Effects of the Formative Approach on Students' Learning and Motivation

Inaugural dissertation
submitted in partial fulfillment of the requirements for the degree
Doctor of Social Sciences
in the
Graduate School of Economic and Social Sciences
at the
University of Mannheim

by

Saraswathi Ramasamy

Supervisors:
Prof. Dr. Oliver Dickhäuser
Prof. Dr. Karina Karst

Dean of the School of Social Sciences:
Prof. Dr. Michael Diehl

Thesis Evaluators:
Prof. Dr. Oliver Dickhäuser
Prof. Dr. Karina Karst
Prof. Dr. Stefan Münzer

Thesis Defense:
25.03.2021

Acknowledgements

First of all I would like to thank my supervisor Prof. Dr. Oliver Dickhäuser for guiding me through my doctoral studies and helping me reach this point. I am deeply grateful and feel fortunate to have worked with you. You have always been encouraging, supportive, and readily available for discussions and provided valuable feedback. My sincere thanks to you for counselling me on my professional career over and beyond my doctoral pursuits. You have inspired me and helped me in numerous ways probably unknown to you. Thanks for believing in me even in moments when I lost faith in myself.

My sincere gratitude to my second supervisor Prof. Dr. Karina Karst, who has helped me with her timely suggestions and who I look up to as my role model for her disciplined work ethics.

I would like to thank Prof. Dr. Stefan Münzer for his valuable suggestions, feedback and most importantly for always being happy to help and discuss ideas.

Thanks to my wonderful colleagues Belinda Merkle and Julia Hein who helped me with their feedback on parts of my dissertation, especially on the initial rough drafts.

Thanks to Marc-Philipp Janson and Benedict Fehringer for always willingly clarifying my questions on statistics.

Thanks to Stefan Janke, Tamara Marksteiner and Tim Kühl for caringly enquiring on the progress of my dissertation and the readiness to offer help and support.

Thanks to my office-mate and my friend Hanna Beißert who has been a great support and whose presence in office made my work days more productive and livelier.

Thanks to Gabi Atkinson for willingly helping out with a whole lot of things, readily lending ears to any queries or concerns and most of all for being a wonderful person.

Thanks to my graduate school friends Sophie Scharf, Liliane Wulff, Mirka Henninger, Stella Guldner, Sebastian Siehl and Lucas Lörch for taking interest in my research, rendering support and actively discussing my research ideas.

What started out as a collegial relationship developed into a profound friendship with Anne-Sophie Waag and Meike Bonefeld. Thank you both for being immensely supportive in every possible way. Thank you very much for your feedback on my dissertation. I am glad to have gotten to know you both and grateful for having you as my friends.

I take this opportunity to thank my parents for enabling me the pursuit of higher education and doctoral studies in a foreign country. Thank you for your unconditional love and support. Thanks to my little sister who takes pride in my accomplishments.

Finally, my loving thanks to my better half who always takes pleasure in my accomplishments, stood by me through the ups and downs of my doctoral career, has been massively supportive in all my endeavours and particularly being helpful in the much needed end phase of my dissertation.

Table of Contents

List of Tables 8
List of Figures 9
Abstract 10
Prologue 13
Research Problem 14
Outcomes of the Formative Approach 15
1 The Formative Approach: A Conceptualization and a Model 17
1.1 Theoretical Background - Perspectives on Formative Approach 18
1.2 The Formative Approach and its Components 21
1.2.1 Diagnostic Assessment 21
1.2.2 Comprehensive Feedback 22
1.2.3 Adaptive Instruction 24
1.3 Definition of Formative Approach 26
1.4 Proposed Model of Formative Approach 27
1.5 A Model of the Formative Approach 27
2 The Formative Approach: A Systematic Review and a Meta-Analysis 29
2.1 Systematic Review 29
2.1.1 Eligibility Criteria 29
2.1.2 Literature Search 30
2.1.3 Study Selection 31
2.1.4 Coding of Primary Studies 33
2.1.5 Results 34
2.1.6 Conclusions 35
2.1.7 Limitations 36
2.1.8 Next Steps 36
2.2 A Meta-Analysis on the Formative Approach 37
2.2.1 Formative Approach and Motivation 38
2.2.2 Formative Approach and Learning 39
2.2.3 Earlier Meta-analyses on Formative Approach 40
2.2.4 Moderators of the Formative Approach 45
2.2.5 The Current Meta-Analysis 50
2.2.6 Research Questions and Hypotheses 50

2.2.7 Methods 53
2.2.8 Analysis 57
2.2.9 Results 58
2.2.10 Discussion 70
2.2.11 Strengths and Limitations 76
2.2.12 Next Steps 77
2.3 An Explorative Study 78
3 An Experimental Study on the Effects of Formative Approach on Learning & Motivation 81
3.1 Formative Approach and Learning 81
3.2 Formative Approach and Motivation 81
3.3 Transfer Learning Effects of the Formative Approach 84
3.4 Current Study 84
3.5 Research Questions 85
3.6 Method 86
3.6.1 Participants 86
3.6.2 Stratified Randomization 86
3.6.3 Test Materials 86
3.6.4 Instructors 93
3.6.5 Ethical Considerations 94
3.7 Analysis 94
3.7.1 Covariate 95
3.8 Results 95
3.9 Discussion 99
3.9.1 Near-transfer Learning 100
3.9.2 Effects on Motivation 102
3.10 Strengths and Limitations of the Experiment 104
3.11 Implications 106
3.12 Future Research 107
3.13 Conclusions 108
General Discussion 110
Conclusions 110
Current Model versus other Formative Approach Models 112
Implications 115

Limitations 117
Future Research 120
References 122
Appendix A: A List of all the Studies included in the Systematic Review and Meta-Analysis 135
Appendix B: Further Results of the Experimental Study on the Formative Approach 152
Appendix C: Test Materials 158

List of Tables

Table 1: Constellations of the Formative Approach Components 33

Table 2: Moderating Effects of the Stages of Schooling on Learning 64

Table 3: Moderating Effects of Special Education and Regular Education on Learning 64

Table 4: Moderating Effects of different Subjects on Learning 68

Table 5: Difficulty and Discrimination Indices of the Multiple-Choice-Questions 92

Table 6: Distractor Analysis: Distribution of Items with Non-functional Distractors 93

List of Figures

Figure 1. The Formative Approach Model 27

Figure 2. PRISMA Flow Diagram of the literature search 32

Figure 3. Forest Plot of a Subgroup of Studies 58

Figure 4. Funnel Plot with Outliers 59

Figure 5. Funnel Plot without Outliers 59

Figure 6. Effect Sizes and CIs of Formative Approach Components on Motivation 61

Figure 7. Effect Sizes and CIs of Formative Approach Components on Learning 63

Figure 8. Differential Effects of the Components on Near-transfer Learning 96

Figure 9. Near-transfer Learning Scores 97

Figure 10. Scores on the open-ended question 98

Figure 11. Motivation Scores 98

Abstract

Formative approach is frequently implemented in learning contexts for its supposed positive effects on students' learning and motivation. Inquiry into what is formative approach revealed multiple definitions and varying conceptualizations. The range of practices under the term formative approach is wide and fluctuating. Consequently, the reported effects of formative approach are also wide ranging. Without a clear conceptualization of formative approach, it is impractical to understand the approach and investigate its effectiveness. Neither is it feasible to amend an approach that is ambiguous. Therefore, it is essential to remove this ambiguity. This called for the research carried out in the current dissertation. In this dissertation, the aim is to conceptualize formative approach and examine its effects on students' learning and motivation.

With an initial overview on what has been understood so far as formative approach and where the gaps are, the need for a clear conceptualization of formative approach was exemplified. With the goal of arriving at a comprehensive conceptualization of the formative approach, a theoretical model of formative approach was postulated that took all specific aspects of the approach into consideration. The proposed model of the formative approach comprised the following three components, diagnostic assessment, comprehensive feedback and adaptive instruction. In the next phase, the goal was to confirm the proposed theoretical model. This goal was pursued by means of a systematic literature review. After a thorough literature review and a stringent screening of an initial pool of over four thousand records, a total of 117 records were finally obtained. As there were multiple studies in each record and thus multiple effect sizes, there were altogether 459 effect sizes from 117 records. Two independent researchers coded the primary studies and arrived at a high agreement on the formative approach components and their prevalence in four different constellations in the primary studies. With a high interrater reliability, the results

of the coding confirmed the theoretical model. The next step was to investigate the effects of the formative approach components and gather empirical evidence for the theoretical model. To test a theoretical model, it is typical to conduct a new study. In the current study, we endeavored on a unique approach of conducting a meta-analysis for testing our theoretical model. Furthermore, examining the existing evidence in the literature firstly would be more efficient than conducting a new study. Thus, conducting a meta-analysis on the effects of the formative approach became the next step.

The investigated effects of the formative approach have so far been cognitive and motivational. Examining the existing literature for empirical evidence for the current theoretical model by means of a meta-analysis resulted in significant medium size effects of the formative approach on learning and motivation. Furthermore, it would be worthwhile to know whether the individual components of the formative approach have differential effects on learning and motivation respectively, as this information, in addition to significantly enhancing the theory, determines the implementation of formative approach in terms of how the available resources are allocated and applied for the implementation. The meta-analysis resulted in significant medium size effects of the formative approach components on learning and motivation. These effects were compared to see if they were statistically different from each other, as this would provide further evidence for the model. This statistical comparison of the effect sizes for the different formative approach components is yet another uniqueness of the current research. But, the effect sizes of the different components were not significantly different from each other. Nevertheless, drawing conclusions on the absence of differential effects of formative approach components based on the meta-analytic results would not be the best idea. This is because, the components existed in varying

intensities and foci in the primary studies which posed a huge challenge to statistically distinguish the effects of the components.

Since the differential effects of the formative approach components is an important aspect to explore, for both theoretical as well as practical purposes, a new experimental study was conducted. In an online experimental study, the formative approach components were operationalized in their purest and simplest forms. The research question that remained unanswered so far could finally be answered by the results of this experiment. The results clearly exemplified the differential effects of the formative approach components on learning. The adaptive instruction component had the strongest effect on learning followed by the comprehensive feedback component and then the diagnostic assessment component. Furthermore, the components adaptive instruction and comprehensive feedback had significant long-term effects on learning respectively. These results confirmed the theoretical model of the formative approach that was proposed in an earlier chapter of this dissertation.

Starting from the theoretical model of the formative approach followed by the systematic review and meta-analysis till the experimental study confirming the differential effects of the components, this dissertation provides a profound understanding of the formative approach. The outcomes of this dissertation have theoretical as well as the practical implications. While the definition and the model of formative approach contribute to the theory, the results of the empirical research provide new perspectives and possibilities for precise implementation of the formative approach to reap maximum benefits out of it.

Prologue

Formative approach is widely applied in learning contexts. Formative approach includes practices that are intended to enhance learning, e.g. gauging students' learning by means of teacher developed assessments (Wiliam, Lee, Harrison, & Black, 2004), providing feedback that guide the students further (Wiliam & Thompson, 2008), clarifying and sharing learning intentions and criteria for success with the students (Black & Wiliam, 2009; Clarke, 2001; Wiliam, 2000), modifying instructions that suit the students' learning needs (Broadfoot et al., 2002).

Formative approach continues to be researched extensively and there are persistent attempts to incorporate formative approach in classroom practice and generally in learning contexts. An initial search we conducted on research databases resulted in tens of thousands of articles. Formative approach has been investigated at all levels of education starting from kindergarten (Simmons et al., 2015) to tertiary level education (Wesson, 2013). Formative approach is also employed in professional settings for continuous learning and further development (Bilotta, 2012). The large number of research articles on formative approach is an indication of how widespread and well-known the practice is. Formative approach is continued to be practiced owing to the seemingly positive effects of formative approach in terms of enhancing students' learning achievements.

There exist both theoretical as well as empirical findings on formative approach. One of the widely known and frequently cited article is the review by Black and Wiliam (1998) [1]. Being an extensive narrative review on the formative approach, it serves as a foundational literature for

[1] In this context, it is important to reiterate that Black & Wiliam (1998) is a narrative review and not a meta-analysis. Despite the authors explicitly mentioning it is not a meta-analysis, Black & Wiliam's (1998) article has been cited widely and incorrectly as a meta-analysis and as a strong evidence for formative approach. The frequently attributed effect size ranging from .4 to .7 is from the meta-analysis by Fuchs & Fuchs (1986) reviewed in the Black & Wiliam's (1998) article. It should be noted that Black & Wiliam's (1998) narrative review had no specific selection criteria for the studies included in their review.

researchers and practitioners of the formative approach. Formative approach practices included in this review are mainly classroom assessments and feedback practices. Based on this seminal article, theoretical models of formative approach have evolved (Cowie & Bell, 1999; Ruiz-Primo & Furtak, 2006; Sadler, 1989) and numerous empirical studies were conducted (Hebbecker & Souvignier, 2018; Ruiz-Primo & Furtak, 2007; Yin et al., 2008) which have altogether enriched our understanding and knowledge on the effects of the formative approach. The studies included in this review provide some evidence for the positive effects of formative approach on learning. But neither the list of included studies is exhaustive nor the criteria for the chosen studies were explicitly stated. However, this review managed to project formative approach in a positive light and has evoked a whole lot of research on formative approach. Despite the extensive research on formative approach for decades now, inconsistencies in the empirical findings regarding the effects of formative approach continue to exist. The theoretical models and frameworks have neither been able to explain nor eliminate the inconsistencies. The question of whether we have a clear and complete understanding of the formative approach and its effects still remains unanswered. The issue that remains open with regard to the formative approach is, the lack of a paradigm definition of formative approach leading to varied understanding, diverse practices and wide-ranging reports for effectiveness under the term formative approach.

Research Problem

It is highly relevant to fill in the current gaps and inconsistencies in conceptualization because it is still unclear why there are positive effects of the formative approach only at some instances and not consistently. Only by means of a clear conceptualization, we can test if there different aspects of the formative approach leading to different effects. As a consequence, the implementation of formative approach becomes more focussed and enables to further improve formative approach.

Identifying if there are factors moderating the effects of formative approach and analysing them will further enhance the formative approach practice and have implications on the theory of formative approach.

Thereby, the goals of this dissertation are, addressing this gap in the conceptualization by means of arriving at a comprehensive definition and model of the formative approach. Following the development of a new model, the next step is to confirm the model by means of a systematic review on the existing literature. Thereafter, the goal is to gather empirical evidence for the model by conducting a meta-analysis on the effects of the formative approach. As a final step, an experimental study is conducted to examine the effects of the individual aspects of the formative approach.

Outcomes of the Formative Approach

Effects of the formative approach could be classified into two categories. One is the cognitive effect and the other is the motivational effect (Black & Wiliam, 2009). Apart from the formative approach inducing effects on students' cognition and motivation, students also constantly attempt to strike a balance between their learning achievement and motivation (Boekaerts & Corno, 2005). But the amount of research and focus on each of the categories is different. While the cognitive effects have been extensively investigated with thousands of articles in the research databases and existing meta-analyses (Fuchs & Fuchs, 1986; Gersten et al., 2009; Graham, Hebert, & Harris, 2015; Kingston & Nash, 2011), research on the motivational effects of the formative approach is relatively limited with just a couple of hundred articles (Förster & Souvignier, 2014; Hebbecker & Souvignier, 2018; Hung, Chiu, & Yeh, 2013; Hwang & Chang, 2011; Lipnevich & Smith, 2009; Wongwatkit, Srisawasdi, Hwang, & Panjaburee, 2017; Yin et al., 2008). The aim is to focus on both cognitive and motivational effects of the formative approach in this dissertation. Although it

might help to unify both categories and investigate their reciprocal effects, the primary focus in my dissertation is on conceptualizing formative approach and determining its effects on cognition and motivation individually.

1 The Formative Approach: A Conceptualization and a Model

Formative approach carried out in learning environments is generally regarded as beneficial for learning (Black & Wiliam, 1998). Formative approach is regarded as enhancing students' learning as evidenced by the gains in their achievement scores (Fuchs & Fuchs, 1986). Simultaneously, formative approach is also regarded as enhancing learners' motivation (Kluger & DeNisi, 1996). However, what the formative approach entails and what impact it has on learning and motivation is still unclear. For several decades now there have been continuous attempts and efforts to better understand formative approach and improve it further to enhance instructional practices and thereby the learning outcomes (Black, 2015). Whether we have accomplished a clear understanding of the formative approach still remains questionable as there is no paradigm definition or framework for formative approach. Although the formative approach in learning contexts are generally regarded as enhancing learning, there is inconsistency on how formative approach is conceptualised. Diverse notions and descriptions of formative approach prevail, leading to various definitions. Previous reviews on formative approach have attempted to address these aspects, but a comprehensive review on formative approach is still missing (Bennett, 2011). There is a large body of research on formative approach, but the findings are not consistent. In this dissertation chapter, the aim is to address the gaps in the conceptualization of formative approach.

The goals of this chapter are two-fold. Firstly, conceptualizing formative approach as a comprehensive theoretical model by identifying and including the individual aspects of the approach. Secondly, confirming the proposed model of formative approach. Doing so would enhance the conceptual understanding of formative approach which will in turn have practical implications in learning environments. Prior to conceptualizing a new theoretical model, let us take an overview of the origin of formative approach. In the following section is an outline of the

course of development of formative approach over the years and a synopsis of the existing perceptions of formative approach.

1.1 Theoretical Background - Perspectives on Formative Approach

The word *formative* first appeared in the context of curriculum evaluation in educational research (Scriven, 1967) which eventually transpired to *formative assessment* (Bloom, 1968). Right from the onset, formative assessment was clearly distinguished from its counterpart *summative assessment* (Bloom, Hastings, & Maduas, 1971). Formative assessment means assessment for the purpose of learning that takes place during the learning process and summative assessment is the assessment that takes place after a period of learning in order to measure the effect of learning (Bloom et al., 1971).

Bloom (1968) conceptualized and implemented formative assessment as a diagnostic tool to gauge students' learning and subsequently to ensure students' mastery of a learning unit. Bloom (1968) posited that by conducting formative assessments regularly, students' learning can be monitored and it can be ensured whether a learning unit has been mastered before moving on to the next unit. The diagnostic tool of formative assessment provides information on where the learning gaps are and what the students still need to learn. But the initial phraseology of formative assessment was actually way more than mere assessment. The definitions of Black and Wiliam (2009), Cowie and Bell (1999), Sadler (1989) as well as Tunstall and Gipps (1996) equate to the explanation that formative assessment is the process of knowing how students are learning and responding to the learning process with the aim of enhancing the learning. This led to the attempts of conceptualizing formative assessment based on the process. Cowie and Bell (1999) proposed a model of formative assessment by classifying it as *planned* and *interactive*. While the planned formative assessment entailed the components *eliciting*, *interpreting* and *taking action*, interactive

formative assessment entailed *noticing*, *recognizing* and *responding*. In the planned formative assessment e*liciting* refers to acquiring information on students' learning, understanding and skills, *interpreting* this information and *taking action* based on the obtained information. In the interactive formative assessment *noticing* refers to observing students' learning, understanding and skills, *recognizing* the importance of aspects being noticed and *responding* to students' learning needs and gaps individually or in small groups. Cowie and Bell's (1999) conceptualization as planned and interactive formative assessment came to be called formal and informal formative assessment by Ruiz-Primo and Furtak (2004). Ruiz-Primo and Furtak's (2004) informal formative assessment components of *eliciting*, *recognizing* and *using information* were analogous to the interactive formative assessment components of noticing, recognizing and responding proposed by Cowie and Bell (1999). Although the above definitions, models and conceptualizations shed more light on the formative approach, the conceptualization is still incomplete in the sense, it does not clearly elucidate how the students' learning will be gauged and what it means to respond to the learning process.

