). Personality traits moderate the big-fish–little-pond effect of academic self-concept. *Learning and Individual Differences, 22*(6), 736-746. http://dx.doi.org/10.1016/j.lindif.2012.07.020	3
Kong, C.-K. (2000). Chinese students' self-concept: Structure, frame of reference, and relation with academic achievement. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 61*(3-A), 880.	1
Kuzmina Yu, V. (2016). Big-frog-in-a-small-pond effect: Is it always good for a child to study in a strong class? *Psikhologiya Psychology, Journal of the Higher School of Economics, 13*(4), 712-740. http://dx.doi.org/10.17323/1813-8918-2016-4-712-740	3
Kwiterovich, P. O., III. (2012). Big-fish-little-pond effect: The impact on academic self-concept. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 72*(7-A), 2333.	1
Ludtke, O., & Koller, O. (2002). Individual reference norm and social comparisons in mathematics classes: The impact of different frames of reference on the domain-specific self-concept of ability. *Zeitschrift Fur Entwicklungspsychologie Und Padagogische Psychologie, 34*(3), 156-166. http://dx.doi.org/10.1026//0049-8637.34.3.156	3
Lüdtke, O., Köller, O., Artelt, C., Stanat, P., & Baumert, J. (2002). Eine Überprüfung von Modellen zur Genese akademischer Selbstkonzepte: Ergebnisse aus der PISA-Studie = Testing models on the genesis of academic self-concepts: Findings of the PISA Study. *Zeitschrift für*	2

Pädagogische Psychologie / German Journal of Educational Psychology, 16(3-4), 151-164. http://dx.doi.org/10.1024//1010-0652.16.34.151

Noh, H., Kim, D., Park, S., & Sohn, W. (2018). Impact of academic achievement on students mathematical and English subject-attitude: An integrated approach for I/E and BFLPE Models. *Korean Soc Edu Evaluation Korean Society for Educational Evaluation, 31*(1), 101-124 2

Xia, Z., Yang, F., Praschan, K., & Xu, Q. (2019). The formation and influence mechanism of mathematics self-concept of left-behind children in mainland china. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues. Advance online publication*. https://doi.org/10.1007/s12144-019-00495-4 3

Zhang, W. (1996). Influences of internal and external frames of reference on the formation of math and verbal self-concepts for gifted and non-gifted tenth grade students. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *56*(10-A), 3894. 3

Table B 2

Articles Excluded During Full-Text Screening

Article	Exclusion criteria
Kadir, M. S., Yeung, A. S., & Diallo, T. M. O. (2017). Simultaneous testing of four decades of academic self-concept models. *Contemporary Educational Psychology, 51*, 429-446. http://dx.doi.org/doi:10.1016/j.cedpsych.2017.09.008	3
Keyserlingk, L., Becker, M., Jansen, M., & Maaz, K. (2020). Leaving the pond-choosing an ocean: Effects of student composition on STEM major choices at university. *Journal of Educational Psychology, 112*(4), 751-764. https://doi.org/10.1037/edu0000378	3
Marsh, H. W., Parker, P. D., Guo, J., Pekrun, R., & Basarkod, G. (2020). Psychological comparison processes and self–concept in relation to five distinct frame–of–reference effects: Pan–human cross–cultural generalizability over 68 countries. *European Journal of Personality, 34*(2), 180-202. https://doi.org/10.1002/per.2232	3
Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology, 54*(2), 263-280. http://dx.doi.org/10.1037/dev0000393	3
Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331-353. http://dx.doi.org/10.1037/edu0000281	3
Rosman, T., Mayer, A., Leichner, N., & Krampen, G. (2020). Putting big fish into a bigger pond: Self-concept changes in psychology undergraduate entrants. *Journal of further and Higher Education, 44*(1), 14-28. https://doi.org/10.1080/0309877X.2018.1493095	3
Song, K.-Y., & Ahn, J.-D. (2014). A comparative analysis of body composition, physical fitness, and physical self-concept between gifted students in math and science and non-gifted students. *The Journal of the Korea Contents Association The Journal of the Korea Contents Association, 14*(11), 450-466. https://doi.org/10.5392/JKCA.2014.14.11.450	3

Appendix C

Mapping the Field: Summary of the Nine Studies in the Literature Review

Study	Sample	Description of study	Measures	Main findings
Chiu (2012)	Countries: 27 countries that participated in the Trends in International Mathematics and Science Study (TIMSS) in 2003. Ireland not included School grade/age: Pupils who received eight years of formal schooling; ages varied per country; average age of pupils in each country varied from 13.7 – 15.5 years old 139,174 pupils in 4231 schools	Paper provided a rationale and conceptual framework for the combined model Analysis sought to investigate whether the combined model would be supported by TIMSS data from 27 countries	Data taken from the 2003 TIMSS -Mathematics Achievement -Science Achievement -Mathematics self-concept -Science self-concept	Support for the combined model for mathematics self-concept. (Fully supported for 16 countries; partially supported 11 countries) Partial support for the combined model for science self-concept. Mathematics achievement did not have a significant negative impact on pupils' science self-concept. (Fully supported for nine countries; partially supported for 15 countries; not supported for four counties) Chi-squared test of association indicated that the combined model was a better fit to the data than the basic I/E model and BFLPE models
Guo, Marsh, Parker and Dicke (2018)	Counties: 15 OECD countries that participated in the 2011 TIMSS and the Progress in International Reading Literacy Study (PIRLS) School grade/age: Grade 4 pupils	Drawing on the BFLPE and the I/E model, the study sought to simultaneously investigate social and dimensional comparison effects across three academic domains within a unified	Data taken from the 2011 TIMSS and PRILS -Mathematics achievement -Mathematics self-concept -English	Moderate correlations between ASC and achievement in reading (r = .399) and mathematics (r = .314). Smaller correlation between science achievement and science self-

	67385 participants in 3808 classes and 2564 schools	model	achievement -English self-concept -Science achievement -Science self-concept	concept (r = .202) Cross culturally, slightly negative or non-significant effect of math/reading achievement to reading/math self-concept (non-matching domains) Small negative impact of individual mathematics achievement on science self-concept. Science achievement did not significantly impact mathematics self-concept Reading achievement had a positive impact on science self-concept. Science achievement did not have a significant impact on reading self-concept Strong negative impact of class-average achievement (BFLPE) on ASC in matching domains; this effect was strongest for mathematics, followed by reading, and a smaller negative effect was found for science; effect for science found to be significant in only nine of the 15 counties

				Degree to which the predictions of the combined model were supported by the data varied per country and also by the relative distinctiveness between the academic domains tested
Kavanagh (2019)	Country: Ireland School grade/age: Pupils in grade six (sixth-class; mean age 12.5 years) 4166 pupils in 220 classes	Study investigated whether the combined model would be supported by data collected from grade 6 (sixth-class) pupils	Data taken from the 2014 National Assessment of Mathematics and English Reading -Mathematics achievement -Mathematics self-concept -Reading achievement -Reading self-concept	The data supported the combined model There was a larger effect size for the BFLPE for mathematics (*ES* = -.42) than for reading (*ES* = -.19) No significant effect of class-average achievement on pupils' self-concept in non-matching academic domains
Lohbeck and Möller (2017)	Country: Germany (Lower Saxony) School grade/age: Pupils in grade 2 (aged seven to nine years old) 291 pupils in 20 classes and 10 schools	Study sought to investigate whether the combined model would be supported by data from grade 2 pupils in Germany	-Mathematics achievement (addition and subtraction with and without ten transition; taken from a standardized German mathematics assessment) -Reading Achievement (standardized German reading comprehension test) -Mathematics self-concept -Reading self-	Assumptions of the combined model not fully supported Significant correlation between mathematics and reading achievement (.41) Significant correlation between mathematics achievement and mathematics self-concept (.40) Significant correlation between reading achievement and reading self-

			concept	concept (.33) Significant BELPE for class-average mathematics achievement on pupils' mathematics self-concept (*ES* = -.23) No effect of class-average reading achievement on reading self-concept Individual achievement in one domain did not have a significant impact on individuals' self-concept in a non-matching domain (slightly negative relationship – not significant)
Marsh (1990c)	Countries: United States School grade/age: Pupils were high-school sophomores 14825 pupils in 1015 high schools	The study sought to investigate the I/E model and the BFLPE within a unified framework	Data taken from the High School and Beyond study (1980 cohort) -Mathematics achievement -English achievement -Mathematics self-concept -English self-concept	Results supported the predictions of the combined model Individual achievement was positively related to ASC in matching domains Individual achievement was negatively related to ASC in non-matching domains, however, these correlations were smaller than the aforementioned positive correlations between achievement and self-concept in matching domains

				School-average achievement had a negative impact on ASC in matching domains There was a small positive correlation between school-average English achievement and mathematics self-concept (.13)
Marsh (1994)	Country: Pupils from the United States School grade/age: Pupils in 10th grade; some achievement measures taken when children in eighth-grade 17544 pupils from the United States	The paper sought to (i) evaluate self-concept measures employed in the NELS:88 study, (ii) compare the responses of pupils in the United States to that of Australian pupils on the Self-Description Questionaire (SDQ) and (iii) investigate whether the I/E model and the BFLPE were supported by the data in relation to pupils' ASCs	Data taken from the National Educational Longitudinal Survey of 1988 -Mathematics self-concept -English self-concept -Mathematics achievement -Reading achievement -Mathematics grades -English grades	The data supported both the I/E model and the BEFPE when tested simultaneously within a unified model School-average reading achievement had a positive impact on mathematics self-concept School-average mathematics achievement did not have a significant impact on English self-concept
Parker, Marsh, Ludtke and Trautwein (2013)	Country: Germany School grade/age: Pupils in final year of secondary school; average age 19.76 years 5016 pupils in 157 schools	Study sought to investigate whether the combined model would be supported by German pupils in traditional and themed (magnet) university track schools	Data from the Transformation of the Secondary School System and Academic Careers project -Mathematics achievement -English as a foreign language achievement -Mathematics self-concept	Mathematics and English achievement were significantly more differentiated in themed schools, compared to traditional track schools Combined model was supported by data from both themed and traditional university track

			-English self-concept	schools BFLPE effect sizes larger in themed schools compared to traditional track schools; BFLPE not significant for English in traditional track schools Significant effect of school-average achievement on pupils' self-concept in non-matching academic domains in themed schools, but not in traditional university track schools
Pinxten et al. (2015)	Country: Belgium (region: Flanders) School grade: Pupils in grade four of elementary school; 4436 pupils in 241 classes	Study sought to investigate whether the combined model would be supported by data from grade 4 pupils in Flanders	Data from SiBO project (Dutch acronym for School Trajectories in Primary Education) -Mathematics achievement -Dutch achievement (spelling and reading comprehension test) -Mathematics self-concept -Dutch self-concept	Predictions of the combined model supported by the data Moderate effect size for the BFLPE in this model Small significant effect of class-average achievement on pupils' self-concept in contrasting academic domains
Schurtz, Pfost, Nagengast and Artelt (2014)	Country: Germany School grade/age: Pupils in fifth and sixth grade; average age of pupils in grade five was 11.5 years	The study sought to investigate the impact of social and dimensional comparisons on pupils' subject-interests in mathematics and English as an	Data taken from the German BiKS project -Mathematics competence -English competence	In line with the combined model, pupils' competencies had an impact on their subject-interests in mathematics and English; this was mediated by pupils'

	1390 pupils in 108 classes	additional language; the combined model was employed as a framework to investigate social and dimensional comparisons	-Grades in mathematics and English (as a foreign language) -ASCs in mathematics and English -Subject-interests in mathematics and English.	subject-specific grades and self-concepts. Class-average achievement in mathematics had a significant negative impact on pupils' mathematics self-concept ($\beta = -.433$; $p < .05$) Pupils' grades had a significant positive impact on pupils' self-concepts in matching domains (Mathematics: $\beta = .407$; English: $\beta = .497$). Pupils' grades had a smaller, but significant, negative impact on their ASCs in non-matching domains (English grades on mathematics self-concept: $\beta = -.156$; Mathematics grades on English self-concept: $\beta = -.151$)

Appendix D

Coding Protocol for Weight of Evidence A

The following protocol was compiled considering (i) the review question, (ii) relevant studies, and (iii) relevant literature documenting quality standards for correlational research (Howitt & Cramer, 2011; Mertens, 2015). Items for the 'practical and clinical significance' section and the 'data analysis' section were taken from Thompson et al.'s (2005) 'Suggested Quality Indicators for Correlational Research'. Studies received a rating of three for WOE A if they met seven or eight of the protocol criteria, a rating of two if they met six of the criteria, and a rating of one if they met five or less of the criteria.

Measures

Evidence of validity for achievement scores (e.g. test constructed specifically to test national/international academic achievement standards in a particular domain, in line with curriculums/ standards that would typically be expected from pupils in a particular school grade; appropriate range of items employed to assess academic achievement in a given domain)

☐ Yes

☐ No

☐ N/A

Evidence of validity for self-concept scores (e.g. reference to prior literature/ study employed an established measure which is typically employed when assessing ASC among a particular cohort/sample, and this measure is supported by literature)

☐ Yes

☐ No

☐ N/A

Score reliability coefficients are reported for all measured variables, based on data from participants who completed the study.

☐ Yes

☐ No

☐ N/A

Reliability coefficients for all measures are .7 or above.

☐ Yes

☐ No

☐ N/A

Practical and Clinical Significance

Effect size statistics are reported for all primary study outcomes, and effect size statistics are clearly identified

☐ Yes

☐ No

☐ N/A

Authors interpret study effect sizes by directly and explicitly comparing study effects with those reported in prior studies

☐ Yes

☐ No

☐ N/A

Authors explicitly consider the study design and effect size statistic limitations as part of effect interpretation

☐ Yes

☐ No

☐ N/A

Data Analysis

Persuasive evidence is explicitly presented that the assumptions of statistical methods are sufficiently well-met for results to be deemed credible

☐ Yes

☐ No

☐ N/A

Appendix E

Coding Protocol for Weight of Evidence B

Criteria for WOE B were established considering the review question and relevant literature documenting the appropriateness of research methodologies and designs for answering particular types of research questions (Howitt & Cramer, 2011; Mertens, 2015). Studies that adhered to all four items received a rating of three. Studies that adhered to three of the items received a rating of two. Studies that adhered to two or less of the items received a rating of one.

Criteria

-As all studies in the current review were cross-sectional, data for all measures were collected within the same academic school year; accordingly, measures were not impacted by factors such as a change in a pupils' attainment or academic self-concept from one school year/class to another

☐ Yes

☐ No

☐ N/A

-Achievement and self-concept domains that were investigated in the study were distinctly different from each other; (if two similar domains were investigated, dimensional comparison processes between pupils' performances in different academic domains may not be evident)

☐ Yes

☐ No

☐ N/A

-Multilevel structural equation modelling was employed to analyse the data. Consistent with Marsh et al. (2009), motivational constructs such as ASC are construed as latent constructs within this model/analysis. The multilevel component is warranted to account for the hierarchical nature of the data.