The idea of providing feedback to learners based on the results of the formative assessment evolved simultaneously, with the argument that it helps to enhance the learning process (Bloom, 1968). Bloom et al. (1971) who initially distinguished formative and summative assessment, emphasized on the feedback that stems from formative assessment followed by Sadler (1989) and the Assessment Reform Group (Broadfoot et al., 2002) with a similar ideology. So far, there are two perspectives on formative assessment, namely, formative assessment as a diagnostic tool and formative assessment as a feedback tool. Another competing perspective is that of formative assessment being an adaptive tool to cater to students' learning needs (Black & Wiliam, 1998). To facilitate and enhance learning, not only assessment and feedback play crucial roles, but also the

adaptations made to the prevailing instructional practices and the associated decisions. The importance of assessment outcomes aiding in adapting instructions has been time and again reiterated (Fuchs & Fuchs, 1986; Weston, McAlpine, & Bordonaro, 1995). The instructional component has often been included in the definitions and conceptualizations of formative assessment (Black & Wiliam, 2009). Over time, understanding of the formative assessment kept evolving leading to more perspectives on the phenomenon rather than reaching a unifying perspective.

Currently prevailing perceptions of formative approach as an instrument or as a process lead to the risk of oversimplifying formative approach (Bennett, 2011). Such perspectives without a closer examination of the various components of the formative approach increase ambiguity in the conceptualization of formative approach. We have multiple perspectives on the formative approach, but we do not have a comprehensive conceptualization or a definition of it. A major constraint is the challenge in discerning the various components of the formative approach amidst their complexity and interplay. Therefore it is important to identify the various components of the formative approach and understand the interplay of the components. Thereby our primary goal was to arrive at a holistic conceptualization of the formative approach.

In this theoretical section on formative approach, we illustrated the various existing conceptualizations on formative approach. Despite several conceptualizations and models on formative approach, the empirical findings regarding the effects of formative approach are inconsistent. This inconsistency is possibly because of the gaps still prevailing in the understanding of formative approach. It is possible to address this inconsistency by conceptualizing a comprehensive theoretical model of formative approach. The new model will fill in the research gap by ensuring the inclusion of all aspects of formative approach.

By starting out with an extensive literature review on formative approach practices and a conceptual analysis of the various constructs discussed in the literature, we arrived at a model of the formative approach. In the following section, we take a closer look at what all formative approach comprises of. From the prevailing concepts three components of the formative approach were identified. Each of them is defined individually in order to understand the components precisely and avoid any ambiguity. By doing so, we aim to conceptualize a holistic model of formative approach that unifies all facets of the approach.

1.2 The Formative Approach and its Components

So far we looked at the existing conceptions of formative approach. In this section, we present the components identified from the literature review and the conceptual analysis. We also elaborate on the individual components of the formative approach and explain how the components operate as a comprehensive model.

1.2.1 Diagnostic Assessment

Assessments are typically conducted to assess learning. But assessments can also be employed to improve learning when used diagnostically and formatively (Huhta, 2008). Initially, formative assessment and diagnostic assessment emerged for different reasons. The idea of formative assessment of learners evolved from the formative evaluation of curriculum (Bloom, 1968; Scriven, 1967). Diagnostic assessment originally emerged for identifying learning difficulties (Delandshere, 1990). Formative assessment also referred to as assessment for learning, typically takes place during the learning process in order to enhance learning. Diagnostic assessment identifies students with learning difficulties by profiling their strengths and weaknesses and assessing their attainment of learning objectives (Delandshere, 1990).

Among the multiple purposes of assessment in a learning context, one purpose is to determine the learners' strengths and areas to improve. Both formative as well as diagnostic assessment serve this purpose. Though diagnostic and formative assessments emerged for different purposes they have been advancing in the same direction with similar goals and practices (Huhta, 2008) to the extent that the two terms are sometimes used synonymously (Black & Wiliam, 1998; Bloom et al., 1971).

Though most of the assessment instruments provide some diagnostic information, not all assessments are conducted for diagnostic purposes. Assessments within the realm of formative practices are conducted with the purpose of diagnosing learners' skills and knowledge. Without the diagnostic purpose, an assessment has no formative role. Therefore, we call the assessment component in our formative approach model *diagnostic assessment*. We define our diagnostic assessment component as, *the assessment conducted to generate information on the learning process and status.*

1.2.2 Comprehensive Feedback

The over-arching work on feedback is the model of feedback proposed by Hattie and Timperley (2007). Hattie and Timperley (2007) define feedback as the "information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one's performance or understanding" (p. 81). The feedback could contain one or more of the following types of information, such as, information on correctness of the task (right/wrong - knowledge of correctness), information on what is the correct response (knowledge of correct response), and elaborate information on the task (Shute, 2008b). Feedback can be provided in various forms at various timing with varying amount of information (Shute, 2008b).

Hattie and Timperley (2007) claim that the goal of feedback is to reduce the discrepancy between current knowledge/ performance and a learning goal. This applies to the overall formative approach as its main aim is to reduce the discrepancy between the current and the desired state. Feedback accomplishes to reduce the discrepancy by providing information on the task, information on the processes underlying the task, information on the learners' self-regulation and information on the learner itself (Hattie & Timperley, 2007).

Shute (2008b) defined formative feedback as "information communicated to the learner that is intended to modify his or her thinking or behaviour to improve learning" (p.153). Shute (2008b) elaborated on the characteristics, types and timing of formative feedback and also provided guidelines on generating formative feedback. The main features of formative feedback include verification of the correctness of response and elaboration on the response (Kulhavy & Stock, 1989; Shute, 2008a); this elaboration was eventually classified as directive and facilitative feedback (Black & Wiliam, 1998). Furthermore, formative feedback is adapted based on the timing and characteristics of learners (Shute, 2008a). Shute (2008b) highlighted the learner end of formative feedback, but (Gipps & Lawton, 1990) pointed the effect of formative feedback on both teaching and learning. So, not only the learners but also the teachers gain insights from formative feedback practice (Brookhart, Moss, & Long, 2009).

Given the wide ranging types and practices of feedback, it is important to establish what kind of feedback the formative approach comprises of. We cannot generalize that all kinds of feedback serve formative purposes merely because feedback is beneficial. Similar to how any assessment is not formative unless it is intended and structured to be formative (Huhta, 2008), any feedback does not become formative unless intended and planned to be so (Shute, 2008b). For example, feedback on performance or the end product is not formative as there is no scope for change. Therefore, our

focus is particularly on feedback that plays a formative role. When the feedback is on the process, during the process and directed towards improving the process, there is scope for change and feedback plays a formative role here. We refer to the feedback component of our formative approach model as *comprehensive feedback* because it provides the learners with information on their learning processes, learning strategies, understanding as well as their performance rather than merely their performance/scores.

We define our comprehensive feedback component as *the information generated from the diagnostic assessment and provided to the learners on their learning process and status.*

But soon after the advent of formative assessment, any information acquired from formative assessment was viewed as a valuable feedback (Bloom, 1968, 1971). What we failed to realize is that the feedback acquired becomes truly valuable only when we act upon it. By assuming that assessments and feedback in general are formative, we are at the risk of oversimplifying the formative approach. In addition to diagnostic assessment and comprehensive feedback, formative approach comprises yet another component called the adaptive instruction which is illustrated in the following section.

1.2.3 Adaptive Instruction

Adaptive Instruction was first used in the context of enabling slow learners to improve their learning (McNeely & Cummings, 1953). Since then adaptive instruction has been implemented and tested in educational contexts. Glaser and Nitko (1970) defined adaptive instruction as the instruction designed to ensure a particular level of mastery by most students. It involves the decision of what kind of instruction a learner requires for learning. The instruction is to be based on the needs of the learner. The underlying premise being, not all individuals will learn equally well with a single form of instruction (Jonassen & Grabowski, 2012).

In a study by Flanagan (1969), although the terminology adaptive instruction was not used, they had implemented adaptive instruction by identifying the needs of the learners and then developing procedures for the learners to acquire the skills and knowledge in their project PLAN (Program for Learning in Accordance with Needs). Similar approach of adapting instruction to learners' needs has been prevailing in the formative approach. Bloom (1968) elucidated the adaptive instruction aspect of the formative approach, although he did not explicitly use the term. Bloom (1968) emphasized how important it is for the instructors to provide alternate explanations or instructions when the learners don't understand the content. And it is by means of formative assessments instructors become aware of the learning gaps of their students.

The instruction component is discussed as an outcome of formative assessment (Bloom et al., 1971; Sadler, 1989) or as formative assessment facilitating the adaptation of instructional practices (Shepard, 2000, 2009). Though these arguments hold some validity, in reality it is challenging to tease apart instruction from assessment and feedback. This is especially the case during informal formative approach in learning scenarios. The indiscernibility of instruction from assessment and feedback components stems from the fact that the latter two would not count as formative if there is no possibility for an impact on instructional decisions and practices (Huhta, 2008). However, there is a need to distinguish the components from each other and define them precisely (Bennett, 2011). Only then we can clearly conceptualize the formative approach precisely and implement it properly. This in turn will enable us to test the effectiveness of the approach reliably. As much as the three components function synergistically, they are likely to produce effects individually as evidenced by studies that have investigated individual components (Azzi, Ramnanan, Smith, Dionne, & Jalali, 2015; Herman, Osmundson, Dai, Ringstaff, & Timms,

2015; Kickmeier-Rust, Hillemann, & Albert, 2014). It is important to investigate and identify these individual effects for practical implications.

We refer to the instructional component of our formative approach model as *adaptive instruction* because it takes into consideration the results of the diagnostic assessment to adapt the instructions and alter the decisions regarding instructional practices.

We define our adaptive instruction component as *the adaptations carried out in the instructional information, decisions and practices based on the assessment outcomes.*

We have identified and defined three components of formative approach, which are, diagnostic assessment, comprehensive feedback and adaptive instruction. Based on the identified components we provide a definition of formative approach in the following section and propose a model of formative approach.

1.3 Definition of Formative Approach

In order to teach students, it is important to be aware of what the students know and what they need to learn, what skills they have already acquired and what needs to be imparted further (Crooks, 1988; Natriello, 1987). This calls for assessing students and gathering information on a continuous basis. Apart from using this information for diagnostic purposes, students also receive information on their learning progress and inputs on what they need to do as learners to enhance their learning (Black & Wiliam, 2009). This process of providing learners with relevant information on their learning is called feedback (Shute, 2008b). Based on the assessment outcomes, it might be desirable or even essential to adapt the instructional inputs and practices (Black & Wiliam, 2009). The three practices, namely, diagnostic assessment, comprehensive feedback and adaptive instruction collectively constitute the formative approach.

Thereby, we define formative approach as characterizing practices of diagnosing students' learning progress, providing them with comprehensive feedback and adapting instructions based on the assessment outcomes.

1.4 Proposed Model of the Formative Approach

Figure 1 is a visual representation of the model of formative approach. In this visual representation, the aim is to illustrate that the formative approach comprises the three components, namely diagnostic assessment, comprehensive feedback and adaptive instruction. As a further step, the aim is to confirm the model.

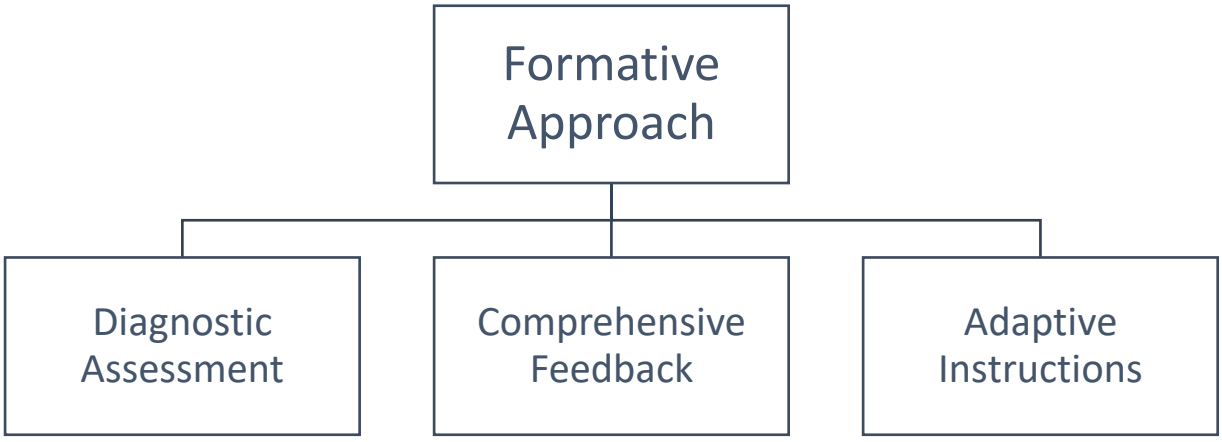

Figure 1. The Formative Approach Model

1.5 A Model of the Formative Approach

After a thorough review of the existing literature on formative approach, we identified the inconsistencies and gaps in the conceptualization. With the aim to address this problem, we proposed a model of formative approach by examining the existing literature. As a next step, we would like to find evidence for our proposed model. To do so, one can conduct a new study or investigate the existing literature for evidence. Before conducting a new study, a systematic review on the formative approach would be a valuable initial step, as the review will provide an overview

of the existing evidence and clearer directions for further research. Conducting a new study based on the solid evidence from a systematic review will be a more efficient approach.

2 The Formative Approach: A Systematic Review and a Meta-Analysis

2.1 Systematic Review

Petticrew and Roberts (2006) defined systematic review as, "A review that strives to comprehensively identify, appraise, and synthesize all the relevant studies on a given topic"(p. 19). A systematic review as the name suggests is a systematic step-by-step method, starting from defining the research question, determining the eligibility criteria of the studies required for answering the question, searching the literature, screening the studies based on the criteria, appraising the included studies, synthesizing the studies and reporting the findings (Petticrew & Roberts, 2006). In the following sections, the methods and results of the systematic review are described and presented in detail.

2.1.1 Eligibility Criteria

Since there is a vast amount of literature on formative approach, we adopted more specific inclusion criteria for our systematic review.

We applied the following criteria:

i. Measures of formative approach: Studies with formative approach by teachers and teacher designed online/ computer-based / paper-based formative approach were included.

   Exclusion criteria: Studies on formative approach by peers and self, such as peer-assessment, peer-feedback and self-assessment; studies on students' experiences, preferences, attitudes and perceptions; studies on teacher intentions, perceptions, attitudes and practices; formative evaluation of programs; formative approach at work place; professional development programs on formative approach; theories, models, comments and reviews on formative approach.

ii. Outcome measures: Studies on learning outcomes and/or motivation as a result of formative approach were included.
iii. Quantitative data: Only empirical studies with quantitative data were included in the coding.
iv. Study Design: Only experimental and quasi-experimental studies were included.
v. Participants: Study participants could be of any age, educational level, or learner status (regular/special education needs).

2.1.2 Literature Search

We conducted our literature search using the databases Academic Search Premier, PsycARTICLES, PsycINFO, PSYNDEX, Teacher Reference Centre and ProQuest. Key words used were formative, formative assessment, formative evaluation, formative practice, formative feedback, formative instruction, assessment for learning, learning progress assessment and progress measurement. No language restrictions were applied.

Possible chances of publication bias were minimized by the following ways: a) by checking the reference sections of studies for additional sources of information such as unpublished reports, conference proceedings, b) by checking with researchers in the field of education and educational psychology for unpublished data, c) by sending out emails to the mailing lists of German Psychological Society (DGPs) and Society for Empirical Educational Research (GEBF) which include members working on research in education and educational psychology and requesting for unpublished data d) by posting the requisition for unpublished data in the Educational Psychology division of American Psychological Association (APA) and in the Learning and Instruction Division of American Educational Research Association (AERA). Using strategy (a) six additional records were included. Through strategy (c) nine records (seven journal articles and two

dissertations) were received out of which the journal articles were already obtained through the database search and the two dissertations did not fulfil the criteria and were not included. Other strategies did not yield additional records or unpublished data. In total, we identified 4,791 empirical and quantitative records till February 2019.

2.1.3 Study Selection

We screened the titles and abstracts of the obtained 4,791 empirical and quantitative records and excluded 4,451 records that did not meet the eligibility criteria. After examining the full text of the eligible 340 articles more closely, we finally included the studies that fulfilled the eligibility criteria. The final sample of 117 records with k = 459 effect sizes were coded on the formative approach components. Figure 2 illustrates the step by step process of the study selection.

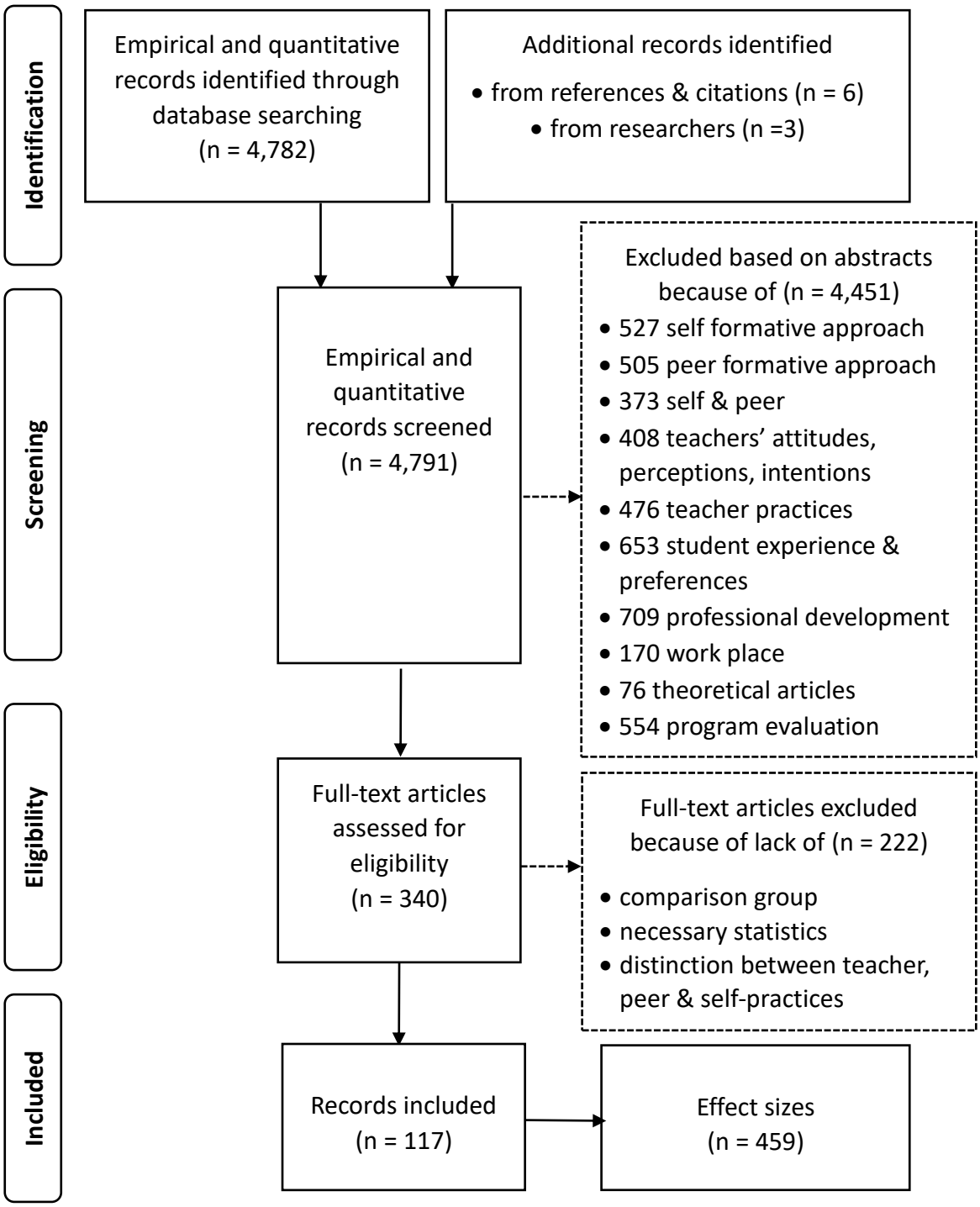

Figure 2. PRISMA Flow Diagram of the literature search

2.1.4 Coding of Primary Studies

Two independent researchers coded the 117 primary studies with 459 effect sizes. Owing to different experimental conditions within a study, there were multiple effect sizes within a study. We coded each effect size individually for the formative approach components. We coded for each study and every effect size, how the formative approach was conceptualized: whether as diagnostic assessment, as comprehensive feedback, or as adaptive instruction. During the coding process it became evident that formative approach was conceptualized as either solely diagnostic assessment or diagnostic assessment with one other component or both components. So, the diagnostic assessment component served as the baseline component. Either comprehensive feedback or adaptive instruction component accompanied diagnostic assessment or all three components were present. Thereby existed four possible constellations of the formative approach components which are illustrated in Table 1.