☐ Yes

☐ No

☐ N/A

-The study design or analysis accounted for the sampling error associated with the class/school-average achievement measure. Specifically, this may be achieved through aiming for a 100% sample rate among sampled schools/classes, or through the class/school-average achievement measure being construed as a latent variable within the analysis.

☐ Yes

☐ No

☐ N/A

Appendix F

Coding Protocol for Weight of Evidence C

WoE C was evaluated using the criteria in the checklist documented below. Studies that adhered to all four criteria received a rating of three. Studies which adhered to three of the criteria received a rating of two. Studies which adhered to two or less of the criteria received a rating of one.

Criteria	Rationale
-Sufficient details noted regarding sample size and characteristics (e.g. age, nationality). Participant sample size was representative of a particular region, country or specified area (including international studies which examined countries separately and/or aggregated scores from multiple countries). Generalisability of findings was discussed in line with sample size and characteristics (e.g. participants from a particular area/ attending a particular school).	-To consider whether the findings may have been impacted by particular characteristics of the sample participants (e.g. cultural factors) -To consider the generalisability of study findings and synthesise results from all studies, in line with the research question -Larger samples that are statistically representative of a particular region provide more robust evidence for the generalisation of the findings to that region and for the interpretation of the findings in line with other studies and the research question
-Evidence that participants were randomly sampled	-To promote generalisability of the findings
-Principle focus of study was to investigate the combined model of academic self-concept formation, as described in the introduction. Evidence obtained from the study was explicitly discussed and critiqued	-Relevance of evidence to review question -In-depth analysis/discussion of findings from the author will provide further insights/interpretations that will likely be of use to answering the current review

in relation to the current review question.	question.
-Both the traditional I/E model and the BFLPE are investigated using separate models, and subsequently within the same model when investigating the combined model.	-Previously, the I/E model and the BFLPE were not supported by data from some countries, and effect sizes varied based on data from different countries. Thus, testing the models separately before testing them within the combined model allowed for a more critical analysis of the findings in line with the review question.

Appendix G

MPLUS Syntax for Each Confirmatory Factor Analysis

Confirmatory Factor Analysis: MPLUS Syntax Model One

```
TITLE:      CFA with Four factor indicators RSI MSI RSC MSC

DATA:       FILE IS File1.dat;

VARIABLE:   NAMES ARE

            IDSCHOOL IDCLASS IDSTUD ITSEX

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G

            ASDAGE TOTWGT HOUWGT SENWGT

            WGTADJ1 WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3

            JKREP JKZONE SACHM SACHR;


            USEVARIABLES =

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G;


            MISSING ARE

            ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)

            JKREP(9) JKZONE(99);

ANALYSIS:   ESTIMATOR = MLR;
```

```
MODEL:    RSI BY ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr;
          MSI BY ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er;
          RSC BY ASBR08Ar ASBR08Br ASBR08C ASBR08G;
          MSC BY ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G;
OUTPUT:   MOD standardized sampstat;
```

Confirmatory Factor Analysis: MPLUS Syntax Model Two

```
DATA:       FILE IS File1.dat;

VARIABLE:   NAMES ARE

            IDSCHOOL IDCLASS IDSTUD ITSEX

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G

            ASDAGE TOTWGT HOUWGT SENWGT

            WGTADJ1 WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3

            JKREP JKZONE SACHM SACHR;


            USEVARIABLES =

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G;


            MISSING ARE

            ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)

            JKREP(9) JKZONE(99);

ANALYSIS:   ESTIMATOR = MLR;

MODEL:      RSI BY ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr;

            MSI BY ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er;

            RSC BY ASBR08Ar ASBR08Br ASBR08C ASBR08G;
```

MSC BY ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G;

! Correlated error terms positively-worded items (i.e. reverse-coded)

ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er ASBM03Ar ASBM03Dr ASBR08Ar ASBR08Br with ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er ASBM03Ar ASBM03Dr ASBR08Ar ASBR08Br;

! Correlated error terms negatively-worded items (i.e. the rest)

ASBR07A ASBR07D ASBM01B ASBM01C ASBM03B ASBM03C ASBM03G ASBR08C ASBR08G with

ASBR07A ASBR07D ASBM01B ASBM01C ASBM03B ASBM03C ASBM03G ASBR08C ASBR08G;

OUTPUT: MOD standardized;

Confirmatory Factor Analysis: MPLUS Syntax Model Three

```
TITLE:      CFA with Four factor indicators RSI MSI RSC MSC

DATA:       FILE IS File1.dat;

VARIABLE:   NAMES ARE

            IDSCHOOL IDCLASS IDSTUD ITSEX

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G

            ASDAGE TOTWGT HOUWGT SENWGT WGTADJ1

            WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3

            JKREP JKZONE SACHM SACHR;


            USEVARIABLES =

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G;


            MISSING ARE

            ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)

            JKREP(9) JKZONE(99);


            CLUSTER = IDCLASS;
```

```
                WEIGHT IS HOUWGT;


ANALYSIS:       ESTIMATOR = MLR; TYPE = COMPLEX;


MODEL:          RSI BY ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr;

                MSI BY ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er;

                RSC BY ASBR08Ar ASBR08Br ASBR08C ASBR08G;

                MSC BY ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G;
! Correlated error terms positively-worded items (i.e. reverse-coded)

ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er ASBM03Ar ASBM03Dr
ASBR08Ar ASBR08Br with ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er
ASBM03Ar ASBM03Dr ASBR08Ar ASBR08Br;

! Correlated error terms negatively-worded items (i.e. the rest)

 ASBR07A ASBR07D ASBM01B ASBM01C ASBM03B ASBM03C ASBM03G ASBR08C ASBR08G with
 ASBR07A ASBR07D ASBM01B ASBM01C ASBM03B ASBM03C ASBM03G ASBR08C ASBR08G;


OUTPUT:         TECH4 standardized;
```

Appendix H

MPLUS Syntax for Each Theoretical Model

MPLUS Syntax for the Basic I/E Model

```
TITLE:      Basic I/E Model

DATA:       FILE IS Filelist.dat;

            TYPE = IMPUTATION;

VARIABLE:  NAMES ARE

            IDSCHOOL IDCLASS IDSTUD ITSEX

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBR08Ar ASBR08Br ASBR08C ASBR08G

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er

            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G

            ASDAGE TOTWGT HOUWGT SENWGT WGTADJ1 WGTADJ2
            WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3

            JKREP JKZONE SACHM SACHR;


            USEVARIABLES =

            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr

            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er SACHM
            SACHR;


            MISSING ARE

            ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)

            JKREP(9) JKZONE(99);


            CLUSTER = IDCLASS;
```

```
                WEIGHT IS HOUWGT;
ANALYSIS: TYPE = COMPLEX;

                ESTIMATOR = MLR;

MODEL:      RI BY ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr;

                MI BY ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er;

                RI ON SACHR SACHM;

                MI ON SACHM SACHR;

! Correlated error terms positively-worded items (i.e. reverse-coded)
 ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er WITH
 ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er;

! Correlated uniqueness for negatively-worded items (i.e. the rest)
ASBR07A ASBR07D ASBM01B ASBM01C WITH
ASBR07A ASBR07D ASBM01B ASBM01C;

 OUTPUT:    TECH4 STANDARDIZED;
```

MPLUS Syntax for the BFLPE for Pupils' Intrinsic Motivation for Mathematics

```
TITLE:      BFLPE for Interest (Maths)
DATA:       FILE IS Filelist.dat;
            TYPE = IMPUTATION;
 VARIABLE: NAMES ARE
            IDSCHOOL IDCLASS IDSTUD ITSEX
            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr
            ASBR08Ar ASBR08Br ASBR08C ASBR08G
            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er
            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G
            ASDAGE TOTWGT HOUWGT SENWGT
            WGTADJ1 WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3
            JKREP JKZONE MACHW RACHW;

            USEVARIABLES ARE
            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er !maths interest
            MACHW !Individual math achievement
            MACHB;      ! Class-level average (defined below)

            MISSING ARE
            ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)
            JKREP(9) JKZONE(99);

            CLUSTER = IDSCHOOL IDCLASS;
```

```
   WEIGHT IS HOUWGT;

   WITHIN = MACHW;

   BETWEEN = MACHB;

   DEFINE:    CENTER MACHW (GRANDMEAN);

              MACHB = CLUSTER_MEAN(MACHW);

ANALYSIS:  TYPE= COMPLEX TWOLEVEL; ESTIMATOR=MLR;

       MODEL: %WITHIN%
       ! the following is the measurement part of the model.
       MINTw BY ASBM01Ar
       ASBM01B
       ASBM01C
       ASBM01Dr
       ASBM01Er;

   ! Correlated error terms for positively-worded items (i.e. reverse-coded)
   ASBM01Ar ASBM01Dr ASBM01Er WITH ASBM01Ar ASBM01Dr ASBM01Er;
   ! Correlated error terms for negatively-worded items (i.e. the rest)
   ASBM01B ASBM01C WITH ASBM01B ASBM01C;

   ! The regressions/structural part is next
```

```
    ! The parts in parentheses label certain estimates
    ! which can be used for other calculations. For example,
    ! b_wMM1 means the beta (b) for within-level effect (w) with maths interest
    ! regressed on maths score (MM). A similar naming convention applies
    ! to the other estimates.

    ! Maths SI predicted by maths achievement
    MINTw on MACHW;
    ! Variances:
    MINTw (wfIntm1);

    %BETWEEN%
    MINTb BY ASBM01Ar
    ASBM01B
    ASBM01C
    ASBM01Dr
    ASBM01Er;

    ! BFLPE in Maths
    MINTb on MACHB (b_bMM1);
    MACHB (bfmach1);

MODEL CONSTRAINT:
! This function allows the BFLPE to be computed
! Contextual effect sizes
NEW(bfmm1);
```

```
! maths bflpe
! Calculation of effect size (Marsh et al., 2009; Guo et al., 2018):
! ES=2 × ß ×  SD_predictor/SD_outcome
! Which they specify as bfmm1 = 2*(b_bMM1)*(bfmach1**.5)/(wfscm1**.5);
! Which implies that it is interpreted as:
! ES=2 × ß_Between  ×  SD_predictorB/(SD_outcome W)
! Using this formula:
! Maths effect
bfmm1 = 2*(b_bMM1)*(bfmach1**.5)/(wfIntm1**.5);
```

OUTPUT: STANDARDIZED;

MPLUS Syntax for the BFLPE for Pupils' Intrinsic Motivation for Reading

```
TITLE:      BFLPE for Interest (Reading)
DATA:       FILE IS Filelist.dat;
            TYPE = IMPUTATION;
VARIABLE:   NAMES ARE
            IDSCHOOL IDCLASS IDSTUD ITSEX
            ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr
            ASBR08Ar ASBR08Br ASBR08C ASBR08G
            ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er
            ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G
            ASDAGE TOTWGT HOUWGT SENWGT
            WGTADJ1 WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3
            JKREP JKZONE MACHW RACHW;

            USEVARIABLES ARE
             ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr
            !reading interest
            RACHW       !Individual reading achievement
            RACHB;      ! Class average reading (defined below)

            MISSING ARE
            ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)
            JKREP(9) JKZONE(99);

            CLUSTER = IDSCHOOL IDCLASS;
```

COMPARISON PROCESSES AND INTRINSIC MOTIVATION

WEIGHT IS HOUWGT;

WITHIN = RACHW;

BETWEEN = RACHB;

DEFINE: CENTER RACHW(GRANDMEAN);

RACHB = CLUSTER_MEAN(RACHW);

ANALYSIS: TYPE= COMPLEX TWOLEVEL; ESTIMATOR=MLR;

MODEL: %WITHIN%

!the following is the measurement part of the model

RINTw BY ASBR07A

ASBR07Br

ASBR07Cr

ASBR07D

ASBR07Er

ASBR07Fr;

! Correlated error terms for positively-worded items (i.e. reverse-coded)

ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr with

ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr;

! Correlated error terms for negatively-worded items (i.e. the rest)

ASBR07A ASBR07D with

```
  ASBR07A ASBR07D;


! The regressions/structural part is next
! The parts in parentheses label certain estimates
! which can be used for other calculations. For example,
! b_wRR1 means the beta (b) for within-level effect (w) with reading interest
! regressed on maths score (RR). A similar naming convention applies
! to the other estimates.


  ! Reading SI predicted by reading achievement
  RINTw on RACHW (b_wRR1);
  ! Variances:
  RINTw (wfIntr1);


  %BETWEEN%
  RINTb BY ASBR07A
        ASBR07Br
        ASBR07Cr
        ASBR07D
        ASBR07Er
        ASBR07Fr;


  ! BFLPE in Reading
  RINTb on RACHB (b_bRR1);


  RACHB (bfrach1);
```

```
MODEL CONSTRAINT:    ! This function allows the bflpe to be computed
                     ! contextual effect sizes
                     NEW(bfrm1); ! reading bflpe

 ! Calculation of effect size (Marsh et al, 2009; Guo et al., 2018):
 ! ES=2 × ß ×  SD_predictor/SD_outcome

bfrm1 = 2*(b_bRR1)*(bfrach1**.5)/(wfIntr1**.5);

OUTPUT: STANDARDIZED;
```

MPLUS Syntax for the Combined Model

```
! This model looks at cross-domain effects at the within level
! and uses interest scores as the criterion variable

TITLE:      BFLPE for Interest (Maths and English)
DATA:       FILE IS Filelist.dat;
            TYPE = IMPUTATION;

VARIABLE: NAMES ARE
                IDSCHOOL IDCLASS IDSTUD ITSEX
                ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr
                ASBR08Ar ASBR08Br ASBR08C ASBR08G
                ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er
                ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G
                ASDAGE TOTWGT HOUWGT SENWGT
                WGTADJ1 WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2
                WGTFAC3
                JKREP JKZONE MACHW RACHW MACHW1 RACHW1;

USEVARIABLES ARE
ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr  !reading interest
ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er     !maths interest
MACHW1 RACHW1 MACHB RACHB;     ! These are the class-level averages (defined
below)

MISSING ARE
```

ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)

JKREP(9) JKZONE(99);

CLUSTER = IDSCHOOL IDCLASS;

WEIGHT IS HOUWGT;

WITHIN = MACHW1 RACHW1;

BETWEEN = MACHB RACHB;

DEFINE:

CENTER MACHW1 RACHW1(GRANDMEAN);

MACHB = CLUSTER_MEAN(MACHW1);