Table 1

Constellations of the Formative Approach Components

Diagnostic Assessment	Diagnostic Assessment + Comprehensive Feedback	Diagnostic Assessment + Adaptive Instruction	Diagnostic Assessment + Comprehensive Feedback + Adaptive Instruction

Both researchers coded the formative approach components independently. For the inter-rater reliability we calculated Cohen's Kappa using the package *irr* in R-Software (Gamer, Lemon, Robinson, & Kendall's, 2012). The obtained Cohen's Kappa of $\kappa = 0.92$ indicated a high objectivity and reliability of the coding procedure.

2.1.5 Results

In this chapter, firstly, we aimed to conceptualize formative approach and propose a comprehensive theoretical model of the formative approach. Secondly, we conducted a systematic review of the literature on formative approach to find evidence for our proposed theoretical model of the formative approach.

We saw that the literature on formative approach is diverse, complex and unclear. It became evident that the lack of clarity is due to the lack of clear conceptualization and the complexity is due to multiple models, methods and diverse findings on formative approach. Some studies conceptualized formative approach as assessment (Förster & Souvignier, 2014; Herman et al., 2015; Kibble et al., 2011) and some others attached more value to the feedback aspect (McNulty, Espiritu, Hoyt, Ensminger, & Chandrasekhar, 2015; Ozogul & Sullivan, 2009; Shih, Ku, & Hung, 2013). Apart from assessment and feedback, some studies attributed utmost importance to adaptations in instructional practices based on the assessment outcomes (Griffin & Murtagh, 2015; Roschelle et al., 2010; van den Berg, Harskamp, & Suhre, 2016). Given the ambiguity, we proposed a comprehensive theoretical model of the formative approach. Our model of formative approach comprises three main components, namely, diagnostic assessment, comprehensive feedback and adaptive instructions (as illustrated in Figure 1).

Through a systematic search procedure, we obtained all the empirical records (117 records) on formative approach which included journal articles, dissertations, research reports and unpublished research. As a next step we conducted a systematic review of the obtained records on formative approach. All the records were screened and coded by two independent raters. In each record the researchers coded the formative approach components that prevailed.

From the process of coding, it became evident that the components exist in four different constellations and the formative approach in primary studies is conceptualized as one of the four constellations. The diagnostic assessment served as the base and always preceded the other two components. The comprehensive feedback and adaptive instruction components were based on the outcomes of diagnostic assessment and were not observed independently. When we examined the distribution of the conceptualization of the formative approach with one or more of the components from our model, in 13% of the records the formative approach was conceptualized as diagnostic assessment. In 33% of the records formative approach was conceptualized as diagnostic assessment followed by comprehensive feedback. In 23% of the records, formative approach was conceptualized as diagnostic assessment followed by adaptive instruction. And 31% of the records conceptualized formative approach as diagnostic assessment followed by comprehensive feedback and adaptive instruction. The inter-rater agreement was $\kappa = 0.92$ for the coding of formative approach components as well as their constellations in the primary studies.

2.1.6 Conclusions

In this chapter we postulated a theoretical model of the formative approach. A systematic review of the literature on formative approach was conducted and 117 records were obtained. Two independent researchers coded all the primary studies. The high inter-rater agreement for formative approach components and their constellations provided reliable and objective evidence for the theoretical model and its components. Once again, we reiterate our findings here. We defined formative approach as characterizing practices of diagnosing students' learning progress, providing them with comprehensive feedback and adapting instructions based on the assessment outcomes. The model of formative approach comprises the following three components: diagnostic assessment, comprehensive feedback and adaptive instructions. The model distinguishes the three

components clearly and avoids room for ambiguity. It became evident that the components prevail in different constellations with the diagnostic assessment functioning as the basis component.

2.1.7 Limitations

What is not included in our model are the persons involved in implementing and receiving the formative approach. Also the outcomes of the formative approach are not included in the model. Although we arrived at our formative approach model after a thorough review and analysis, the possibility of additional formative approach components cannot be entirely excluded. One way of addressing this is to be open to additions and adaptations in the model.

2.1.8 Next Steps

Having conceptualized the model of formative approach, our next step is to investigate the effects of the formative approach and its components. In the literature, it is most often reported on the effects of the formative approach on learning achievement followed by relatively fewer reports on the effects on motivational aspects. Hence we find it worthwhile to investigate the effects of formative approach on learning and motivation. Additionally, it is essential to identify aspects that could moderate the effect of formative approach on learning and motivation and analyze them. An experimental study on the formative approach will shed light on its effects on learning and motivation. However, given the vast amount of literature on formative approach, a meta-analysis on the existing literature would be economical and eloquent as a subsequent step. Therefore, we decided on conducting a meta-analysis prior to a new empirical study.

2.2 A Meta-Analysis on the Formative Approach

In the previous section we conceptualized the formative approach. We defined the formative approach as characterizing practices of diagnosing students' learning progress, providing them with comprehensive feedback and adapting instructions based on the assessment outcomes. Thereby, the formative approach comprises the three components diagnostic assessment, comprehensive feedback and adaptive instructions. A comprehensive theoretical model of the formative approach was developed. A logical next step is to investigate the effects of the formative approach and its components. The existing empirical literature on formative approach is quite extensive (Wiliam, 2011). Therefore, a meta-analysis on the existing empirical data is consequential. While a vast amount of literature reports positive effects on learners' achievement, relatively smaller fraction of the literature reports positive effects on learners' motivation. It is therefore reasonable to focus on the effects of the formative approach on motivation and learning. Apart from investigating the overall effect of the formative approach on learning and motivation, it would be useful to investigate if one or more of the formative approach components moderate the effects on learning and motivation. Thereby, we can identify if the components have different effects. The knowledge of differential effects will further substantiate the theoretical model of formative approach and will facilitate better implementation of the approach in learning environments. Identification of any additional moderators and analysing their effects would be meaningful for theoretical and practical implications.

In the following section we attempt to draw links between the theories and findings on motivation and learning and the formative approach components of our theoretical model.

2.2.1 Formative Approach and Motivation

Motivation seems to be an important predictor of school achievement (Steinmayr & Spinath, 2009). Achievement motivation could arise out of one's expectation to succeed in a task and how one values the task as illustrated by the expectancy-value theory (Wigfield & Eccles, 2000). According to the expectancy-value-theory, the following processes underlie the expectation of success and subjective task value: a) when individuals perceive their abilities to be matching the demands of the task, they expect to succeed in the task; b) a task is valued when it is perceived meaningful and useful. During diagnostic assessments students get an impression of whether their abilities and the task match. Their active participation in the task is determined by whether they perceive the task as useful. Hattie and Timperley (2007) reported an effect size of $d = 0.55$ when learners received feedback on their current performance in relation to their performance in previous trials. Students benefit in terms of both motivation as well as cognition from the awareness of discrepancies between what they do and what they are actually capable of (Salomon & Globerson, 1987). Therefore, when students are provided with comprehensive feedback based on the diagnostic assessment outcomes, they will be better able to understand if the task and their abilities match. Their initial impression on their ability and the task demands gets confirmed or revised during the diagnostic assessment. With adaptive instructions, the tasks are more likely to match students' perceived abilities and thereby increase their expectations of success. Comprehensive feedback and adaptive instructions could influence students' perception of the utility value of the task. There seems a plausibility for the formative approach components to facilitate expectations of success and alter one's perceptions of a task and the value attributed to the task, and thereby impact students' achievement motivation.

According to the self-determination theory, individuals are intrinsically motivated when their need for autonomy, competence and relatedness are fulfilled (Ryan & Deci, 2000). Interventions in learning environments that support and facilitate "innate psychological needs" (Ryan & Deci, 2000, p. 68) have been found to increase individuals' motivation (Lazowski & Hulleman, 2016). When students receive comprehensive feedback on their learning and performance, it is possible for students to feel more in control of their learning and their efforts and thereby feel competent in accomplishing their learning goals. Instructions adapted to suit their learning needs could facilitate not only stronger feelings of relatedness but also a sense of autonomy (Wehmeyer, 1999).

Deriving from the above two theories it would be worthwhile to investigate whether formative approach in classroom enhance students' motivation. In the following section we will take a look at why it is important to investigate the effects of the formative approach on learning and the existing findings on the effects on learning and motivation.

2.2.2 Formative Approach and Learning

In the current study the term *learning* comprises the learning gains of students across all domains and across the different stages of schooling including tertiary education. The learning gains in this study are operationalized as scores from all forms of achievement tests, performance assessments and standardized measures as these scores are a valid indicator of students' learning gains. It is crucial to investigate whether formative approach leads to better learning gains as this knowledge would facilitate better decisions and more effective and efficient implementation of formative approach in classrooms and other learning environments. Before conducting the planned meta-analysis on the effects of the formative approach on learning and motivation, it would be advantageous to have an overview on the existing meta-analyses and their findings. This will help

enhance the current meta-analysis and preclude flaws and weaknesses. The overview of the existing meta-analyses follows in the next section.

2.2.3 Earlier Meta-analyses on Formative Approach

Given the large interest of researchers on the effects of formative approach on learning, a large number of studies exist on formative approach. The first meta-analysis on formative approach dates back to more than three decades – 1986. Prior to conducting a new meta-analysis, it would be beneficial to have a synopsis of the existing meta-analyses on the formative approach. In this section we provide an overview of the existing meta-analyses on formative approach. Each meta-analysis had specific goals and therefore a specific procedure. This overview of existing meta-analyses shows the strengths as well as the research gaps, based on which we will illustrate how the current meta-analysis contributes further in understanding and implementing formative approach.

Fuchs and Fuchs (1986): This was the first ever meta-analysis on formative approach. The purpose of this meta-analysis was to determine the effect of formative approach on academic achievement. The authors defined their formative approach as "systematic formative evaluation, which includes ongoing evaluation and modification of educational programs" (p. 200). Though the evaluation part of the definition was clear, the main drawback of this meta-analysis was the ambiguity with regard to the latter half of their definition. It was unclear what the modification of educational programs actually entailed. The meta-analytic sample comprised of 21 studies from 23 records (eight journal articles, four dissertations and eleven unpublished research studies) with 96 effect sizes which included samples from kindergarten to twelfth grade. The meta-analysis resulted in a significant average weighted effect size of .70 (k=96). The meta-analysis included the moderating effects (1) behaviour modification of students as part of the formative evaluation

treatment, (2) graphical display of students' performance data by the teachers to the students, (3) whether programmatic changes were made by the teachers based on the data, (4) grade level, (5) special education needs, (6) measurement frequency and (7) treatment duration. The positive aspect of this meta-analysis is that the moderators helped in understanding formative approach better. Regarding the moderator (3), namely, the changes made to the programme by the teachers based on the data, the results of the meta-analysis revealed that changes based on explicit and systematic rules of when and how they are to be carried out, impacted learning largely (d = .91, k = 49, p < .001), whereas freely decided flexible practices impacted learning moderately (d = .42, k = 47, p < .001). However, what kind of programmatic changes were carried out in the primary studies or what the explicit and systematic rules were, were not mentioned in the meta-analysis. It is therefore unclear what the authors attribute these effects to. A weakness of this meta-analysis is that it was not considered whether additional aspects apart from formative evaluation and modification of programs were implemented in the included studies, which could have impacted the effect sizes.

Gersten et al. (2009): This meta-analysis determined the effects of different instructional approaches on mathematics performance of students with learning disabilities. Among the four categories of instructional components analysed in this meta-analysis, two of the categories were relevant for the current study on formative approach. One category was "providing teachers with ongoing formative assessment data on students' mathematics performance and feedback for addressing instructional needs" (p. 1120). The other relevant category was "providing formative assessment data and feedback for students with learning disabilities on their mathematics performance" (p. 1221). Providing teachers with bimonthly formative assessment data of students had a small but significant effect on students' performance g = .21, k = 5, p = .04. In another sample

of studies, when teachers were provided with inputs on planning and fine-tuning their instruction to meet the students' needs based on the formative assessment data, there was no significant effect on students' performance $g = .34$, $k = 3$, $p = .10$. It should be noted that the number of effect sizes in the above two analyses was very small and the resulting effect sizes should therefore be interpreted with caution. However, when we try to explain these findings, there are two possibilities of what might have happened when teachers were provided only with the formative assessment data without any additional inputs. One possibility is that teachers drew their own inferences from the formative assessment data and did what they felt was needed. Or the teachers did not act upon the formative assessment data that was provided to them and the students made learning progress as a result of being assessed and becoming aware of their learning gaps. However, the small but significant effect size implies that teachers could have acted upon the formative assessment data in ways that were not reported or documented.

With regard to the second category, when provided with feedback on their mathematics performance, students with learning disabilities improved significantly in their scores with an effect size of $g = .23$, $k = 7$, $p = .01$. In six out of the seven effect sizes here, feedback was only on performance and the effect sizes ranged between -0.17 and 0.24. In one study in which the feedback was on students' effort, the effect size was 0.60. Possibly this single large effect size elevated the overall effect size for feedback. Also, the sources of feedback were diverse (adults, peers and software programs). Therefore, the results should be interpreted with caution. Based on the formative assessment data, when students were provided with feedback on their progress towards their goals (newly set learning goals or pre-set ones), there was no significant effect on learning $g = .17$, $k = 5$, $p = .29$.

The major limitation of this meta-analysis was that there was no operational definition of formative assessment. It can only be inferred from the article that they refer to some sort of assessment data when they mention formative assessment data. Though the meta-analysis provided some insights, we should bear in mind the small number of records for each effect size and the lack of clarity on what effects the authors are actually discussing and what the effect sizes imply for practice. Moreover, the effect size for feedback on learning is questionable as one particular effect size with a very specific operationalisation of feedback is elevating the overall effect size.

Kingston and Nash (2011): This meta-analysis set out to determine the efficacy of formative assessment and arrived at a weighted mean effect size of .20 and a median effect size of .25. The meta-analytic sample comprised of 13 records (five journal articles, five dissertations, three conference presentations) with 42 effect sizes including samples from kindergarten to twelfth grade. This meta-analysis was criticised by Briggs, Ruiz-Primo, Furtak, Shepard, and Yin (2012) for its study retrieval procedures, application of inclusion criteria, effect size calculations and the impact of outcome variable on effect size variability. Though Kingston and Nash (2011) provided their justifications for the above criticisms in a subsequent article, our concerns are with regard to the conceptualization of formative assessment. Though the article addressed the fact that each included study had one or more components of the formative approach, these components were neither considered nor included in the analysis. So, the effect sizes in this meta-analysis indicate the impact of a collection of formative approach practices on students' learning. McMillan, Venable, and Varier (2013) attempted to augment Kingston and Nash's (2011) meta-analysis and pointed out the different conceptualizations of formative approach in the individual studies and developed guidelines to code the formative approach components. However, no further analyses were conducted based on the coding by McMillan et al. (2013) or by the original authors (Kingston

& Nash, 2011) on the effects of the formative approach components on students' achievement. The criticisms of McMillan et al. (2013) on the poor quality of the studies in the meta-analysis render almost all the included studies unfit. This is rather a reflection of the gaps and inconsistencies in the conceptualization of formative approach resulting in diverse practices and outcomes rather than the weakness of the meta-analysis.

Graham et al. (2015): This meta-analysis set out to find if feedback on formative assessments of students' writing enhanced their writing performance. The meta-analytic sample for feedback effects comprised of 25 records (seventeen journal articles and eight dissertations) with 29 effect sizes with samples from grades one to eight. They found an effect size of $g = 0.87$, $k = 7$, $p < .001$ for feedback from adults, an effect size of $g = 0.38$, $k = 4$, $p = .001$ for feedback from computers and $g = 0.58$, $k = 8$, $p < .001$ and $g = 0.62$, $k = 10$, $p < .001$ respectively for feedback from peers and self. Though this meta-analysis reports the effects of feedback, the operationalisation of feedback also included instructional inputs. The potential effect of the initial formative assessment on students' performance should not be overlooked either. The reported effect sizes are therefore not completely attributable to feedback.

To sum up, earlier attempts to identify the efficacy of formative approach and their components by means of meta-analysis (Kingston & Nash, 2011) were criticised for methodological flaws (Briggs et al., 2012; McMillan et al., 2013). Apart from the methodological issues, there are issues with regard to the conceptualizations of formative approach in the earlier meta-analyses. Formative approach is not conceptualized consistently, to the extent that interventions are carried out with unclear conceptual definitions and vague operationalisations. The reported effects of formative evaluation in Fuchs and Fuchs (1986); formative assessment and feedback in Gersten et al. (2009); formative assessment in Kingston and Nash (2011); and

feedback on formative assessment in Graham et al. (2015) do not specifically pertain to the constructs defined in the respective meta-analysis.

Consequently, to identify the efficacy of formative approach and its specific components, replicating or extending the existing meta-analyses is not optimal. Therefore, a new meta-analysis is necessary to ascertain the effects of the components of formative approach. Ascertaining the effects of formative approach and its components has practical implications in terms of implementing the approach, improving it, and framing policies with regard to the approach.

2.2.4 Moderators of the Formative Approach

Ambiguity regarding the effect of formative approach on learning prevails and is yet to be clarified. Although the formative approach is commonly believed to be enhancing learning, the findings are inconsistent (Bennett, 2011). The inconsistency could probably be due to the role of different moderators. Therefore, apart from the formative approach components, we are interested in potential additional moderators of the effect of formative approach on learning. We selected the additional moderators a priori and categorised them as learner, teacher, and formative approach intervention related moderators. We wanted to be judicious about our choice of moderators as there is a chance of one in ten moderators becoming significant merely by chance (Viechtbauer, 2007).

To be further sure, we ran the multiple moderator model, wherein all the significant individual moderators were included in one model to assess if they still remain significant (Assink & Wibbelink, 2016). By means of the multiple moderator model we can address the issue of multicollinearity and identify the most relevant moderators (Hox, Moerbeek, & van de Schoot, 2017).

Learner related Moderators

Regarding learner related moderators, we plan to analyse the moderating effect of the learners' stage of schooling and their special education status. When formative approach is implemented in classrooms, it is questionable whether learners across different stages of schooling receive, process and benefit from formative approach similarly. Learners in different stages of development could perceive the demands of learning tasks differently and attribute values to the tasks differently, thereby experiencing different levels of motivation (Wigfield & Eccles, 2000). According to the information-processing approach, the capacity for children to process information increases gradually with their development, enabling them to acquire more complex knowledge and skills (Siegler, 1998). It is possible that with increase in age learners are better able to process the feedback and instructions they receive. While the elementary school learners are more likely to profit from adapted instructions, secondary and high-school learners are likely to benefit more from comprehensive feedback. However, with increase in information processing ability with age it is possible that they are able to generate feedback for themselves from the diagnostic assessment outcomes. This could mean that beyond a certain level of development, learners do not profit much from formative approach. Therefore, we are interested if different stages of learners are benefitting differently from the formative approach in terms of learning gains and their motivation.

The Education of Persons with Special Education Needs (EPSEN) Act defines "Special education needs are a restriction in the capacity of the person to participate in and benefit from education on account of an enduring physical, sensory, mental health or learning disability, or any other condition which results in a person learning differently from a person without that condition" (Department of Education and Science, 2004). It is possible that learners with special education needs profit largely, show significant learning gains, experience higher motivation when

instructions are adapted to their needs. On the other hand, the possibility of them getting overwhelmed with comprehensive feedback could curtail their learning gains or demotivate them. Therefore, if learners with special education needs tend to have different emotional needs, process information and acquire knowledge and skills differently from those of regular education needs, we are interested in finding if the effect of formative approach is different for regular and special education needs learners on their learning and motivation. With regard to the learners with regular education needs, their emotional needs are different and they are already likely to be making steady learning progress that the gains due to formative approach are not as steep as the learners with special education needs.

Teacher related Moderators

Regarding teacher related moderators, we intend to analyse the moderating effect of training teachers on formative approach and teachers' experience (novice versus experts). Teachers' perceptions and attitudes towards teaching and assessment practices are not always the same and keep varying based on their experiences. Attitudes and perceptions on formative approach of trained teachers were significantly positive than those teachers who were not trained (Young & Jackman, 2014). Trained teachers encouraged students to engage in reflective journal writing significantly more than the untrained teachers (Young & Jackman, 2014). A meta-analysis by Blank and Las Alas (2009) shows evidence for gains in student achievement as a result of teacher professional development. Training teachers seems to impact their perceptions, attitudes, and practices of formative approach. Therefore, we wanted to find out if training teachers on formative approach increases the effect of formative approach on students' learning and motivation.

The Dreyfus Five-Stage Model of Adult Skill Acquisition illustrates the transition of an adult learner from being a novice to becoming an expert (Dreyfus & Dreyfus, 1980). In teaching

practice, the expert-novice difference has been noticed in teachers' planning, interactive teaching and post lesson reflections (Livingston & Borko, 1989). This led to the question of whether teachers' experience, as an indicator of their skill level (novice versus experts), moderates the effect of formative approach on students' learning and motivation. We expect that the expert-teachers are more skilled at formative approach and therefore elicit higher learning gains and have larger impact on students' motivation and learning.

Formative Approach Intervention related Moderators

Regarding the intervention related moderators, we propose to analyse the moderating effect of the following:

(a) formal vs. informal formative approach

(b) subject (math, science, reading and writing)

(c) individual vs. group implementation

(d) mode of formative approach implementation (paper-pencil, face to face, computer based).

Practices of formative approach range from informal to formal (Ruiz-Primo & Furtak, 2006). Formal formative approach practices are structured and based on a framework, whereas informal formative approach provides more flexibility, is conversational, allows spontaneous reactions and modifications (Bell & Cowie, 2001). While the formal formative approach is planned and is therefore more structured, informal formative approach takes place spontaneously and is therefore interactive. In other words, formal formative approach is curriculum based or even embedded in the curriculum and informal formative approach takes place on-the-fly (Shavelson et al., 2008). Owing to the above differences in the practice of formal and informal formative

approach, it is worthwhile to know if and how formal-informal implementation is moderating the effect of formative approach on learning and motivation.

With regard to the subject domains, the acquisition of knowledge and skills are different across domains. The demands of the domains are different and so are the instructional inputs (Tricot & Sweller, 2014). Steinmayr and Spinath (2009) illustrated the differences in motivation within specific subject domains and across domains in general. Therefore, we would like to find if the formative approach effects on learning and motivation are different for different subject domains.

Formative approach seems to be commonly practiced in special education settings for diagnostic purposes and to draw up individualised education plans (Cornelius, 2013). In regular education settings, formative approach is used for individual learners as well as in groups (Ruiz-Primo & Furtak, 2006). Though there are practical reasons for group and individual implementation, it is possible that individual implementation caters more precisely to the learning needs and students' emotional needs as opposed to the group implementation where the formative approach caters to a range of learning and emotional needs. Students who underwent both individual as well as group formative approach have reported different experiences for the two forms of formative practices on their learning and their motivation (Weurlander, Söderberg, Scheja, Hult, & Wernerson, 2012). Therefore, it would be worthwhile to analyse whether and how much of differential effects individual versus group implementation has on learning and motivation.

Whether the formative approach is implemented by the teacher, via computer or as paper-pencil method could moderate the effects of formative approach on learning and motivation as each of the modes has its pros and cons. While teacher implemented formative approach could

increase the sense of affiliation for the learners (Brackett, Reyes, Rivers, Elbertson, & Salovey, 2011), computer based formative approach gives scope for more systematic implementation (Bull & McKenna, 2004). Given the possibility of different modes of formative approach implementation moderating the effects on students' learning and motivation differently, the mode of implementation is included in the moderator analysis.

2.2.5 The Current Meta-Analysis

The main aim of the current meta-analysis is to identify the effects of the formative approach and its components on learning and motivation. Apart from that we aim to find out if moderators related to the learners, teachers, and the formative approach intervention additionally moderate the effect of the formative approach on learning and motivation. For the additional moderator analyses, we will investigate the overall formative approach and not the components, as the number of effect sizes are relatively small for the specific components. In the following section we will go through the details of the methodological approach to this meta-analysis. The literature search and the eligibility criteria were illustrated in the previous chapter on our systematic review. We will now delve into the research questions of the current meta-analysis, study selection, coding of empirical data and the analysis.

2.2.6 Research Questions and Hypotheses

We addressed the following research questions and hypotheses in our meta-analysis:

i. What is the overall effect of formative approach on students' learning and motivation?
ii. What are the effects of the formative approach components on learning and motivation?
iii. Moderator hypotheses:

The various aspects that could moderate the effects of the formative approach on learning and motivation were discussed in the theory section. Following are the hypotheses for the moderators tested in our meta-analysis:

a. Learners' stage of schooling:
    i. Learning: The stage of schooling significantly moderates the effect of the formative approach on learning, with higher levels of schooling having larger learning effects.
    ii. Motivation: The stage of schooling significantly moderates the effect of the formative approach on motivation, with higher levels of schooling having larger effects on motivation.
b. Educational needs of learners (special education Vs regular education):
    i. Learning: Learners' educational needs significantly moderates the effect of the formative approach on learning, with special education learners having larger effect on learning.
    ii. Motivation: Learners' educational needs significantly moderates the effect of the formative approach on motivation, with special education learners having larger effect on motivation.
c. Training teachers on formative approach:
    i. Learning: Training teachers on formative approach significantly moderates the effect of formative approach on learning with teachers with training having larger effect on learning.

ii. Motivation: Training teachers on formative approach significantly moderates the effect of formative approach on motivation with teachers with training having higher effect on learners' motivation.

d. Novice vs. Expert Teachers:

i. Learning: Teachers' expertise significantly moderates the effect of formative approach on learning, with expert-teachers having larger effect on learning.

ii. Motivation: Teachers' expertise significantly moderates the effect of formative approach on learning, with expert-teachers having larger effect on learners' motivation.

e. Formal vs. Informal:

i. Learning: Formal versus informal implementation of the formative approach significantly moderate the effect of formative approach on learning and have differential effects on learning.

ii. Motivation: Formal versus informal implementation of the formative approach significantly moderate the effect of formative approach on motivation and have differential effects on motivation.

f. Subject:

i. Learning: The subject domain significantly moderates the effect of the formative approach on learning and the effects are different for different subject domains.

ii. Motivation: The subject domain significantly moderates the effect of the formative approach on motivation and the effects are different for different subject domains.

g. Group vs. individual:

i. Learning: Group versus individual implementation of the formative approach significantly moderates the effect of formative approach on learning and have differential effects on learning.

ii. Motivation: Group versus individual implementation of the formative approach significantly moderates the effect of formative approach on motivation and have differential effects on motivation.

h. Mode of implementation (in person/ on computer / paper-pencil):

i. Learning: The mode of implementation of formative approach significantly moderates the effect of formative approach on learning with differential effects on the learning.

ii. Motivation: The mode of implementation of formative approach significantly moderates the effect of formative approach on motivation with differential effects on motivation.

2.2.7 Methods

In this section we describe our study selection procedure and coding of primary studies for the meta-analysis. The eligibility criteria and our literature search procedure for the retrieval of studies have been explained in the previous chapter on systematic review. Also, a graphical representation of the process of searching and selecting the studies is in the previous chapter (Figure 2).

Study Selection

We screened the 4,791 quantitative records and excluded 4,451 records that did not meet the eligibility criteria. We screened the titles and abstracts of all the obtained records. After examining the eligible 340 full text articles more closely, we contacted the authors of articles and dissertations that did not report the descriptive data for calculating the effect sizes. During our correspondence with authors for the descriptive data we also inquired for unpublished data on this topic. No unpublished data was obtained during these inquiries. We finally included the studies that fulfilled the eligibility criteria either immediately or after obtaining the necessary data from the authors. The meta-analysis was based on a final sample of 117 records with k = 459 effect sizes. Figure 2 in the previous section illustrates the step by step process of the study selection.

Coding of Primary Studies

We coded the effects of the formative approach on learning and motivation in all of the included studies that resulted from a comparison of the formative approach intervention with no-intervention/ conventional treatment control group or from a pre-post comparison of the intervention group. We coded the raw data (i.e., the means and standard deviations of the learning and motivation measures) and computed a standardized effect size metric for each study. There were three types of data available from the primary studies and the standardized mean difference (SMD) was calculated for each study. The first category included pre-post means and standard deviations of the intervention group from which the SMD was calculated. The second category included post intervention means and standard deviations of the treatment and control group from which the SMD was calculated. The third category included the pre and post intervention means and standard deviations of treatment and control group. In these studies, the pre and post learning measures were either the same or comparable in nature. The SMD in these studies were calculated

from the pre to post gain scores and the pooled pre-test standard deviations. We used pooled pre-test standard deviations because they provide an unbiased estimate of the population effect size as opposed to two other possible methods, namely, using the pooled standard deviations across pre- and post-test scores or taking only the pre-test standard deviations (Morris, 2008).

For each included record, we coded the formative approach components and the additional potential moderators. As there were multiple experimental conditions within a record, we coded each condition separately, resulting in multiple effect sizes from a single record. We coded whether the formative approach in the primary studies was conceptualized as one or more of the three components, diagnostic assessment, comprehensive feedback, and adaptive instruction. We used dummy coding for this. From the systematic review we know that the three components prevail in four constellations, namely, diagnostic assessment (DA); diagnostic assessment and comprehensive feedback (DA+CF); diagnostic assessment and adaptive instructions (DA+AI); diagnostic assessment, comprehensive feedback and adaptive instructions (DA+CF+AI). An effect size was coded as DA, when the assessment generated information on the learning status. An effect size was coded as DA+CF, when assessment was conducted and information on learning was generated and based on this information, feedback was provided to the learners. A record was coded as DA+CF, when the given feedback was on the performance of the task and/or on the processes and strategies involved in solving the task. An effect size was coded as DA+AI, when information on learning was generated and based on this information, instructional information, decisions and practices were adapted. An effect size was coded as DA+CF+AI, when information on learning was generated and based on this information, feedback was provided to the learners and instructional information, decisions and practices were adapted.

Using dummy coding, we coded the four constellations resulting in four dummy variables. Each dummy variable was coded 0 (absent) or 1 (present). DA was coded as 1 = present and 0 = absent. Similarly, DA+CF; DA+AI; DA+CF+AI each was coded as 1 = present and 0 = absent. As additional potential moderators, we coded the learner, teacher, and formative approach intervention characteristics. As learner characteristics we included learners' stage of schooling and special education needs. As teacher characteristics we included training on implementing formative approach and teacher expertise. The formative approach characteristics we coded include formal/ informal implementation, subject (Medicine, Science, Math, Reading, Writing), group/ individual implementation and the mode of implementation (paper-pencil/ computer/ face-to-face).

The raw data for the dependent variables – learners' achievement, and learners' motivation (when available) – were coded and the standardized mean difference (Hedges' g) was calculated for every study. The first author of this meta-analysis coded all the records. A subset of 37% percent of the records was additionally coded by a second coder. We calculated Cohen's Kappa to analyse the agreement between the two coders. There was high agreement between the two coders, ($\kappa = 0.95$), indicating a high objectivity and reliability of the coding procedure.

We calculated Hedges' *g* for the standardized mean difference as it is an unbiased estimator of the mean difference for unequal sample sizes for the experimental and control groups and corrects for the positive bias by using n-1 for calculating the pooled variance. The effect sizes were integrated by a random-effects model as it accounts for variability at study level effect sizes due to the variability in the population effect sizes enabling to generalize the findings (Borenstein, Hedges, Higgins, & Rothstein, 2011).

2.2.8 Analysis

We analysed the data using the *metafor* package (Viechtbauer, 2010) in R software (R Core Team, 2013). We conducted a multivariate meta-analysis to analyse the effects of the formative approach on learning and motivation. All analyses were conducted separately for the outcomes learning and motivation as the pool of primary studies differed for each of the outcomes. Owing to the heterogeneity in the conceptualizations of formative approach between studies, we opted for a random-effects model, as this model accounts for the variance within as well as between studies. We performed a test of homogeneity of effects and calculated the Cochran *Q* statistic (Hedges & Olkin, 2014) to assess heterogeneity between studies. As the *Q* statistic only indicates the presence/absence of heterogeneity and does not quantify the amount of heterogeneity (Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006), we estimated the amount of residual heterogeneity (τ^2) and the ratio of true to total variance (I^2).

As there are multiple effect sizes from the same sample, there is a possibility for dependency of effect sizes. Therefore, we used the restricted maximum likelihood multivariate meta-analysis method to determine the overall effect of formative approach as well as the moderating effects of the formative approach components, learner characteristics, teacher characteristics and formative approach intervention characteristics. The multivariate meta-analysis addresses this dependency (van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2015).

After a visual inspection of the forest plot of the effect sizes, we statistically identified the outliers by calculating the Mahalanobis distance (Acuna & Rodriguez, 2004) and removed them. After a visual inspection of the funnel plot of the effect sizes, we calculated the Eggers' regression value (Egger, Smith, Schneider, & Minder, 1997) to assess the presence of publication bias. By

using the Trim and Fill method (Duval & Tweedie, 2000) we imputed the missing effect sizes due to publication bias. Finally, we performed fail-safe *N* (Orwin, 1983; Rosenthal, 1979) analyses to assess the robustness of the effect of formative approach.

2.2.9 Results

A total of 117 records with 459 effect sizes were included in the meta-analysis. The included records comprised of 68 journal articles and 49 dissertations. Of the included 459 effect sizes, 9% had conceptualized formative approach with only the diagnostic assessment component, 27% conceptualized formative approach with diagnostic assessment and comprehensive feedback components, 24% with diagnostic assessment and adaptive instruction and 40% with all three components of the formative approach.

Publication Bias

In a first step, to visualize the heterogeneity of effect sizes with their confidence intervals and to identify outliers, we created a forest plot (Figure 3). Although most of the effect sizes lined up in a row, there were some outliers in the data.

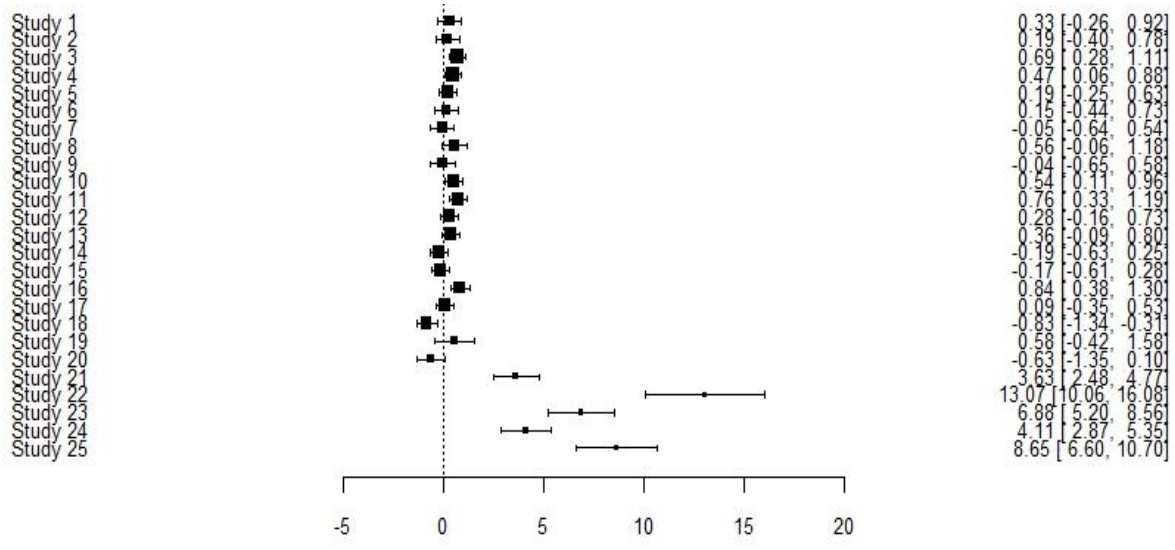

Figure 3. Forest Plot of a Subgroup of Studies

From the total sample of 459 effect sizes, six effect sizes from three records were above the critical Mahalanobis value of $\chi^2(4, 367) = 11.07$. With the outliers the weighted mean effect size was $g = 0.52$ [0.42, 0.63], $k = 459$, $p < .001$. Without the outliers the weighted mean effect size was $g = 0.45$ [0.37, 0.52], $k = 453$, $p < .001$. The outliers were removed, and the final dataset comprised of 115 records with 453 effect sizes.

To check for publication bias (Rothstein, Sutton, & Borenstein, 2006), in a next step, we created a funnel plot with the effect size on the x-axis and standard error on the y-axis. The effect sizes were concentrated at the top of the funnel (Figure 4). But, applying the trim and fill method

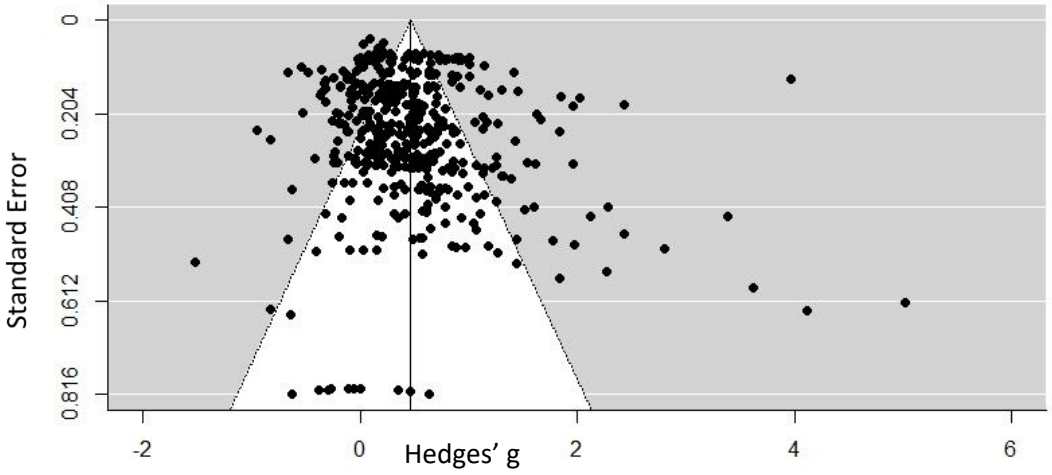

Figure 4. Funnel Plot with Outliers

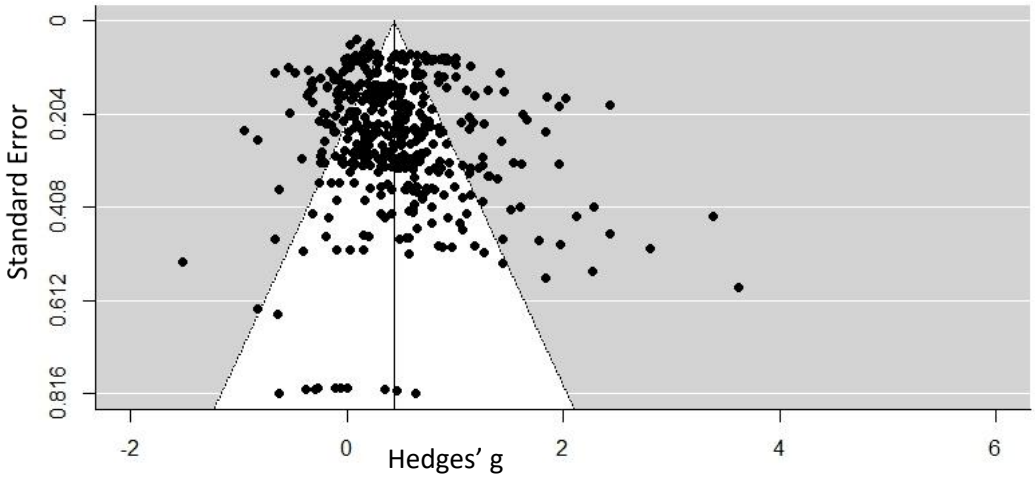

Figure 5. Funnel Plot without Outliers

(Duval & Tweedie, 2000) did not add additional effect sizes to the distribution to make it more symmetric (Figure 5). This indicates that the publication bias is too small to be detected with the trim-and-fill method. As funnel plot asymmetry is not always due to publication bias and has only a limited capacity to detect publication bias, we additionally ran the Egger's regression test (Egger et al., 1997; Rothstein, Sutton, & Borenstein, 2005). Egger's regression test for random effects with standard error as the predictor indicated asymmetry ($Z = 3.45$, $p < .001$), thus indicating some publication bias. The Orwin fail-safe N analysis (Orwin, 1983) showed that 453 studies averaging null results will have to be added to reduce the unweighted average effect size of 0.45 to 0.23. The Rosenthal fail-safe N analysis (Rosenthal, 1979) showed that 360,956 studies averaging null results will have to be added to reduce the observed significance level of $p < .001$ to $p = .05$. The above analyses confirm the robustness of the current meta-analytic findings.