RACHB = CLUSTER_MEAN(RACHW1);

ANALYSIS: TYPE= COMPLEX TWOLEVEL; ESTIMATOR=MLR; H1ITERATIONS = 20000;ITERATIONS = 4000;

MODEL:

%WITHIN%

! the following is the measurement part of the model. MINT and RINT are the maths and

! reading interest latent variables

```
RINTw BY ASBR07A
        ASBR07Br
        ASBR07Cr
        ASBR07D
        ASBR07Er
        ASBR07Fr;
MINTw BY ASBM01Ar
        ASBM01B
        ASBM01C
        ASBM01Dr
        ASBM01Er;

! Correlated error terms for positively-worded items (i.e. reverse-coded)
ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er with
ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er;

! Correlated error terms for negatively-worded items (i.e. the rest)
ASBR07A ASBR07D ASBM01B ASBM01C with
ASBR07A ASBR07D ASBM01B ASBM01C;

MACHW1 RACHW1 with MACHW1 RACHW1;
MINTw RINTw with MINTw RINTw;

! The regressions/structural part is next
! The parts in parentheses are how you can tell MPLUS to label certain estimates
! which you can use for other calculations. For example,
```

```
! b_wMM1 means the beta (b) for within-level effect (w) with maths interest
! regressed on maths score (MM). A similar naming convention applies
! to the other estimates.

! Maths SC predicted by maths achievement
MINTw on MACHW1 (b_wMM1)
        RACHW1 (b_wMR1);
! Reading SC predicted by reading achievement
RINTw on RACHW1 (b_wRR1)
        MACHW1 (b_wRM1);

! Variances:
RINTw (wfIntr1);
MINTw (wfIntm1);
MACHW1 (wfmach1);
RACHW1 (wfrach1);

%BETWEEN%
RINTb BY ASBR07A
        ASBR07Br
        ASBR07Cr
        ASBR07D
        ASBR07Er
        ASBR07Fr;
MINTb BY ASBM01Ar
        ASBM01B
```

```
        ASBM01C
        ASBM01Dr
        ASBM01Er;


! BFLPE in Maths
MINTb on MACHB (b_bMM1)
          RACHB@0 (b_bMR1);


! BFLPE in Reading
RINTb on RACHB (b_bRR1)
          MACHB@0 (b_bRM1);


RINTb (bfIntr1);
MINTb (bfIntm1);
MACHB (bfmach1);
RACHB (bfrach1);


MODEL CONSTRAINT:
! This function allows the bflpe to be computed
! contextual effect sizes
NEW(bfmm1);          ! maths bflpe
NEW(bfrm1);          ! reading bflpe


! Equation for effect size (Guo et al., 2018)
! ES=2 × ß ×  SD_predictor/SD_outcome
! Which they specify as bfmm1 = 2*(b_bMM1)*(bfmach1**.5)/(wfscm1**.5);
```

```
! Which implies that it is interpreted as:

! ES=2 × ß_Between  ×  SD_predictorB/(SD_outcome W)

! Using this formula:

! Maths effect

bfmm1 = 2*(b_bMM1)*(bfmach1**.5)/(wfIntm1**.5);

! Reading effect

bfrm1 = 2*(b_bRR1)*(bfrach1**.5)/(wfIntr1**.5);


OUTPUT: STANDARDIZED;
```

MPLUS Syntax for the Mediation Analysis

```
! This model looks at cross-domain effects at the within level
! and uses interest scores as the criterion variable
! and self-concept as a mediator variable

! This dataset also presents problems with invariant clusters,
! so based on a previous output warning,
! the following cases were removed in SPSS prior to analysis:
!       Variable   Cluster (i.e. Class) IDs with no within-cluster variation

 !    ASBR08AR    4401 3802 14702
 !    ASBR08BR    4002
 !    ASBR08G     1501 3802 4601 4301 14001 13301
 !    ASBM03AR    1101
 !    ASBM03B     1301

TITLE:          BFLPE for Interest (Maths and English)
DATA:           FILE IS Filelist.dat;
                TYPE = IMPUTATION;

VARIABLE:   NAMES ARE
        IDSCHOOL IDCLASS IDSTUD ITSEX
        ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr
        ASBR08Ar ASBR08Br ASBR08C ASBR08G
        ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er
        ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G
```

```
    ASDAGE TOTWGT HOUWGT SENWGT
    WGTADJ1 WGTADJ2 WGTADJ3 WGTFAC1 WGTFAC2 WGTFAC3
    JKREP JKZONE MACHW RACHW MACHW1 RACHW1;

USEVARIABLES ARE
ASBR07A ASBR07Br ASBR07Cr ASBR07D ASBR07Er ASBR07Fr  !reading interest
ASBM01Ar ASBM01B ASBM01C ASBM01Dr ASBM01Er     !maths interest
ASBR08Ar ASBR08Br ASBR08C ASBR08G !Reading Self-concept
ASBM03Ar ASBM03B ASBM03C ASBM03Dr ASBM03G  ! Maths self-concept
MACHW1 RACHW1 MACHB RACHB;     ! These are the class-level averages (defined
below)

MISSING ARE
ITSEX(9) ASBR07A-ASBM03G(9) TOTWGT-WGTFAC3(999999.000000)
JKREP(9) JKZONE(99);

CLUSTER = IDSCHOOL IDCLASS;

WEIGHT IS HOUWGT;

!The following identifies the variables with ONLY within-level variance
WITHIN = MACHW1 RACHW1;

!variables with ONLY between-level variance
! Here, the mean classroom achievement scores
BETWEEN = MACHB RACHB;
```

!variables not specified here have variances on both levels (i.e. Interest, SC)

DEFINE:

CENTER MACHW1 RACHW1(GRANDMEAN);

MACHB = CLUSTER_MEAN(MACHW1);

RACHB = CLUSTER_MEAN(RACHW);

ANALYSIS: ESTIMATOR=MLR; TYPE= COMPLEX TWOLEVEL; H1ITERATIONS = 20000;ITERATIONS = 4000;

MODEL: %WITHIN%

! the following is the measurement part of the model. MINT and RINT are the maths and

! reading interest latent variables. MSC and RSC are measures of self-concept.

RINTw BY ASBR07A

ASBR07Br

ASBR07Cr

ASBR07D

ASBR07Er

ASBR07Fr;

MINTw BY ASBM01Ar

ASBM01B

ASBM01C

ASBM01Dr

ASBM01Er;

RSCw BY ASBR08Ar

ASBR08Br

ASBR08C

ASBR08G;

MSCw BY ASBM03Ar

ASBM03B

ASBM03C

ASBM03Dr

ASBM03G;

! Correlated error terms for positively-worded items (i.e. reverse-coded)

ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr ASBM01Ar ASBM01Dr ASBM01Er

ASBM03Ar ASBM03Dr ASBR08Ar ASBR08Br with ASBR07Br ASBR07Cr ASBR07Er ASBR07Fr

ASBM01Ar ASBM01Dr ASBM01Er ASBM03Ar

ASBM03Dr ASBR08Ar ASBR08Br;

! Correlated error terms for negatively-worded items (i.e. the rest)

ASBR07A ASBR07D ASBM01B ASBM01C ASBM03B ASBM03C ASBM03G ASBR08C ASBR08G with

ASBR07A ASBR07D ASBM01B ASBM01C ASBM03B ASBM03C ASBM03G ASBR08C ASBR08G;

! The regressions/structural part is next

! The parts in parentheses are for MPLUS to label certain estimates

! which can be used for other calculations. Given the complexity

! of the multiple mediation paths, the regression coefficients

! will be labelled for simplicity. MA corresponds to maths achievement

! MS corresponds with maths self-concept, and MI corresponds with

! maths interest (similar for reading, but replace the M with R).

```
! The coefficients are labelled according to a combination of these
! with predictor followed by criterion, followed by an underscore
! and a number to indicate the level (1 or 2)
! For example, the coefficient for maths SC regressed on reading achievement
! is labelled RAMS_1

! Maths Interest predicted by maths achievement and reading achievement
! Includes SC also
MINTw on MACHW1 (MAMI_1)
        RACHW1 (RAMI_1)
        MSCw (MSMI_1)
        RSCw (RSMI_1);

! Reading Interest predicted by reading achievement, etc.
RINTw on RACHW1 (RARI_1)
        MACHW1 (MARI_1)
        RSCw (RSRI_1)
        MSCw (MSRI_1);

!SC predicted by achievement variables
MSCw on MACHW1 (MAMS_1)
       RACHW1 (RAMS_1);
RSCw on RACHW1 (RARS_1)
       MACHw1 (MARS_1);
```

```
! Residual variances:
RINTw (wfIntr1);
MINTw (wfIntm1);
MACHW1 (wfmach1);
RACHW1 (wfrach1);
RSCw (wfrsc1);
MSCw (wfmsc1);

%BETWEEN%

!latent measures
RINTb BY ASBR07A
        ASBR07Br
        ASBR07Cr
        ASBR07D
        ASBR07Er
        ASBR07Fr;
MINTb BY ASBM01Ar
        ASBM01B
        ASBM01C
        ASBM01Dr
        ASBM01Er;
RSCb BY ASBR08Ar
       ASBR08Br
       ASBR08C
       ASBR08G;
```

```
MSCb BY ASBM03Ar
       ASBM03B
       ASBM03C
       ASBM03Dr
       ASBM03G;

MINTb on MACHB (MAMI_2) ! BFLPE in Maths
          RACHB@0 (RAMI_2)       ! Set at zero variance even though it's
   MSCb@0 (MSMI_2)        ! not essential. Easier to change
   RSCb@0 (RSMI_2);       ! if L2 variance starts to be
                           ! of interest

RINTb on RACHB (RARI_2) ! BFLPE in Reading
          MACHB@0 (MARI_2)
   RSCb@0 (RSRI_2)
   MSCb@0 (MSRI_2);

MSCb on MACHB (MAMS_2); !this is the L2 maths mediation
RSCb on RACHB (RARS_2); !this is the L2 reading mediation

!Residual variances
RINTb (bfIntr1);
MINTb (bfIntm1);
MACHB (bfmach1);
RACHB (bfrach1);
RSCb (bfrsc1);
```

MSCb (bfmsc1);

MODEL CONSTRAINT:

! This function allows the bflpe to be computed

! contextual effect sizes for SC and SI

NEW(bfmm1); ! maths bflpe for interest

NEW(bfmm2); ! maths bflpe for self-concept

NEW(bfrm1); ! reading bflpe for interest

NEW(bfrm2); ! reading bflpe for self-concept

! Effect size formula (Marsh et al., 2009; Guo, 2018)

! ES=2 × ß × SD_predictor/SD_outcome

! Which they specify as bfmm1 = 2*(b_bMM1)*(bfmach1**.5)/(wfscm1**.5);

! Which implies that it is interpreted as:

! ES=2 × ß_Between × SD_predictorB/(SD_outcome W)

! Using this formula:

! Maths bflpe for interest (bfmm1) and SC (bfmm2)

bfmm1 = 2*(MAMI_2)*(bfmach1**.5)/(wfIntm1**.5);

bfmm2 = 2*(MAMS_2)*(bfmach1**.5)/(wfmsc1**.5);

! Reading effect for interest (bfrm1) and SC (bfrm2)

bfrm1 = 2*(RARI_2)*(bfrach1**.5)/(wfIntr1**.5);

bfrm2 = 2*(RARS_2)*(bfrach1**.5)/(wfrsc1**.5);

```
! The following model constraint is the indirect effect of the
! combined variances of slopes a and b in each mediation pathway
! (there are four of interest here for each subject interest DV)

! The naming convention is the same as above, but with
! each of the three variables in an indirect mediation pathway
! represented by two letters each, and
! concatenated in the order x -> m -> y (i.e. XXMMYY)
! followed by a number to indicate level (1 = within, 2 = between).
! This is just in case the hypotheses are extended into L2
! - at present there are no L2 computations here.

NEW(MAMSMI_1); ! Maths Ach -> Maths SC -> Maths interest at level 1
NEW(MARSMI_1); ! Maths Ach -> Read SC -> Maths interest at level 1
NEW(RAMSMI_1); ! Read Ach -> Maths SC -> Maths interest at level 1
NEW(RARSMI_1); ! Read Ach -> Read SC -> Maths interest at level 1

MAMSMI_1 = MAMS_1*MSMI_1; !Within-level indirect effect
MARSMI_1 = MARS_1*RSMI_1; !Within-level indirect effect
RAMSMI_1 = RAMS_1*MSMI_1; !Within-level indirect effect
RARSMI_1 = RARS_1*RSMI_1; !Within-level indirect effect

NEW(RAMSRI_1); ! Read Ach -> Maths SC -> Reading interest at level 1
NEW(RARSRI_1); ! Read Ach -> Read SC -> Reading interest at level 1
NEW(MAMSRI_1); ! Maths Ach -> Maths SC -> Reading interest at level 1
```

```
NEW(MARSRI_1); ! Maths Ach -> Read SC -> Reading interest at level 1


RAMSRI_1 = RAMS_1*MSRI_1; !Within-level indirect effect

RARSRI_1 = RARS_1*RSRI_1; !Within-level indirect effect

MAMSRI_1 = MAMS_1*MSRI_1; !Within-level indirect effect

MARSRI_1 = MARS_1*RSRI_1; !Within-level indirect effect



OUTPUT: STANDARDIZED CINTERVAL;
```

Appendix I

CFA Model One: Correlations Between Factor Indicators of Latent Constructs

Item Code	Item Code				
	ASBR07A	ASBR07BR	ASBR07CR	ASBR07D	ASBR07ER
ASBR07A	1.000				
ASBR07BR	0.154	1.000			
ASBR07CR	0.278	0.364	1.000		
ASBR07D	0.466	0.268	0.501	1.000	
ASBR07ER	0.293	0.355	0.531	0.520	1.000
ASBR07FR	0.411	0.362	0.584	0.690	0.637
ASBM01AR	0.028	0.181	0.127	0.091	0.117
ASBM01B	0.128	0.150	0.122	0.180	0.124
ASBM01C	0.094	0.160	0.130	0.180	0.118
ASBM01DR	0.001	0.213	0.167	0.104	0.148
ASBM01ER	0.047	0.187	0.152	0.115	0.129
ASBR08AR	0.234	0.210	0.283	0.326	0.295
ASBR08BR	0.195	0.132	0.206	0.234	0.219
ASBR08C	0.268	0.026	0.094	0.241	0.111
ASBR08G	0.293	0.045	0.128	0.306	0.136
ASBM03AR	0.084	0.111	0.099	0.108	0.099
ASBM03B	0.182	-0.006	0.020	0.095	0.012
ASBM03C	0.150	0.047	0.024	0.140	0.020
ASBM03DR	0.054	0.100	0.080	0.051	0.104
ASBM03G	0.119	0.005	0.003	0.067	0.004
	Item Code				
	ASBR07FR	ASBM01AR	ASBM01B	ASBM01C	ASBM01DR
ASBR07FR	1.000				
ASBM01AR	0.130	1.000			
ASBM01B	0.136	0.562	1.000		
ASBM01C	0.134	0.645	0.671	1.000	
ASBM01DR	0.165	0.612	0.424	0.490	1.000
ASBM01ER	0.152	0.823	0.578	0.682	0.639
ASBR08AR	0.371	0.116	0.085	0.078	0.120