Effect of the Formative Approach on Motivation

To estimate the mean effect of the formative approach on motivation, we ran a multivariate meta-analysis with restricted maximum likelihood. The weighted mean effect size (Hedges' g) of the 53 effect sizes was 0.34. The 95% confidence intervals [0.05, 0.63] did not include zero with $p = .02$. The homogeneity statistic Cochran Q indicated the presence of heterogeneity as the variation in effects is statistically significant ($Q = 622.07$; $k = 53$; $p < .001$). The amount of variance of the true effect sizes as indicated by the τ^2 value is 0.34. The proportion of true to total variance indicated by $I^2 = 93.6\%$ implied a large heterogeneity of the effect sizes which in turn implied that the effect of the formative approach on motivation is moderated by other variables. Therefore, it was essential to analyze whether the formative approach components moderated the effect of formative approach on learning. This analysis is illustrated in the following section.

Effects of the Formative Approach Components on Motivation

When analyzed whether the formative approach components moderated the effect of formative approach on motivation, the effect was close to being significant with $F(4, 49) = 2.42$, $p = .06$. A multivariate meta-analysis with the different constellations of the formative approach components as moderators produced the following results.

The multivariate meta-analytic effect of the diagnostic assessment on motivation was $g = 0.35$ [0.05, 0.66], $k = 10$, $p = .03$; the effect of diagnostic assessment and comprehensive feedback was $g = 0.44$ [0.14, 0.73], $k = 24$, $p = .01$, the effect of diagnostic assessment and adaptive instruction was not significant $g = 0.09$ [-0.41, 0.59], $k = 7$, $p = .71$; and the effect of all three components together was $g = 0.33$ [0.03, 0.63], $k = 12$, $p = .03$. On taking a look at the effects of the formative approach components on motivation and comparing them, they are not significantly different from each other. Even the confidence intervals of all the effect sizes are largely overlapping as illustrated in figure 6.

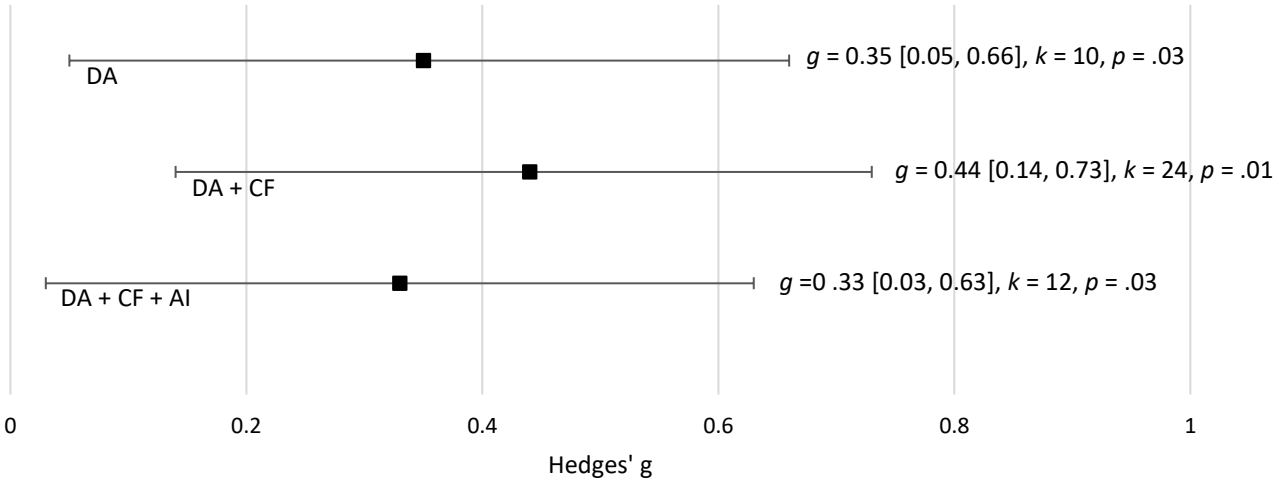

Figure 6. Effect Sizes and Confidence Intervals of Formative Approach Components on Motivation

Effect of the Formative Approach on Learning

To estimate the mean effect of the formative approach on learning, we ran a multivariate meta-analysis with restricted maximum likelihood. The weighted mean effect size (Hedges' g) of the 453 effect sizes was 0.45. The 95% confidence intervals [0.37, 0.52] did not include zero with $p <$ .001. The homogeneity statistic Cochran *Q* indicated the presence of heterogeneity as the variation in effects is statistically significant ($Q = 3279.84$; $k = 453$; $p < .001$). The amount of variance of the true effect sizes as indicated by the τ^2 value is 0.20. The proportion of true to total variance indicated by $I^2 = 88.78\%$ implies a large heterogeneity of the effect sizes which in turn implies that the effect of the formative approach on learning is moderated by other variables.

Effects of the Formative Approach Components on Learning

Among the 453 effect sizes from 115 records, 42 effect sizes are from studies which conceptualized formative approach as only diagnostic assessment component, 124 effect sizes are from studies which conceptualized formative approach as diagnostic assessment and comprehensive feedback components, 106 effect sizes are from studies which conceptualized formative approach as diagnostic assessment and adaptive instruction components and 181 effect sizes are from studies which conceptualized formative approach as all the three components. When analyzed whether the components moderate the effect of formative approach on learning, there was a significant moderating effect $F(4, 449) = 34.15$, $p < .001$. A multivariate meta-analysis with the formative approach components as moderators produced the following results.

The multivariate meta-analytic effect of the diagnostic assessment on learning was $g = 0.37$ [0.18, 0.56], $k = 42$, $p < .001$; the effect of diagnostic assessment and comprehensive feedback was $g = 0.51$ [0.39, 0.64], $k = 124$, $p < .001$, the effect of diagnostic assessment and adaptive instruction

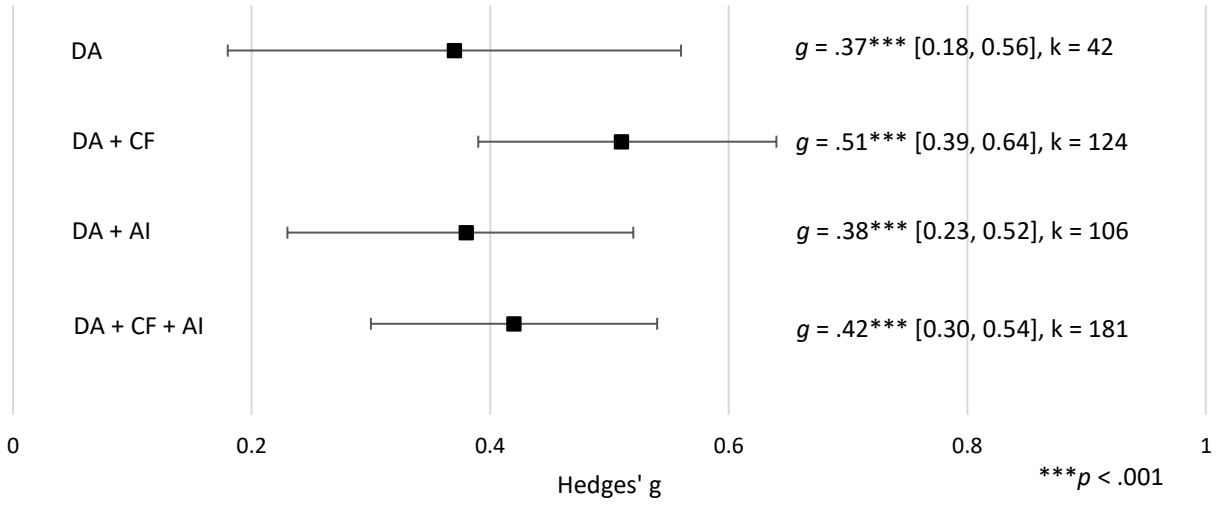

Figure 7. Effect Sizes and Confidence Intervals of Formative Approach Components on Learning

was g = 0.38 [0.23, 0.52], k = 106, p < .001; and the effect of all three components together was g = 0.42 [0.30, 0.54], k = 181, p < .001. On comparing the effect sizes of the formative approach components, it was clear that they are not significantly different from each other. Even the confidence intervals of all the effect sizes are largely overlapping as shown in figure 7.

Moderators

Learner related Moderators – Motivation

The omnibus test indicated that the learners' stage of schooling significantly moderated the effect of formative approach on motivation: F(2, 49) = 3.10, p = .04. Among the 52 effect sizes from 16 records, there were 31 effect sizes for tertiary level, 6 effect sizes for secondary level and 15 effect sizes for elementary level. The learners' motivation was significant at the tertiary level: g = 0.53 [0.11, 0.95], k = 31, p = .01.

With regard to learners' educational needs, there were no primary studies that measured the motivation of special education learners.

Learner related Moderators - Learning

With regard to the learner characteristics, we checked whether the stage of schooling moderated the effect of formative approach on learning. There is a moderating effect of the stage of schooling, as the results of the omnibus test indicated a significant moderating effect: $F(5, 435) = 32.94$, $p < .001$. The effect sizes are listed in Table 2.

Table 2

Moderating Effects of the Stages of Schooling on Learning

	Kindergarten	Elementary	Secondary	High School	Tertiary
Formative Approach	$g = 0.32$ [0.01, 0.63] $k = 55$ $p = .05$	$g = 0.35$ [0.23, 0.48] $k = 112$ $p < .001$	$g = 0.32$ [0.20, 0.44] $k = 113$ $p < .001$	$g = 0.76$ [0.53, 0.99] $k = 43$ $p < .001$	$g = 0.56$ [0.43, 0.69] $k = 117$ $p < .001$

When compared between the different levels of schooling, effect of formative approach at high school level was significantly larger than kindergarten, elementary and secondary levels. The effect of formative approach at tertiary level was significantly larger than elementary and secondary levels but not the high school. The high school and tertiary level learners are close in terms of their age and comprehension level and therefore there was no significant difference between these two levels. But these two levels seem to benefit the most from the formative approach.

The omnibus test indicated that the education status had a significant moderating effect on learning: $F(2, 431) = 64.97$, $p < .001$. The effect sizes are listed in Table 3.

Table 3

Moderating Effects of Special Education and Regular Education on Learning

	Regular Education	Special Education
Formative Approach	g = 0.44 [0.36, 0.51] k = 348 p < .001	g = 0.42 [0.19, 0.65] k = 85 p < .001

Although the learners' education status significantly moderates the effect of formative approach on learning, there is no significant difference between the effect on learning for regular and special education learners. The significant omnibus test indicate that at least one of the two categories (special education or regular education) is significantly different from zero (Assink & Wibbelink, 2016). In this case, both moderators are significantly different from zero. But that does not necessarily mean that the effects of the two categories differ significantly from each other. The results of the Test for Residual Heterogeneity show that there is significant unexplained variance left between all effect sizes in the data set (Q = 3085.368, k = 431, p < .001), after education status has been added to the meta-analytic model.

Teacher related Moderators – Motivation

With regard to teacher related moderators, there were very few records with the necessary data. There were totally nine records with 28 effect sizes. Out of these, there were 17 effect sizes from six records where teachers were trained on formative approach and 11 effect sizes from three records where teachers were not trained, and the formative approach was implemented via computer. When analyzed whether training teachers on formative approach moderated the effect of formative approach on learners' motivation, the omnibus test indicated a significant moderating effect: $F(2, 26) = 4.23$, $p = .03$. But contrary to our hypothesis, no training on formative approach had a significant effect g = 0.92 [0.27, 1.57], k = 11, p = .007 as opposed to training on formative approach g = 0.01 [-0.43, 0.45], k = 17, p = .96. Therefore we reject the hypothesis that teachers

with formative approach training have larger effect on motivation. This result seems puzzling and counterintuitive in the first instance. However, it is possible that the training on formative approach increased the variance in teachers' formative approach implementation. On the contrary, computer-based formative approach implementation without training was pre-programmed and thereby had lesser variance.

When analysed whether teacher expertise moderated learners' motivation, the omnibus test indicated no moderating effect: $F(2, 8) = 0.03$, $p = .97$.

Teacher related Moderators - Learning

With regard to teacher related moderators, we examined the moderating effect of teacher training on formative approach implementation and teacher expertise. The omnibus test indicates that training teachers on formative approach has a significant moderating effect on learning: $F(3, 419) = 49.10$, $p < 0.001$. The significant omnibus test indicates that at least one of the three categories is significantly different from zero (Assink & Wibbelink, 2016). When checked whether the effect on learning differed between trained and untrained teachers, there was no significant difference. When teachers were trained on the formative approach intervention, the effect of formative approach on learning was $g = 0.42$ [0.31, 0.52], $k = 279$, $p < .001$. When teachers were not trained on formative approach, the effect of formative approach on learning was $g = 0.44$ [0.26, 0.61], $k = 69$, $p < .001$. When formative approach was implemented via teacher designed online tools and applications or offline quizzes on mobile devices and computer, formative approach had an effect of $g = 0.53$ [0.37, 0.69], $k = 74$, $p < .001$ on learning.

Teacher expertise significantly moderates the effect of formative approach on learning, $F(2, 78) = 6.21$, $p = 0.003$. Teachers with less than five years of experience were grouped under "novice-teachers" and the ones with more than ten years of experience were grouped under

"expert-teachers." Among the primary studies there was none with teaching experience between five and ten years. The effect of formative approach on learning for novice-teachers was not significant g = 0.20 [-0.07, 0.47], k = 18, p = .14, whereas the effect for expert-teachers was significant g = 0.27 [0.10, 0.43], k = 62, p = .002. With higher teaching expertise, teachers' formative approach implementation significantly moderated students' learning.

Formative Approach Intervention related Moderators – Motivation

With regard to formative approach intervention related moderators, there were no records for informal formative approach. The results for other moderators such as the subject domain, group vs individual implementation, mode of implementation are as follows.

In order to find out if formative approach has differential motivational effects on various subjects, we ran an analysis with subject as moderator. Subject has a significant moderating effect: $F(5, 41) = 5.17$, $p < .001$. Only 'writing' significantly moderated the effect of formative approach on motivation: g = 1.87 [1.05, 2.70], k = 2, $p < .001$. But it should be noted that there were only two effect sizes for writing.

The omnibus test indicated that the effect of formative approach on motivation is moderated by group or individual implementation: $F(2, 51) = 3.40$, p = .04 with individual formative approach implementation having a significant effect on learners' motivation g = 0.44 [0.10, 0.78], k = 44, p = .01.

With regard to the mode of formative approach implementation, there was no moderating effect on motivation: $F(3, 50) = 2.55$, p = .07.

Formative Approach Intervention related Moderators - Learning

With regard to the implementation of the formative approach intervention, it can either be implemented formally in a structured manner by planning the steps ahead of implementation and designing the tasks (Ayala et al., 2008). Or the formative approach can be implemented informally or on-the-fly, which means, as and when the opportunity arises in terms of assessment conversations (Ruiz-Primo, 2011). In this meta-analysis, formal or informal formative approach had a significant moderation effect: $F(2, 451) = 69.70$, $p < .001$. Both formal and informal formative approach had significant medium effects on learning. Formal formative approach had an effect of $g = 0.43$ [0.35, 0.51], $k = 333$, $p < .001$ on learning and informal formative approach had an effect of $g = 0.51$ [0.33, 0.69], $k = 120$, $p < .001$.

In order to find out if formative approach has differential learning effects on various subjects, we ran an analysis with subject as moderator. Subject is a significant moderator of formative approach on learning: $F(4, 358) = 25.55$, $p < .001$. The moderating effect of the various subjects are presented in Table 4. Among the subjects, the moderating effect of reading was significantly larger than that of math.

Table 4

Moderating Effects of different Subjects on Learning

Subject	Effect size
Mathematics	$g = 0.35$ [0.25, 0.46], $k = 109$, $p < .001$
Science	$g = 0.42$ [0.24, 0.59], $k = 95$, $p < .001$
Reading	$g = 0.50$ [0.39, 0.61], $k = 136$, $p < .001$
Writing	$g = 0.51$ [0.28, 0.75], $k = 22$, $p < .001$

Formative approach can be implemented in groups or for individual learners. When analyzed whether the effect of formative approach on learning is moderated by group or individual implementation, there was a significant moderating effect, $F(2, 447) = 68.03$, $p < .001$. When implemented in groups, the effect of formative approach on learning was $g = 0.38$ [0.27, 0.50], $k = 200$, $p < .001$; when implemented individually, the effect was $g = 0.48$ [0.38, 0.57], $k = 249$, $p < .001$. Individual implementation of the formative approach has a larger effect on learning than group implementation, although not significantly larger.

Formative approach was implemented either in person or with the assistance of computer or by using paper-pencil tests. However, it should be noted that not all the three formative approach components can be implemented by all modes in reality. When tested whether the mode of implementation moderates the effect of formative approach on learning, it was significant: $F(3, 446) = 44.71$, $p < .001$. Paper-pencil tests were used predominantly for diagnostic assessment and had an effect of $g = 0.47$ [0.23, 0.70], $k = 38$, $p < .001$. Comprehensive feedback and adaptive instruction were implemented either in person or on computer. When formative approach components were implemented in person, the effect of formative approach on learning was $g = 0.41$ [0.31, 0.51], $k = 284$, $p < .001$. When implemented via computer, the effect was $g = 0.48$ [0.36, 0.60], $k = 127$, $p < .001$.

Multiple Moderator Analysis – Motivation

Moderator analyses discussed in the previous sections were conducted individually. As some of the moderating variables are interrelated, it could lead to multicollinearity (Hox et al., 2017). This precludes us from identifying the most relevant moderators. A solution to this problem is including all moderators in a single analysis. A multiple moderator analysis enables to identify the most relevant moderators (Hox et al., 2017).