ASBR08BR	0.289	0.059	0.035	0.012	0.068
ASBR08C	0.166	0.003	0.084	0.047	-0.010
ASBR08G	0.220	-0.053	0.054	0.000	-0.020
ASBM03AR	0.150	0.498	0.308	0.335	0.380
ASBM03B	0.057	0.330	0.334	0.332	0.211
ASBM03C	0.074	0.412	0.364	0.388	0.289
ASBM03DR	0.110	0.480	0.321	0.355	0.402
ASBM03G	0.031	0.442	0.382	0.399	0.283

	Item Code				
	ASBM01ER	ASBR08AR	ASBR08BR	ASBR08C	ASBR08G
ASBM01ER	1.000				
ASBR08AR	0.113	1.000			
ASBR08BR	0.059	0.563	1.000		
ASBR08C	0.008	0.355	0.426	1.000	
ASBR08G	-0.029	0.355	0.428	0.519	1.000
ASBM03AR	0.477	0.306	0.217	0.142	0.058
ASBM03B	0.334	0.147	0.150	0.329	0.208
ASBM03C	0.405	0.156	0.137	0.217	0.176
ASBM03DR	0.463	0.230	0.223	0.133	0.057
ASBM03G	0.430	0.082	0.089	0.191	0.141

	Item Code				
	ASBM03AR	ASBM03B	ASBM03C	ASBM03DR	ASBM03G
ASBM03AR	1.000				
ASBM03B	0.471	1.000			
ASBM03C	0.531	0.600	1.000		
ASBM03DR	0.602	0.446	0.494	1.000	
ASBM03G	0.451	0.632	0.572	0.460	1.000

Note. ASBR07A = I read only if I have to; ASBR07Br = I like talking about what I read with other people; ASBR07Cr = I would be happy if someone gave me a book as a present; ASBR07D = I think reading is boring; ASBR07Er = I would like to have more time for reading; ASBR07Fr = I enjoy reading; ASBM01Ar = I enjoy learning mathematics; ASBM01B = I wish I did not have to study mathematics; ASBM01C = Mathematics is boring; ASBM01Dr = I learn many interesting things in mathematics; ASBM01Er = I like

mathematics; ASBR08Ar = I usually do well in reading; ASBR08Br = Reading is easy for me; ASBR08C = Reading is harder for me than for many of my classmates; ASBR08G = Reading is harder for me than any other subject; ASBM03Ar = I usually do well in mathematics; ASBM03B = Mathematics is harder for me than for many of my classmates; ASBM03C = I am just not good at mathematics; ASBM03Dr = I learn things quickly in mathematics; ASBM03G = Mathematics is harder for me than any other subject.

Appendix J

SPSS Syntax

SPSS Syntax: Class Identification Numbers of Removed Classes due to Cluster Invariance in Pupils' Responses to Intrinsic Motivation Scales

```
Select if (IDCLASS ne 601 and IDCLASS ne 701 and IDCLASS ne 1001 and IDCLASS ne
1101 and IDCLASS ne 1301
and IDCLASS ne 1501 and IDCLASS ne 1701 and IDCLASS ne 1801 and IDCLASS ne
2001 and IDCLASS ne 3802 and
IDCLASS ne 4002 and IDCLASS ne 4301 and IDCLASS ne 5801 and IDCLASS ne 6002
and IDCLASS ne 14502).
EXECUTE.
```

List of Additional Class Identification Numbers which were Removed Prior to Completing the Mediation Analysis.

Class identification numbers (IDCLASS): 4401, 4601, 13301, 14001, 14702

Appendix K

Results for Each Model with Cluster Invariances Removed

Results for CFA model three, the basic I/E model, the BFLPE for pupils' intrinsic motivation for mathematics, the BFLPE for pupils' intrinsic motivation for reading, and the combined model, with cluster invariances for ASC and intrinsic motivation measures removed.

Table K 1

Model Fit Indices for Each Model

Model	χ^2 *df*	SRMR (*SD*)	CFI (*SD*)	TLI (*SD*)	RMSEA (*SD*)	AIC (*SD*)	BIC (*SD*)
CFA Model 3	513.381 *df* = 94	.035	.979	.957	.033	180825.549	181682.893
Basic I/E	455.953 *df* = 34	.043 (.001)	.964 (.005)	.919 (.011)	.055 (.004)	103853.241 (26.592)	104263.001 (26.592)
BFLPE M	57.178 *df* = 14	.015 (<.001) [within] .044 (.002) [between]	.995 (<.001)	.989 (.001)	.028 (.001)	46227.483 (10.616)	46435.515 (10.616)
BFLPE R	695.146 *df* = 21	.046 (.001) [within] .119 (.009) [between]	.916 (.003)	.833 (.005)	.089 (.001)	61013.758 (24.985)	61272.222 (24.985)
Combined	1239.245 *df* = 98	.046 (.001) [within] .174 (.002) [between]	.953 (.001)	.926 (.002)	.054 (.001)	120715.852 (56.124)	121339.947 (56.124)

Note. χ^2 = Chi-Square Test of Model Fit, in all three models $p < .001$; df = degrees of freedom; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = Root Mean Square Error of Approximation; CI = Confidence Interval; AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion.

Table K 2

Estimated Factor Loadings for Indicators of Each Latent Construct, as per CFA Model Three

Indicator	Scale			
	Intrinsic Motivation Mathematics	Mathematics Self-Concept	Intrinsic Motivation Reading	Reading Self-Concept
1	.889	.686	.431	.764
2	.638	.603	.424	.819
3	.736	.703	.674	.483
4	.689	.702	.759	.495
5	.928	.694	.735	
6			.884	

Note. The number of each scale indicator corresponds with the number in the measures section of the empirical paper (see section 3.2.3.). For all estimated factor loadings, $p < .001$.

Table K 3

Model Three: Estimated Correlation Matrix for the Latent Variables

Latent Construct	Latent Construct			
	IMR	IMM	RSC	MSC
IMR	1.000			
IMM	0.198	1.000		
RSC	0.433	0.019	1.000	
MSC	0.105	0.678	0.205	1.000

Note. IMR = Intrinsic Motivation for Reading; IMM = Intrinsic Motivation for Mathematics; RSC = Reading Self-Concept; MSC = Mathematics Self-Concept.