When all the significant moderators of motivation were included in a single model, there was a significant moderating effect as indicated by the omnibus test $F(13, 40) = 2.23$, $p = .03$. However, none of the previously significant moderators uniquely moderated the effect of formative approach on motivation.

Multiple Moderator Analysis - Learning

In this meta-analysis, when all the significant moderators were included in a single model, there was a significant moderating effect as indicated by the omnibus test $F(22, 431) = 9.18$, $p < .001$. Among all the moderators, high school ($g = 0.67$ [0.10, 1.24], $k = 43$, $p = .02$) and tertiary ($g = 0.64$ [0.09, 1.18], $k = 117$, $p = .02$) stages of schooling, regular ($g = -0.54$ [-0.83, -0.26], $k = 348$, $p < .001$) and special education needs ($g = -0.56$ [-0.94, -0.18], $k = 85$, $p = .004$) were significantly moderating the effect of formative approach on learning. The results indicate that the high school, tertiary education, regular and special education needs uniquely moderate the effect of formative approach on learning.

2.2.10 Discussion

In this meta-analysis we investigated the existing empirical data on the formative approach. We quantitatively examined the effects of the four constellations of the formative approach components on learning and motivation. We also investigated whether the effects of the formative approach components significantly differed from each other in order to see if the components had differential effects.

The meta-analysis resulted in an overall small to medium size effect of the formative approach on motivation ($g = 0.34$) and a medium size effect of the formative approach on learning ($g = 0.45$). The formative approach components operating in four different constellations all had significant medium size effects on learning [$g = 0.37$ (DA); 0.51 (DA+CF); 0.38 (DA+AI); 0.42

(DA+CF+AI)]. With regard to motivation, three out of the four constellations had significant effects on motivation [g = 0.35 (DA); 0.44 (DA+CF); 0.33 (DA+CF+AI)], with DA+AI not impacting learners' motivation. The significant effects for the components provide empirical evidence to the theoretical model of the formative approach. Apart from identifying the effects of the formative approach components, our aim was to investigate if these effects differed significantly from each other. These effects did not significantly differ from each other. Nevertheless, we need to be cautious in interpreting what this implies, and we cannot directly affirm that the components did not impact learning differentially or motivation differentially. The formative approach should ideally enable accurate diagnosis of the learning status and learning process which in turn facilitates comprehensive feedback and adaptive instructions. Our purpose of looking precisely into the components of formative approach was to chiefly investigate and explain the reported variability in the effects of formative approach on learning. But the variability could not be explained further by discerning the formative approach into its components. Though the components of the formative approach are theoretically distinct, empirically distinguishing the effects of the components remains a challenge. This could possibly be due to the ways the components are implemented and measured in practice. We once again looked into the primary studies and even more precisely. There seemed to be discrepancies in how the authors conceptualized the formative approach and how it was implemented. Though this was not the case in the majority of the studies, some studies had discrepancies. Let us take a look at examples of primary studies with such discrepancies, which make it challenging to discern the effects of the formative approach components.

The formative approach intervention by Yin et al. (2008) aimed at embedding formative assessments in a science unit and investigated whether these assessments increased the learning

outcomes of students either directly or indirectly by increasing their motivation. A team of experts developed, piloted embedded formative assessment and engaged teachers to implement the embedded formative assessment for middle-school students on the science concept of why things sink and float (Ayala et al., 2008). In a five-day training, the experimental group teachers (N = 6) were introduced to embedded formative assessment and trained on implementing them with a model teacher and researchers demonstrating them the methods (Yin et al., 2008). Both raters of the current meta-analysis coded Yin et al. (2008) as diagnostic assessment component of the formative approach. Though the primary focus of the embedded formative assessments was to diagnose students' understanding on a specific science topic, students also received feedback on their conceptual understanding and were provided with instructions adapted to suit their learning needs. Despite the five-day training program for the teachers on embedded formative assessment, the quantity and the quality of feedback and instruction they provided to the students varied across the six teachers. The wide-ranging practices of the teachers despite receiving the same training could possibly be due to the lack of clarity on what the "embedded formative assessment" entails. The control group teachers in some instances provided comprehensive feedback and adapted instructions spontaneously and more effectively than the trained experimental group teachers. Though this intervention study focused on "embedded formative assessment", it is evident that it is not limited to diagnostic assessment, but also includes feedback and instructional components. Therefore, the effects are also not limited to the diagnostic assessment component.

In the intervention by Ozogul and Sullivan (2009), (coded as diagnostic assessment and comprehensive feedback) teacher education students' draft lesson plans were formatively assessed and they were provided with feedback. The conceptualization of diagnostic assessment and comprehensive feedback seemed clear as they were based on a scoring rubric. The students had

the opportunity to revise their draft lesson plans based on the feedback and submit their final versions. Taking a close look at the feedback of teachers showed that it contained instructional inputs on the content as well. The seemingly clear initial conceptualization of diagnostic assessment and comprehensive feedback was not actually complete. In practice the intervention included adaptive instructions. So, the learning effects of this intervention in terms of the progress they made from their draft versions to final versions cannot be exclusively attributed to diagnostic assessment and comprehensive feedback.

In the formative assessment intervention by Ponce, Mayer, Figueroa, and López (2018), (coded as diagnostic assessment and adaptive instruction) teachers used an interactive software for language instructions. Students read a text on computer and highlighted the words that they did not understand. These words appeared on the teacher's monitor in different colours ranging from most to least highlighted by students with the respective frequencies. The teachers displayed their screen to the students and adapted their instructions according to students learning needs. Although the focus in this intervention was to adapt teachers' instructions to cater to the students' learning needs, the process of displaying the highlighted words with their frequencies, accounts for feedback on their learning status. Though the goal of the study was to investigate the learning effects of the adaptive instructions, neither the implementation nor the learning effect is limited to the instructional component.

The primary focus in each of the above interventions was on one of the components of our proposed model of the formative approach. But in reality, the interventions contained other components as well. In practice, limiting the implementation to the specific components poses to be a challenge. Especially, the distinction between feedback and instruction seems to be vague for researchers as well as practitioners. The challenge could arise from perceiving feedback and

instruction on a continuum. Hattie and Timperley (2007) in their theoretical model of feedback recommended considering feedback and instruction on a continuum in order to understand the purpose, types and effects of feedback better. This perception of continuum does not seem to contribute to the theory or the practice. Feedback and instruction are two distinct constructs and viewing them on a continuum only adds to the chaos in the conceptualization of the formative approach. The ambiguity and the challenge that arose from the continuum perspective could be a possible explanation for the similar distribution of literature on feedback component (33%) and instruction component (31%). This unintended inclusion of formative approach components in practice is firstly due to unclear conceptualization of the formative approach and secondly due to practical implementation challenges. From the above examples of interventions and the implementations, it is clear that investigating the effects of individual components of the formative approach in classrooms and other teaching contexts is highly challenging and at times nearly impossible. As multiple factors play a role in the classroom scenario, clear limitations of the formative approach components during implementation is impractical. This explains why the effects of each of the formative approach components on learning although significant, are not significantly different from each other. Even the confidence intervals of the effect sizes are largely overlapping, leaving no room to draw inferences on differential effects of the formative approach components.

With regard to motivation, the current meta-analysis resulted in an overall small to medium size effect of the formative approach on motivation. The formative approach components operating in different combinations did not all have significant effects on motivation. The components diagnostic assessment and adaptive instruction did not have a significant effect on learning. On the contrary, the diagnostic assessment and comprehensive feedback components had a significant

medium size effect on motivation. Though this sheds light on how comprehensive feedback component impacts motivation to a larger degree, this effect was not significantly different from those of only diagnostic assessment component and the constellation of all three components. Similar to the findings for learning, these results do not yet clearly address the question of whether there are differential effects of the formative approach components on motivation.

The theoretical model of the formative approach model includes three components. Following the theoretical model, we wanted to find out if the components have differential effects on learning and motivation, respectively. But in practice, the components are almost always together varying only in focus and intensity. In the study by Hebbecker and Souvignier (2018), the authors had three experimental conditions, which are parallel to the three components of our formative approach model. Teachers in all three conditions were provided with diagnostic assessment data, one group of teachers received feedback on this assessment data (learning status and progress of students) and the third group received feedback as well as instructional material for supporting students' learning based on the assessment outcomes. Teachers in the feedback and instruction groups were also trained in using the feedback and instruction material and were also provided with a manual corresponding to their experimental condition. After a year of implementing the components, students' learning was assessed. All three groups had made progress, but the feedback and instruction groups did not show additional progress compared to the assessment group. This experiment illustrates the difficulty of disentangling the formative approach components in practice. It is possible that the teachers in the diagnostic assessment group provided feedback and adapted instructions spontaneously. Despite meticulous planning and designing of the three conditions, all components were present in all three conditions, presumably in different intensities.

2.2.11 Strengths and Limitations

The meta-analytic procedure has its own strengths and limitations. In our meta-analysis, we have ensured the key requirements for the validity of our results. We started out with clearly defined objectives of finding out the effects of the formative approach and its components on learning and motivation. By defining our protocol with the objectives, method and analysis plan as a first step, we eliminate the limitation of meta-analysis being a retrospective research procedure. Our step-by-step approach to study identification and the detailed documentation of the selection procedure accompanied by the PRISMA flow diagram contribute to the soundness of our results. Possible bias in study selection was precluded procedurally by means of clearly defined eligibility criteria, by acquiring unpublished research, and statistically by means of sensitivity analyses such as checking the funnel plot symmetry, checking for publication bias and removing the outliers. One major drawback is that of carrying over the weaknesses in the designs of the primary studies to the meta-analysis (Spector & Thompson, 1991). To address this issue primary studies could be rated on their quality. However, weighing the quality of the primary studies is tricky as it has the risk of becoming subjective and arbitrary. This might work against the primary purpose of the meta-analysis providing an objective evidence. Therefore, suitable analytic method for the variability in primary studies is a better option than rating the quality of the primary studies. By implementing the most appropriate analysis for our data, namely, multilevel meta-analysis, we ruled out sample dependency and possible deviations in the results. Finally, the fail-safe analysis confirmed the robustness of our results.

In our theoretical model we conceptualized formative approach as comprising the components of diagnostic assessment, comprehensive feedback and adaptive instruction. From the process of coding the components, it became evident that in practice the components exist in four

possible constellations with the diagnostic assessment component preceding the other two components. The prevalence of the components in this manner could lead to the risk of viewing the components as a chain of events, although that is not how we conceptualized formative approach components in our theoretical model. Our model clearly states the individual components of the formative approach but does not make claims on the sequence of their operation. A chain of events has a clear starting and an ending point, which need not necessarily be the case with formative approach components. Though it is possible to start with diagnostic assessment that leads to comprehensive feedback and ends with adaptive instruction, this chain is likely to occur multiple times and in various sequences in a learning scenario. A clear start and a clear end cannot always be well defined because, opportunities for further diagnostic assessment could arise in the middle of adaptive instruction or comprehensive feedback. Hence, an optimal way of viewing the components is to see them as a cycle of processes that can take place endlessly, similar to the multiple iteration cycles described by Ruiz-Primo and Furtak (2006) in their model of informal formative approach. This might be more helpful because, in a typical learning scenario the formative approach components are likely to repeat themselves. Therefore, the *cycle of processes* view would be more suitable than the *chain of events* view.

2.2.12 Next Steps

An alternate method to identify the effects of the formative approach components would be to investigate the time spent on each of the components in classrooms. In actual practice all three components seem to prevail, but in varying amounts of focus (Yin et al., 2008). Therefore, coding the component in focus is an alternate idea as opposed to coding the prevalence of components in primary studies. The drawback in this approach is that relatively fewer studies report the information on time spent on formative approach implementation (Förster & Souvignier,

2014; Hebbecker & Souvignier, 2018; Hung et al., 2013; Hwang & Chang, 2011; Wongwatkit et al., 2017). Also, the discrepancy in conceptualization and implementation of the formative approach in primary studies would persist in this approach too.

Besides the meta-analytic approach and now with a clear conceptualization of the formative approach, another option is to conduct a new experimental study to identify the effects of the formative approach components on learning and motivation. One prerequisite to clearly identify the effects of each of the formative approach components is a flawless pure implementation of the components. As this poses a challenge in a classroom setting due to the interplay of numerous factors, an online experimental study could be a feasible option. Firstly, in an online study the possibility of manipulating the components distinctly is more realistic. Secondly, we can spare the implementation fidelity aspect as the components are implemented by means of a foolproof software. Thirdly, we can plan and design the teacher-training program on formative approach more efficiently. Once we identify if there are significantly differential effects of the components on learning and/ or on motivation, we can invest more prudently in the training and implementation of the formative approach in classrooms.

2.3 An Explorative Study

Prior to operationalizing *diagnostic assessment* in the formative approach intervention, we wanted to find out whether *diagnostic assessment* and *testing* impact learning differently. Although it is possible to operationalize both diagnostic assessment and testing in a similar manner, they are two different constructs. Diagnostic assessment is the assessment conducted to diagnose the learning process and status. Diagnostic assessments are conducted mainly with the purpose of enhancing the learning. Delandshere (1990) defined diagnostic assessment as follows: "Diagnostic assessment draws a profile of student achievement, considering (a) the discrepancies between

expected and actual achievement, (b) the cause for such discrepancies, and (c) appropriate remedial treatment" (p.341). Testing primarily refers to assessing the amount of information retained by an individual and takes place mostly after a learning phase. Positive learning effects for testing has been reported in the literature (Pan & Rickard, 2018). We were interested in finding out whether diagnostic assessment and testing differentially impact learning. Based on the outcome we wanted to decide on the operationalization of diagnostic assessment in our current experimental study.

In our pilot study, we investigated whether priming the learners on the diagnostic purpose of the assessment leads to better learning outcomes than only testing them. We conducted an online experimental study with a pre-post-test design that included a learning phase in-between. The learning material used in this study is a text on the topic "How a cold virus attacks the body" (Mayer, Griffith, Jurkowitz, & Rothman, 2008). We used a German version of the text which contained around 500 words. We developed 15 multiple-choice-questions based on this text for our study. Out of the 15 questions, five were retention questions and ten were transfer questions. Each question had three options with one correct answer. The sample comprised of 74 teacher education students who were randomly assigned to one of the three conditions, namely, the diagnostic assessment, the testing, and the control condition. The participants of the diagnostic assessment and testing groups answered the 15 multiple-choice-questions in the pre-test, read the text on cold virus and answered the same set of questions for the post-test. The control group read the text on cold virus and answered the 15 multiple-choice-questions. To balance the amount of effort across the three conditions, the control group answered a quiz on random topics for the pre-test. While the participants in the testing and control groups were solely asked to answer the questions, the participants in the diagnostic assessment condition were primed on the diagnostic purpose of the questions. The diagnostic assessment group received the following prime before

the pre-test: "By answering these questions you will know what information you need further to understand this topic. Therefore, not knowing an answer provides you with important information."

There was no difference between the three groups $F(2,71) = 1.953$, $p = .149$. Contrast analyses did not indicate any significant difference between the learning outcomes of the diagnostic assessment and testing groups. Thereby testing the learners without priming them on the diagnostic purpose is just as effective as informing them of the purpose. Based on the outcomes of our pre-test, we decided to operationalize diagnostic assessment in the current study as testing the participants.

3 An Experimental Study on the Effects of Formative Approach on Learning & Motivation

Formative Approach

As conceptualized in an earlier chapter of this dissertation, the formative approach is characterized as practices of diagnosing students' learning progress, providing the students with comprehensive feedback, and adapting instructions based on the diagnostic outcomes.

3.1 Formative Approach and Learning

Research findings on the effect of formative approach on learning is wide ranging. While some studies report positive findings (Paiva, Ferreira, & Frade, 2017; Ponce et al., 2018; Wesson, 2013), others report minimal or null effect of the formative approach on learning (McNulty et al., 2015; van den Berg et al., 2016; Yin et al., 2008). It is unclear why some studies report positive findings and others do not. This leads to the question of whether different components of formative approach have different effects. And whether the positive findings of formative approach are due to one or more of the formative approach components.

Alternatively, the varying effects of formative approach on learning might be explained by learners' motivation. The role of formative approach for learners' motivation will be illustrated in the following section.

3.2 Formative Approach and Motivation

Similar to the effects on learning, the effects of the formative approach on motivation is also wide ranging. There are studies reporting positive effects (Yin et al., 2008) as well studies reporting null effects (Förster & Souvignier, 2014) or even negative effects (Hebbecker & Souvignier, 2018) of the formative approach on motivation. We are interested in finding out if one or more components of the formative approach impact motivation differently. Motivation is a broad construct and entails multiple aspects, e.g., achievement motivation (Dweck, 1986), goal-orientation (Dweck &

Elliot, 1988), self-efficacy (Bandura, 1977). In the current study we focus on the achievement motivation of the learners as it is the most relevant construct for the formative approach intervention.

Achievement motivation can be exemplified based on multiple theoretical frameworks. In the following section is an outline of the theories on achievement motivation along with links drawn to the formative approach.

Social Cognitive Theory: This theory illustrates the self-efficacy aspect of achievement motivation. Facilitating learners to believe in their efforts could make them feel self-efficacious which in turn motivate them to invest further effort (Bandura, 1977). When practices of formative approach inform learners on how they can learn, how to invest effort in a focused manner or modify their learning strategies, the learners' beliefs are directed towards their ability.

Attribution Theory: This theory explains how learners' attribution of their success or failure determines their achievement motivation (Weiner, 1985). The attribution could be on one or more of the following three dimensions, locus (internal vs. external), stability (temporally stable vs. unstable), controllability (controllable vs. uncontrollable). When learners receive feedback on how they can improve and what they need to do to improve, their locus of control is directed to one-self, failure is seen as temporally unstable and they are likely to interpret the situation controllable (Weiner, 1985).

Expectancy-Value Theory: Wigfield and Eccles' (2000) expanded model of the initial expectancy-value theory (Atkinson, 1957) illustrates how achievement motivation results from one's expectations for success and the value of the tasks. One's expectations for success is determined by the self-concept of their ability and their perception of task difficulty. The value for

task is determined by the significance for self-image, intrinsic value or interest, usefulness for future goals and the costs involved. When the formative approach intervention caters to the expectations for success and value for tasks, learners are motivated to engage in achievement related attitudes and choices.

Self-Determination Theory: Autonomy, competence and relatedness are the three basic psychological human needs that facilitate self-determined motivation (Ryan & Deci, 2000). Autonomy refers to the perception of free-choice and decision; competence refers to the perceived ability to perform; and relatedness refers to the interpersonal connections with others. In a learning environment, gratification of these psychological needs of the learners result in self-determined motivation which in turn lead them to invest more effort in learning. Possibilities for the formative approach components to fulfill one or more of these needs exist.

Achievement Goal Theory: This theory focusses on achievement related behaviours unlike the above-mentioned other theories that focus on other aspects of motivation. Achievement related behaviours include investing more effort in learning and enhancing one's competencies or displaying one's abilities and competencies. Focusing on the former leads to learning/ mastery/ task goals and focusing on the latter leads to performance/ ability/ ego goals (Dweck, 1986; Dweck & Elliot, 1988; Nicholls, 1984). While individuals with learning goals pursue knowledge and skills with the aim of learning and mastering, individuals with performance goals pursue affirmation of their knowledge and skills and avoid any criticism (Dweck, 1986). As the formative approach intervention in the current study directs learners' focus towards their learning, our focus in this study is on the motivational construct *learning goal orientation*.

Formative approach being an intervention that directs learners' focus towards the task, it is possible that the learners are more motivated to engage in the task. The task involvement facilitated

by formative approach could foster learning goal orientation (VandeWalle, Cron, & Slocum Jr, 2001).