Table K 4

Standardised Parameter Estimates and Standard Errors for Each Model

Model	Intrinsic Motivation Reading β *(SE)*	intrinsic motivation Mathematics β *(SE)*
Basic I/E Model		
Reading achievement	.376 (.043)***	-.017 (.123)
Mathematics achievement	-.021 (.035)	.096 (.132)
BFLPE Mathematics		
Individual achievement		.134 (.023)***
Class-average achievement		-.305 (.106)**
BFLPE Reading		
Individual achievement	0.368 (.032)***	
Class-average achievement	-.031 (.302)	
Combined Model		
Individual reading achievement	.413 (.037)***	-.101 (.088)
Individual mathematics achievement	-.035 (.034)	.222 (.092)*
Class-average reading achievement	-.079 (.259)	
Class-average mathematics achievement		-.306 (.110)**

Note. *p < .05, **p < .01, ***p < .001

Note. For the basic I/E model, correlations between the manifest and latent constructs indicated that mathematics achievement and reading achievement were strongly correlated with each other (r = .790), while the there was only a small correlation between intrinsic

motivation for mathematics and intrinsic motivation for reading (r = .193); There was a significant contextual effect for the BFLPE for pupils' intrinsic motivation for mathematics, (ES = .220. p < .01); There was no significant contextual effect for the BFLPE for pupils' intrinsic motivation for reading (ES = .007, p = .912); Within the combined model, there was a significant contextual effect for mathematics (ES = -.217, p = .01), but not for reading (ES = -.016, p = .682).

Appendix L

Parent/Guardian Information and Consent Form Detailing Changes to the Study

Title of Study: 'The Relevance of the Integrated Internal/External Frame of Reference Model to Second Class Pupils' Mathematics and Reading Self-Concepts'

Name of Investigator: Aoife Cassidy, Mary Immaculate College (University of Limerick)

Introduction

Thank you for taking the time to read this and for your interest in this study.

Before the school closures, you gave permission for your child to participate in the study entitled 'The Relevance of the Integrated Internal/External Frame of Reference Model to Second Class Pupils' Mathematics and Reading Self-Concepts'. Children were then told about the study and asked if they would like to participate.

Your child then completed a questionnaire in which they rated their ability to learn mathematics and their ability to learn how to read.

It was hoped that pupils second-class standardized test scores could be used to look at relationships between pupils' beliefs in their ability to learn and pupils' performance on standardized tests. Due to the recent school closures, second-class standardized tests have been cancelled.

The research is now hoping to use pupils' first-class standardized test scores instead of their second-class standardized test scores. Permission is needed from parents of participants to access these scores.

Who is undertaking the study?

My name is Aoife Cassidy. I am a second-year student on the Doctorate in Educational and Child Psychology programme at Mary Immaculate College and this research will form part of my thesis. The research is being undertaken under the supervision of Dr. Stella Long and Dr. Paul Mulcahy.

What is the project about?

This project aims to investigate second class pupils' beliefs about their ability to learn mathematics and their ability to learn reading skills. The study will investigate relationships

between pupils' beliefs about their ability to learn and their first-class end-of-year assessment scores.

What will the changes to the study involve?

Access to pupils' first-class standardized test scores for mathematics and reading, instead of their second-class standardized test scores.

Should I tell my son about the changes?

Your son was told about the study before he took part. He was asked if he would like to participate. It is advised that he is told about changes to the study too, and that any questions that he has about the study and/or his participation are answered.

What will stay the same?

Voluntary participation and right to withdraw: Participation in this study is completely voluntary. You or your child can decide to withdraw from the study at any time, without giving any reason.

Benefits: The questionnaire that your child completed in school for the purpose of this study may also be used by the school to inform future planning, teaching and learning.

Confidentiality: Only the principal researcher (Aoife Cassidy), and her supervisors for the doctoral thesis will have access to the data that is collected. It will not be shared with anybody else. It will be stored, anonymously, in a locked storing file which may only be accessed by the principal researcher and her supervisors for the purpose of completing the doctoral thesis. While it is intended that the findings will be published in the research thesis, at relevant conferences, and/or in research journals, your individual child's participation and data will remain completely anonymous and will not be identifiable in any way.

Who to Contact?

If you have any questions you are welcome to contact either myself or either of my research supervisors.

Aoife Cassidy Email: 11111801@micstudent.mic.ul.ie

Dr. Stella Long Email: stella.long@mic.ul.ie

Dr. Paul Mulcahy Email: paul.mulcahy@mic.ul.ie

If you have concerns about this study and wish to contact someone independent, you may contact: Mary Collins, MIREC Administrator, Research and Graduate School, Mary Immaculate College, South Circular Road, Limerick. Telephone: 061-204980 / E-mail: mirec@mic.ul.ie

Consent Form for Parents

I _____ give consent for my son to continue participating in the research project entitled 'The Relevance of the Integrated Internal/External Frame of Reference Model to Second Class Pupils' Mathematics and Reading Self-Concepts'

I have read the 'Parent/Guardian Information Sheet Detailing Changes to the Study'.

I know that pupils' first-class standardised test scores will now be used in the study.

Appendix M

Confirmation of Consent to Access and Use the 2011 TIMSS and PIRLS International Database; Confirmation that the Database May be Freely Used by Researchers

RE: Permission to use 2011 TIMSS and PIRLS data from Irish participants for doctoral thesis.

Permission Requests <permission.requests@iea.nl>
Mon 29/06/2020 14:44

To: AOIFE CASSIDY (Student) <11111801@micstudent.mic.ul.ie>
Cc: Stella Long <Stella.Long@mic.ul.ie>; Paul Mulcahy <Paul.Mulcahy@mic.ul.ie>; Permission Requests <permission.requests@iea.nl>

Dear Aoife Cassidy,

Thank you for your email and your interest in using IEA data in your own work. If I understand your request correctly, you would like to analyze data from TIMSS/PIRLS 2011 as part of your doctoral thesis. Please note that all IEA data is open-access and freely accessible through the IEA data repository (https://www.iea.nl/data-tools/repository). From the same website you can also download the IEA IDB Analyzer, a software tool that provides assistance with the analysis of data from large-scale assessments. We strongly recommend that you use the IDB Analyzer when working with IEA data.

Please note that prior permission from IEA is only necessary if you would like to use the instruments (or items) from TIMSS and/or PIRLS in your research, for example in order to collect data yourself. In that case, please provide more information about the intended use of the TIMSS/PIRLS materials.

I hope this information is helpful.

Please do not hesitate to contact us if you have any questions or concerns.

Have a nice day.
Best wishes,
Jan

Jan-Philipp Wagner, M.A.
Junior Research Officer
j.wagner@iea.nl

IEA

Researching education, improving learning

International Association for the Evaluation of Educational Achievement
Tel. +31 20 625 3625 | Fax +31 20 420 7136 | **www.iea.nl**
IEA Amsterdam | Keizersgracht 311 | 1016 EE Amsterdam | The Netherlands

Appendix N

Academic Self-Concept Questionnaires Which Were to Be Used in The Original Study (Pre-Covid-19)

Mathematics Self-Concept Questionnaire to be used in Original Study (Pre-COVID)

Mathematics Self-concept Questionnaire

1. I usually do well in **mathematics**

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

2. I learn things quickly in **mathematics**

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

3. **Mathematics** is harder for me than for many of my classmates

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

4. **Mathematics** is easy for me

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

5. I am just not good at **mathematics**

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

6. **Mathematics** is harder for me than any other subject

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

Reading Self-Concept Questionnaire to be used in Original Study (Pre-COVID)

Reading Self-concept Questionnaire

1. I usually do well in **reading**

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

2. I learn new **reading** skills quickly

(for example, learning to read new words that I did not come across before)

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

3. **Reading** is harder for me than for many of my classmates

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

4. **Reading** is easy for me

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

5. I am just not good at **reading**

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree

6. **Reading** is harder for me than any other subject

Strongly agree	Agree	Neither agree or disagree	Disagree	Strongly disagree