3.3 Transfer Learning Effects of the Formative Approach

"To say that learning has occurred means that the person can display that learning later" (Perkins & Salomon, 1992). When the display happens by putting the learning into use it is referred to as transfer of learning (Perkins & Salomon, 1992). Transfer of learning takes place across temporal contexts and knowledge domains (Carpenter, 2012). When the transfer of learning occurs in similar contexts and domains it is called near-transfer and when it occurs in novel contexts and across domains it is called far-transfer (Perkins & Salomon, 1992).

In our current study, we operationalized near-transfer learning in two ways. Firstly as the participants' display of learning in terms of scores in the same multiple-choice-test across temporal contexts. Secondly as the participants' display of learning in terms of scores on an overarching open-ended-question across temporal contexts.

3.4 Current Study

The formative approach as conceptualized and defined in an earlier chapter of this dissertation, is characterized as practices of diagnosing students' learning progress, providing them comprehensive feedback, and adapting instructions based on the diagnostic outcomes. Thereby, the components of formative approach are diagnostic assessment, comprehensive feedback, and adaptive instructions. The current study aims to examine the effects of the formative approach components on learning and motivation. Although there is some research evidence that formative approach has positive effects on learning and motivation, this study investigates whether the positive effects are due to one component of formative approach or different combinations of formative approach components. While the research on the effects of formative approach on

learning is extensive, the research on the effects of formative approach on motivation is minimal. Therefore, investigating the effects of formative approach on learning and motivation will shed more light on the effects of each of the formative approach components. The current study was an online experiment, wherein the formative approach components were implemented online. The main reason for an online experiment was to ensure pure implementation of the formative approach components and to eliminate the confounding effects of implementation fidelity variable.

From the systematic review and the meta-analysis, we have evidence for the prevalence of the three components and their positive effects on learning and motivation. Through this experiment we investigate whether the formative approach components have differential effects on learning and motivation.

3.5 Research Questions

What are the effects of the formative approach components diagnostic assessment (DA), comprehensive feedback (DA+CF) and adaptive instructions (DA+CF+AI) on near-transfer learning in the a) short-term and b) in the long-term?

What are the effects of the formative approach components diagnostic assessment (DA), comprehensive feedback (DA+CF) and adaptive instructions (DA+CF+AI) on motivation in the a) short-term and b) in the long-term?

Are there differential effects of the formative approach components diagnostic assessment (DA), comprehensive feedback (DA+CF) and adaptive instructions (DA+CF+AI) on learning?

Are there differential effects of the formative approach components diagnostic assessment (DA), comprehensive feedback (DA+CF) and adaptive instructions (DA+CF+AI) on motivation?

3.6 Method

3.6.1 Participants

Teacher education students enrolled in the educational psychology seminar participated in this study. This is a preparatory seminar for them before they start their school internship. There were six seminar groups and the expected sample size was 150. The expected effect size calculated with G*Power analysis program (Faul, Erdfelder, Lang, & Buchner, 2007) for a sample size of 150 with a statistical power of ≥.90 was a medium size effect of $d \geq .30$.

Students participated voluntarily in the study. Totally 119 students (81F, 38M) with a mean age of 20.8 years (SD = 2.49) participated in the study. The students were in the second or fourth semester of their bachelor program in teacher education. For the actual sample of 119 participants we still obtained a medium size effect of $d \geq .30$ with a statistical power of ≥.90.

3.6.2 Stratified Randomization

The educational psychology seminar took place in six groups for three days each. Although the total duration of the seminar for all the six groups was the same, the time interval between the first two seminar sessions ranged between one and fifteen days and the time interval between the second and third seminar sessions ranged between forty and eighty days for the six groups. It was therefore important to ensure randomized distribution of participants in the three formative approach conditions across the six groups of participants. The participants within each group were randomly assigned to the three conditions. The online study enabled the random distribution without any hassle.

3.6.3 Test Materials

We conducted an online study with a pre-, post- and follow-up test with three experimental conditions. During the pre-test, the demographic data were collected, and the formative approach

intervention was conducted while answering the multiple-choice-questions. Following the multiple-choice-questions, participants responded to items on motivation and at the end answered a content related open-ended-question. The sequence of the three sections (multiple-choice-questions, motivation items and open-ended question) remained the same across all three time points. The multiple-choice-questions appeared in random sequence at every time-point. The diagnostic assessment (DA) group answered all the items at all three time points. The comprehensive feedback (DA+CF) group received feedback on whether their chosen option was correct or not during the pre-test. During post and follow-up test they did not receive any feedback. The adaptive instruction group (DA+CF+AI) received content related instructions along with the feedback on correctness of their response during the pre-test. During post and follow-up test they did not receive any feedback or instructions. The comprehensive feedback and adaptive instructions for every option in each question is presented in appendix C.

Learning

The learning effects in this study were measured as near-transfer effects of the formative approach intervention. The study participants were tested on their seminar content on educational psychology. For measuring near-transfer learning, 20 multiple-choice-items on the seminar content were implemented. The process of developing and validating the items is explained in a subsequent section. In addition, the participants answered an open-ended question on "classroom management." The development of scoring rubric for this question is also illustrated in a later section.

Motivation

For measuring motivation, we used four items on learning goal orientation which were adapted from the German scale for measuring learning and achievement motivation *SELLMO, Skalen zur*

Erfassung der Lern-und Leistungsmotivation (Spinath, 2012). The items were implemented in German language but adapted to suit the task in the current study. The internal consistency of the items was $\alpha = 0.81$. The items were on a five-point rating scale ranging from 1 = *completely disagree* to 5 = *completely agree*. The adapted items used in the current study were: "While answering the questions it was important for me…

a) to learn as much as possible on the topic of "psychology for the teaching profession"
b) to really understand what I can learn by performing this task
c) to use this task to recognize my potential for improvement
d) to acquire insights on how I can improve my psychological knowledge for the teaching profession.

Developing the Formative Approach Intervention Materials

The formative approach intervention was planned for a seminar on educational psychology for bachelor students in teacher education program. The course content included topics such as occupational profile of teachers, types of teacher behaviours and emotional experiences, diagnosing the behavioural patterns leading to burnout in teachers, self-regulation in teachers, characteristics of a good classroom, lesson planning and classroom management.

Determining the Time-points for the Tests and the Intervention

The goals of the three-day seminar were to enhance the students' understanding of the teaching profession, impart the fundamental elements of classroom teaching and enable the students to acquire a clearer and a more realistic perception of the profession. After identifying the goals of the seminar, it was crucial to determine when to incorporate the formative approach intervention in the seminar, when to test the short-term and long-term effects on learning and motivation. The optimal time-point for the formative approach intervention is before the commencement of

theoretical inputs. Therefore, we decided to implement the formative approach intervention during the pre-test. The optimal time point for testing short-term learning effects is immediately after the conclusion of theoretical inputs and therefore decided for the end of the second seminar day. A longer time interval is essential to test the long-term learning effects. Therefore, we decided to test the long-term effects on the third and final day of the seminar, when the students came back from their four-week internship in a school. It is presumable that the students had the opportunity for deep learning and reflection during their internship.

We wanted to investigate whether our formative approach intervention had learning-transfer effects. As the intervention was directly based on the seminar content and we had incorporated the comprehensive feedback and adaptive instructions in the multiple-choice-questions, we decided to test the near-transfer effects by means of these multiple-choice-questions and an open-ended question on the central theme of the seminar "classroom management." Details on the development of the multiple-choice-questions and the coding scheme for the open-ended question are elucidated in the ensuing sections.

Development of the Multiple-Choice-Questions

Based on the seminar content we developed a pool of multiple-choice-questions. The questions were developed based on the guidelines for developing multiple-choice-items provided by Haladyna, Downing, and Rodriguez (2002). Each question had three options, with one option being the correct answer and the remaining two options serving as distractors. We did not add additional distractors, as they neither enhance the psychometric properties nor is it efficient in terms of time and resource to generate them. Meta-analytic data show that three option questions are optimal for multiple choice questions (Rodriguez, 2005) and are equally good as four option questions (Tarrant & Ware, 2010). Furthermore three option questions are most suitable for ability

and achievement tests (Delgado & Prieto, 1998). Three option questions consume less time for answering and have more facility value (Vegada, Shukla, Khilnani, Charan, & Desai, 2016). With regard to the validity, reliability and item discrimination, there is no difference between five, four and three option questions (Vegada et al., 2016).

Development of the Test Materials

The test materials for the current study were developed systematically in an iterative manner. The multiple-choice-test was developed based on the seminar content on educational psychology. The multiple-choice-test assessed the near-transfer learning effects of the formative approach intervention. An initial pool of 35 items (with three options each) were generated on the seminar content. These items were then screened for recurrences and redundancies of item content and the item options. This screening resulted in removing ten items from the initial pool. The remaining 25 items were validated by two experts in the field of educational psychology independently. These two experts regularly lecture the teacher education students on the topic of educational psychology. The current seminar is a supplement to the lecture series of these two experts that the students attended in their previous semester. The construct validation by the two experts resulted in refining the items and their options. The final pool comprised of 20 items.

Development of the Coding Scheme for the Open-ended Question

The purpose of an open-ended question in the experiment was to assess the transfer learning effects of the formative approach intervention. The open-ended question was based on the overall seminar content and is as follows: "How would you explain an efficient classroom-management in about 100-150 words?" A scoring rubric for this question was developed based on the seminar content, more specifically based on the power-point slides of the seminar. As a first step all aspects of classroom management from the slides were listed. As a second step all these aspects were

classified into broader categories. In the third step, a scoring rubric was developed, wherein the broader categories received a score each and further explanations for each of these categories received additional points. For example, "Dealing with disruptions" was a broader category which received a score of one. When elaborated with explanations such as "anticipate and deflect disruptive behavior", "intervene early on and deescalate disruptions quickly" then an additional score is provided for each explanation. The maximum possible score for the answer was 10.

Cognitive Pretesting of the Test Materials

Following the meticulous screening and expert validation of the items, the items were cognitively pretested as a final step. Five participants who participated in the cognitive pretest reported their experiences on taking the test, how understandable the items were and how long it took them to complete the test. Based on their feedback that the sequence of questions in two places facilitated answering the questions that followed, we changed the order in which the items were presented. The participants' feedback also helped in rephrasing and refining the adaptive instructions for some of the item options.

Item Difficulty, Discrimination Index and Distractor Efficiency

Based on the pre-test scores of the participants we calculated the item difficulty index (DIF I), discrimination index (DI) and distractor efficiency (DE). Following are the tables with the indices presented in the same format as that of D'Sa and Visbal-Dionaldo (2017).

Table 5

Difficulty and Discrimination Indices of the Multiple-Choice-Questions

Parameters	Items (n = 20)	Percentage	Interpretation	Mean (SD)
Difficulty Index (DIF I)				72.94 (22.78)
> 70	14	70	Easy	
30-70	5	25	Good	
< 30	1	5	Difficult	
Discrimination Index (DI)				0.48 (0.35)
> 0.35	10	50	Excellent	
0.25 – 0.34	4	20	Good	
0.15 – 0.24	4	20	Marginal	
< 0.15	2	10	Poor	

The indices in the above table indicate that 25% of the items had good difficulty index and 70% of the items had good discrimination index. After combining the two indices totally 70% of the items (14 out of 20 items) can be declared ideal.

The quality of the multiple-choice-questions is determined by the quality of the distractors. Distractors are efficient when they are high on plausibility, based on typical errors or misconceptions and are not blatantly illogical (Haladyna et al., 2002). A distractor is statistically

efficient if more than 5% of the participants have opted for it and non-functional distractors are the ones opted by less than 5% of the participants (Tarrant, Ware, & Mohammed, 2009).

In order to identify how efficient the distractors of the multiple-choice-questions in the current study are, we calculated the distractor efficiency. Out of the 40 distractors, 23 distractors were efficient, functional, and opted by more than 5% of the participants. Remaining 17 distractors were non-functional. Thereby, the distractor efficiency was diverse among the items.

Table 6

Distractor Analysis: Distribution of Items with Non-functional Distractors

Non-functional Distractors	Distractor Efficiency	Number of Items (n = 20)	Percentage of Items
0	100%	8	40
1	50%	7	35
2	0%	5	25

3.6.4 Instructors

With regard to the seminar instructors, there were five of them for the six groups. One of them instructed two groups. Although the instructors were not involved in the formative approach intervention or in the online study, it is possible that they deal with the seminar topics in different intensities. Two measures were taken to reduce the variance in their instruction and thereby the variance in students' learning as a result of instruction. Firstly, all instructors had the exact same power-point slides for their seminar. The slides on which the test materials were developed were marked and commented that they should not be omitted or overlooked during the instruction. Secondly, all instructors were provided with the test material of twenty multiple-choice-questions

items. They were requested not to skip or overlook the content related to these questions. However, instructors' adherence to the two measures were not observed, recorded, or documented.

3.6.5 Ethical Considerations

When conducting an experimental study, there are four essential ethical principles we need to ensure (Vaus, 2001). They are (a) voluntary participation; (b) informed consent; (c) no harm to participants and (d) anonymity and confidentiality. Participants in our study are adults and all required information was obtained directly from them. Their participation was completely voluntary, and they had the option of withdrawing from the study at any point without any consequences for them. As the participants were adults, they were entitled to provide their consent for participation. Only after providing their informed consent participants proceeded with the study. The participants were in no way harmed by their participation in this study. Participants were neither denied of anything nor harmed as a result of their participation. We ensured the anonymity of the participants. No demographical data could directly retrieve personal identification data. No personal data was obtained from the participants. All data obtained were treated confidentially and used only for the purpose of this experiment.

3.7 Analysis

To answer our experimental research questions, we used Analysis of Covariance (ANCOVA). ANCOVA allowed us to investigate the effects of the independent variables (formative approach components) on the dependent variables (learning and motivation), while controlling for the effect of the covariate (pre-test scores). Participants who did not take part in the pre-test but had taken part in the post-test (five participants) and follow-up test (seven participants) were eliminated from the data, as the pre-test served as baseline scores and the formative approach intervention was carried out during the pre-test. This occurred merely because of the participants' absenteeism

during the first seminar session. The experimental attrition at the post-test and follow-up-test were 4.2% (five participants) and 17.65% (21 participants) respectively. This attrition was due to absenteeism as all participants who attended the seminar participated in the study. As the dropout was at random, we verified whether the data were missing completely at random (Little, 1988). By conducting Little's MCAR test, it was confirmed that the data were missing completely at random (χ^2 (79) = 67.85, p = .81).

3.7.1 Covariate

We wanted to verify whether the three groups started out at the same level. Therefore, we analysed the groups on the covariate "pre-test scores." There was a significant difference between the groups on pre-test scores, $F(2,116) = 3.725$, $p = .027$. The pre-test score of the DA group ($M = 15.49$, $SD = 2.28$, $n = 37$) was significantly larger than the DA+CF+AI group ($M = 14.05$, $SD = 2.62$, $n = 40$). Since the groups vary on the pre-test scores, the validity of ANCOVA for the analysis is questionable. Although ANCOVA is considered invalid for a significant covariate, there are additional aspects to be considered. Firstly, random assignment of participants to experimental conditions, which is fulfilled in our experiment. Secondly, covariate is not a function of the dependent variable of formative approach groups as the group assignment did not cause the difference in the covariate (Field, 2013; Miller & Chapman, 2001). Therefore, ANCOVA in our analysis helps in removing noise variance in the groups.

3.8 Results

After controlling for the pre-test scores, there was a significant effect of the formative approach components on near-transfer short-term learning, $F(2,111) = 7.694$, $p = .001$, but no significant effect on long-term learning, $F(2,95) = 2.424$, $p = .09$ (Figure 8). However, all three groups showed progress in their learning scores across time (Figure 9). With regard to the open-ended

question, there was neither short-term effect $F(2,111) = 1.211$, $p = .302$, nor long-term effect $F(2,95) = 0.138$, $p = .872$. Although not significant, the DA+CF and DA+CF+AI groups showed progress in their scores at the short-term level, which did not sustain though (Figure 10)

When analysed whether the formative approach components have differential effects on learning, DA+CF+AI performed the best ($t = 3.92$, $p < .001$), followed by DA+CF ($t = 2.07$, $p = .041$) and then followed by the DA group in the near-transfer short-term learning (as illustrated in the post-test scores in Figure 8). Similarly, in the near-transfer long-term learning, DA+CF+AI performed the best ($t = 2.01$, $p = .047$) followed by DA+CF ($t = 1.73$, $p = .087$) and finally followed by the DA group (as illustrated in the follow-up scores in Figure 8).

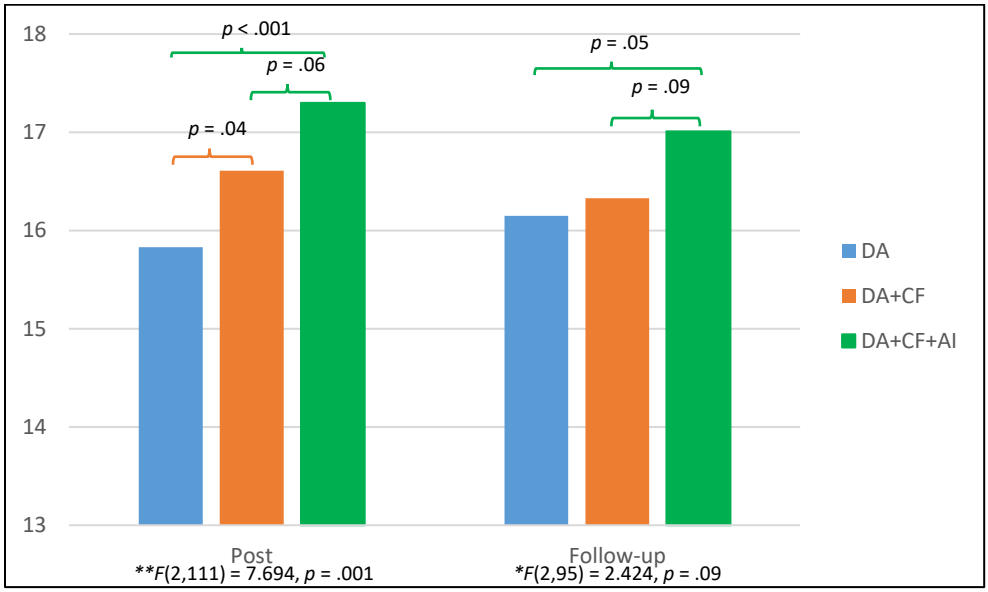

Figure 8. Differential Effects of the Components on Near-transfer Learning

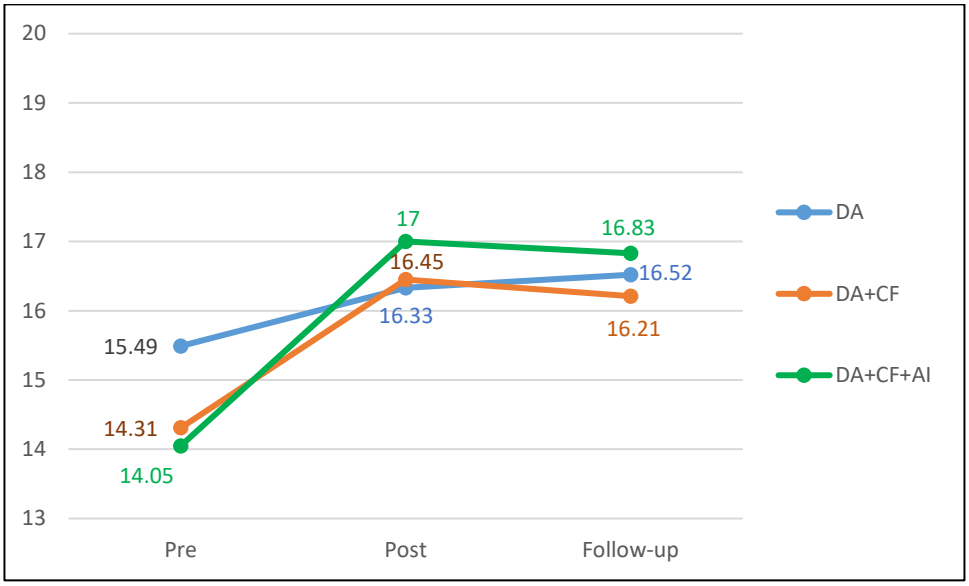

Figure 9. Near-transfer Learning Scores

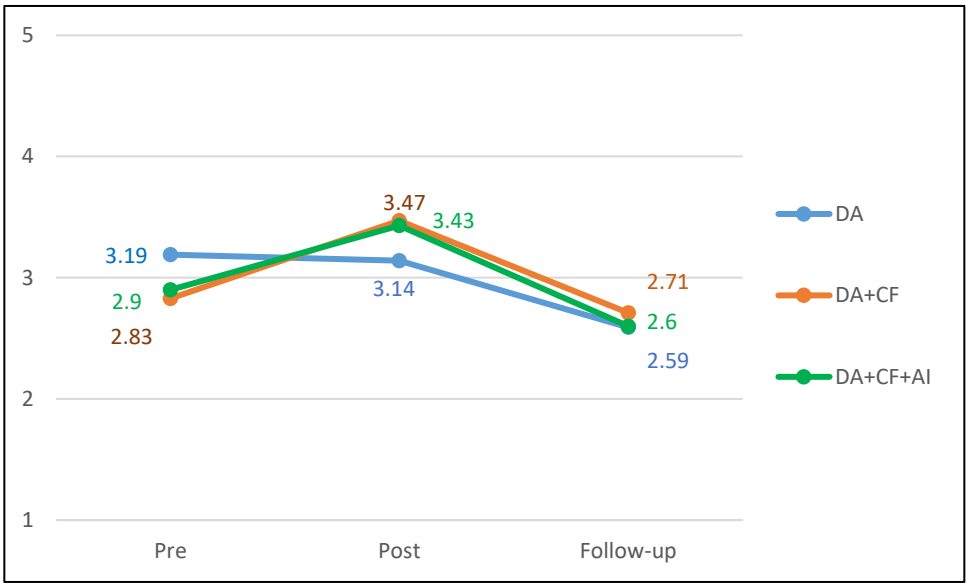

Figure 10. Scores on the open-ended question

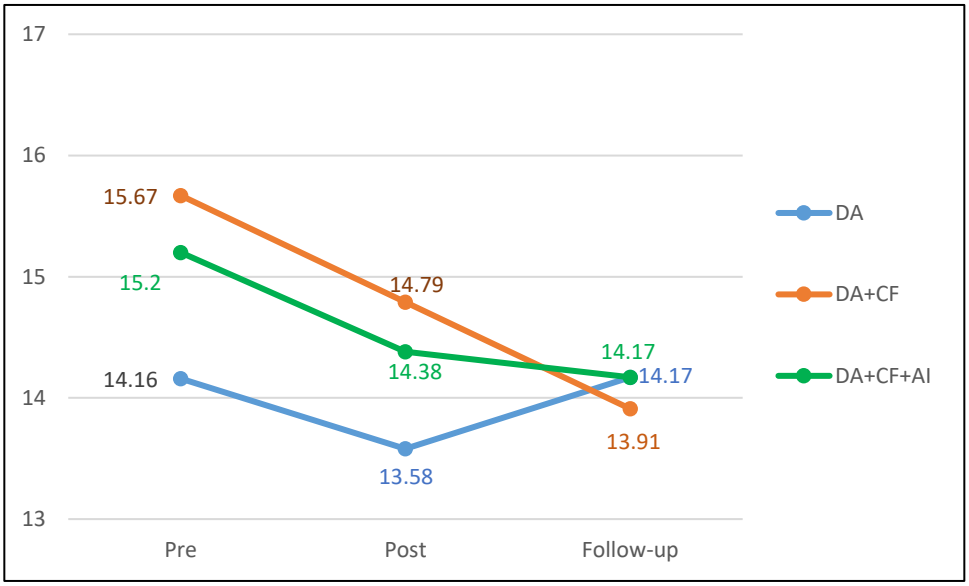

Figure 11. Motivation Scores

When analysed within-group differences over time, all three groups had made significant near-transfer short-term learning progress. But in near-transfer long-term learning, only DA+CF and DA+CF+AI groups had made significant progress (Appendix B).

After controlling for the pre-test motivation scores, there was no significant short-term effect $F(2,111) = 0.100$, $p = .91$, or long-term effect $F(2,93) = 0.445$, $p = .64$ of the formative approach components on motivation. Group wise comparisons did not show any significant difference over time. Motivation scores dropped over time across all three groups (Figure 11). However, at the pre-test, DA+CF group had reported significantly higher motivation than the DA group ($t = 1.93$, $p = .056$).

3.9 Discussion

In this experimental study, we investigated the effects of the formative approach components, namely diagnostic assessment, comprehensive feedback, and adaptive instructions on near-transfer learning and motivation. In our theoretical model, we proposed that the formative approach comprises the three components diagnostic assessment, comprehensive feedback, and adaptive instructions. Our systematic review of the literature on formative approach and coding the formative approach components in primary studies resulted in the finding that formative approach components prevail in constellations, namely, diagnostic assessment (DA), comprehensive feedback (DA+CF) and adaptive instruction (DA+CF+AI). We postulated that the three constellations of formative approach components have differential effects on learning and motivation. Although the constellation DA+AI prevailed, we refrained from including this constellation in our experiment for two reasons. Firstly, to keep the experimental design simple. Secondly, the effect of the AI component should become evident in the DA+CF+AI constellation.

Following the systematic review and the coding, the meta-analysis yielded empirical evidence for the formative approach model and its components. The meta-analysis resulted in significant medium size effects of the formative approach on learning and motivation. The different constellations of the formative approach components all had significant medium size effects on learning. Although it seemed that the components all had similar effects on learning, the different width of the confidence intervals of the effect sizes indicated large variance. A statistical comparison of the effect sizes did not result in significant difference between the effect sizes. A visual interpretation on the difference between effect sizes from the confidence intervals was also not possible as the confidence intervals were largely overlapping (Cumming, 2009). As a result, the question of whether the different components impact learning differently remained

unanswered. With regard to motivation, diagnostic assessment (DA) and comprehensive feedback (DA+CF) components had significant effect on motivation, but the effect sizes did not differ significantly from each other. Thereby the question of differential impact of the formative approach components remained unanswered for learning as well as motivation. A possible reason for the indistinct effects is that the formative approach components in the primary studies prevailed in varying amounts or intensities. The coding of formative approach components in the primary studies was based on their predominance, which was the only possible way to tease apart the components. Therefore, we could not completely exclude the impact of formative approach components that accompanied in minimal amounts and the impact of those that were implemented unintentionally. As a result, it was challenging to identify the effects of the individual components. Therefore, we decided to conduct our own experiment, wherein we operationalized and implemented the formative approach components precisely and investigated the effects of the components on learning and motivation. By preventing any discrepancy between operationalization and implementation and controlling for confounding variables, we were able to better investigate the effects.

3.9.1 Near-transfer Learning

Near-transfer learning was tested by means of the 20 multiple-choice-questions based on the seminar content. Following the results of the meta-analysis on the effects of the formative approach on learning, the results of the current experiment once again confirms that the formative approach can be an effective tool to enhance learning. By operationalizing the components of the formative approach in the most straightforward forms and implementing them precisely in an online experiment, we have a clear picture of the differential effects of the components. The results indicate that the formative approach components had differential effects on near-transfer learning

both in the short-term (post-test) as well as in the long-term (follow-up test). Adaptive instruction (DA+CF+AI) had the highest effect followed by comprehensive feedback (DA+CF) component followed by diagnostic assessment (DA).

From the diagnostic assessment component, it is possible that the learners became somewhat aware of what they knew and what they needed to learn. Unlike open ended questions where it is difficult to answer without prior knowledge, MCQs provide the possibility of choosing answers from the provided options. The availability of the options, apart from providing the scope for guessing (Bereby-Meyer, Meyer, & Flascher, 2002), induces other answering strategies. When one does not exactly know the answer for an MCQ, it is likely that they apply one of the following two methods. Either choosing the most plausible option or eliminating the less plausible options (Parikh, Sai, Nema, & Khapra, 2018). This process of choosing and eliminating options could evoke an awareness of the gaps in their knowledge. The slight awareness of the learners however remains uncertain without any feedback. By receiving comprehensive feedback on the correctness of their response, the awareness of their learning gaps possibly increased and became more concrete. Still, filling their learning gaps was left to the learners at this point. When the learners received instructions that were adapted to their learning needs, the certainty of reducing their learning gaps increased, thereby enhancing their learning.

Positive effects for feedback reported in the literature so far are partly effects of instruction on learning. Most research on feedback includes instruction, although feedback and instruction are two different aspects. In fact, one of the seminal articles on feedback by Hattie and Timperley (2007) suggests considering feedback and instruction on a continuum in order to understand the model of feedback. Although the challenge in distinguishing feedback and instruction is understandable, considering both as feedback is misleading for researchers as well as practitioners.

The adaptive instruction (DA + CF + AI) and comprehensive feedback (DA+CF) components of the formative approach had not only immediate effects but also long-term significant effects on learning. The diagnostic assessment (DA) component had only significant short-term effect on learning. These results confirm the long-term impact of adaptive instructions and comprehensive feedback on students' learning. The instructions that were adapted to address their learning gaps have possibly enabled them to reflect more and engage in deep learning. The feedback they received on the correctness of their response could have enabled them to reflect to an extent, but not as much as the adaptive instructions.

With regard to the participants' scores on the open-ended question, there was no significant increase in the scores in the short-term (post-test) as well as in the long-term (follow-up test). The students were asked to describe "effective classroom management", which was one of the main topics in their seminar. Although the teacher education students received inputs on classroom management during the seminar, they did not receive comprehensive feedback or adaptive instructions for their open-ended answer in our formative approach intervention. For reasons of practical implication, the implementation of the formative approach components was limited to the MCQs. As the students were left to reflect on their own, only a small percentage of the students showed increase in scores for their open-ended answers, however not significant. We verified whether the formative approach components at the pre-test had any effect on their open-ended answers at the pre-test itself, but there was no effect.

3.9.2 Effects on Motivation

With regard to motivation (learning goal orientation), there was no significant effect of the formative approach components. Across all the formative approach components there was an overall decline of motivation from pre-test to follow-up test. The reason for higher motivation in

the pre-test could be due to the sequence of test items and the timing of formative approach intervention. The formative approach components were implemented at the beginning of the pre-test along with the multiple-choice-questions and subsequently the items on motivation were presented. The students thereby underwent the formative approach intervention before they responded to the items on motivation. Undergoing the intervention could have resulted in higher motivation in the pre-test as opposed to the post and follow-up tests, where they were only assessed. When taking a test or answering multiple-choice-questions, learning becomes an innate goal (Carpenter, 2012; Pan & Rickard, 2018) and consequently the efforts are directed to obtaining as many answers correct. The learners automatically orient themselves toward an achievement goal (Dweck, 1986; Dweck & Elliot, 1988; Nicholls, 1984). So, when learners receive information that directs them towards their goal, they are motivated (Shute, 2008b). In this study, the formative approach intervention at the pre-test provided information that directed the learners toward their innate goal, which resulted in higher motivation at pre-test.

Though all the three groups reported higher motivation at the pre-test, there seemed to be differences between the groups. Contrast analysis at the pre-test level indicated slightly higher motivation for the comprehensive feedback (DA+CF) group than the diagnostic assessment (DA) group (p = .06). It is possible that receiving feedback on the correctness of their response had an immediate effect on students' motivation. This increase in motivation did not sustain over time and therefore there was no effect at the post-test or follow-up time points.

According to the self-regulation model (Boekaerts & Corno, 2005), the learning goals are adjusted as per the performance. Although the learners did not receive any score, it is possible that they realized how well they performed in the pre-test. Apart from the DA group, the other two groups received feedback on the correctness of response and were better aware of their

performance. Based on this awareness the learners possibly downregulated their learning goals, which explains the drop in motivation scores across time.

The results of the meta-analysis by Kluger and DeNisi (1996) on the effects of feedback illustrate how feedback at task level (correct or incorrect) has only minimal augmenting effects on learning and motivational processes. This is because, with increasing complexity of the task and for meta-task processing, task level feedback is rather insufficient to enhance learning and task-motivation in the long-term and thereby insufficient for transfer of learning (Kluger & DeNisi, 1996).

3.10 Strengths and Limitations of the Experiment

In the current experiment, we see how every additional component of the formative approach increases learning. The differential effects that were not evident from the meta-analytical results have become clearer from the current experiment. This was possible because of the following aspects. Firstly, by means of an online experiment we were able to precisely implement the components of the formative approach. Through our exact implementation of the formative approach components, we established high internal validity.

Secondly, through our online experiment we could eliminate and control multiple confounding variables. We could eliminate the confounding variable "implementation fidelity" in our online treatment, which is often a challenge in field experiments. The fidelity of implementation and treatment effectiveness are closely related (Furtak et al., 2008). Pertaining classroom learning context we controlled for the confounding variables through the following measures. All the seminar groups were from one university and from the same cohort. All instructors used the same material for their instruction. In order to further reduce the variance in

the instructions, all instructors were provided with primes in the instruction material to not skip specific content. The instructors were also provided with the test material (multiple-choice-questions) for the reason of reducing the variance in their instructions. Altogether our online experiment meticulously dealt with "implementation fidelity", classroom related variables and the variance in "quality of instructions."

Thirdly, we restricted the teacher related variables, which were cited as a challenging variable in most of the field experiments. For example, in the randomized experiment by Yin et al. (2008) a formative approach intervention was developed by an expert team and embedded in the curriculum. Despite a systematic planning and designing there was no significant difference between the experimental and control groups in their learning. It was subsequently found that the lack of learning effects was due to experimental teachers' lack of thorough implementation of the formative approach and control teachers' spontaneous implementation of formative approach methods (Furtak et al., 2008). Teachers tend to engage in informal formative approach and the effects are hard to tease apart from formal formative approach. So, by eliminating teachers in this experiment we were able to eliminate informal formative approach effects, which could confound the effects of formal formative approach. Although eliminating the teachers seems far-fetched from reality, the aim here was to identify the effects of the components which in turn will help in channelizing teachers' efforts and practices.

Fourthly, by randomizing the participants across the three formative approach conditions and across the six seminar groups with different time intervals we established external validity. The findings of this study can be generalized to the student population.

Finally, by integrating this experiment in an actual learning context this study has established ecological validity and the findings can be generalized to different learning scenarios.

Although the online experiment enabled a clear picture of the effects of the formative approach components, there are real challenges involved in applying this approach in educational contexts. These challenges remain unaddressed in the current experiment. To address the practical challenges, it would be effective to involve the teachers in planning and designing the formative approach. When teachers are involved only at the implementation stage there is a risk of leaving out the pragmatic aspects of the intervention which the teachers are better aware of. The ultimate goal is that the teachers are able to integrate formative approach in their practice to an extent that it enhances both learning as well teaching.

3.11 Implications

This study clearly highlighted the differential effects of the formative approach components on learning and motivation. With every additional formative approach component (DA; DA+CF; DA+CF+AI), there is an incremental effect on learning. For DA+CF and DA+CF+AI the effects on learning are long-term. With regard to motivation, only DA+CF had an initial significant effect, followed by a downward trend across formative approach components and across time. These results have implications for the theoretical model as well as the practical implementation of the formative approach. The results confirm the theoretical model of the formative approach.

With regard to practical implications for classroom teaching, the implementation of the formative approach components plays a crucial role. For an effective implementation of formative approach in learning environments, teachers need to be trained in formative approach. All the more helpful is, when teachers themselves design the formative approach intervention. In case it is not

practical, teachers should be involved in designing the intervention as they are the ones who are finally implementing it and are aware of the practicality aspects. Additionally, an ongoing support system for teachers would be helpful in addressing the challenges faced by teachers while implementing formative approach.

3.12 Future Research

Future experiments could retain the online format and investigate the effects of different types of feedback and different types of instructions as the operationalization in the current experiment was limited to their simplest forms. The effects of specific combinations of feedback and instruction would be worthwhile to investigate, as the two constructs are often confused and misunderstood. Also considering the effects of the learners' stage of development and level of competency in the investigation would be of additional advantage, as they were found to be the most relevant moderators and uniquely moderating the effects of the formative approach on learning in the meta-analysis of this dissertation.

The current experiment investigated the effect of formative approach components on the motivational construct of *learning goal orientation*. Future experiments could investigate the effects on other motivational constructs. The formative approach components were operationalized in their simplest forms in the current experiment. Future studies could explore the effects of more complex forms of the components on motivation.

Future research could try to replicate the current findings in a field experiment. It is neither mandatory nor pragmatic to realize the formative approach components in their purest forms in classrooms or any other learning environment. But replicating the current experiment in a field study would shed light on the practical aspects and challenges in the implementation of formative

approach. These findings can serve to plan professional development programs for teachers, support them in ways that are necessary and thereby enhance the quality of implementation.

3.13 Conclusions

After confirming the theoretical model of the formative approach and finding evidence for the model by means of a meta-analysis, we wanted to identify the effects of the individual components of our model. Therefore, we conducted an online experiment to investigate the effects of the formative approach components on learning and motivation. In this study, we answered the following research questions: Do the formative approach components diagnostic assessment (DA), comprehensive feedback (DA+CF) and adaptive instructions (DA+CF+AI) have differential impact on

(1) near-transfer learning in the a) short-term and b) in the long-term?

(2) motivation in the a) short-term and b) in the long-term?

The findings of this study clearly illustrated the differential effects of the formative approach components on near-transfer learning both short-term as well long-term on the multiple-choice questions. The formative approach components showed a cumulative effect on learning. The effects were such that the adaptive instruction (DA+CF+AI) had the highest effect on learning, followed by the comprehensive feedback (DA+CF) and finally followed by the diagnostic assessment (DA) component with minimal effects. Nevertheless, the learning effects of the formative approach could not be observed in the open-ended question.

With regard to motivation, the scores declined over time across all three groups. The scores increased significantly for the comprehensive feedback (DA+CF) group in the pre-test immediately after they received the feedback.

The major outcome of this study is that, the differential effects of the formative approach components on learning confirm the theoretical model of the formative approach.

General Discussion

In this dissertation, the following were accomplished:

a) conceptualizing formative approach and arriving at a theoretical model of the approach
b) confirming the model of formative approach by means of a systematic review
c) gathering empirical evidence for the model of formative approach by means of a meta-analysis
d) investigating the effects of the formative approach components experimentally

In this section, I draw conclusions from the entire dissertation, discuss the implications and limitations, and recommend directions for future research.

Conclusions

In the first chapter of this dissertation, we postulated an all-inclusive definition and theoretical model of the formative approach. Based on the existing literature and a conceptual analysis, we identified the individual components of the formative approach and arrived at a clear definition and a comprehensive model. We defined formative approach as practices of diagnosing students' learning progress, providing the students with comprehensive feedback, and adapting instructions based on the diagnostic outcomes. Our theoretical model of formative approach comprised the components, diagnostic assessment, comprehensive feedback and adaptive instruction. As a next step, we pursued the goal of confirming the model. This was accomplished by a systematic review of the literature on formative approach. Two independent researchers coded the final pool of 117 records and arrived at 92% agreement on the formative approach components in the primary studies and the constellations in which they prevailed. As a result, we have a clear-cut understanding of the specific aspects of the approach. This clarity should enable precise practice of the formative approach and avert ambiguity. Following the formative approach

conceptualization and the theoretical model, the next milestone was to gather empirical evidence for the theoretical model by examining the effectiveness of the formative approach and its components.

The need for empirical evidence on the effectiveness of formative approach and its components called for the meta-analysis, which is the second major part of this dissertation. Cognitive and motivational consequences being the meaningful outcomes for learning scenarios and have been extensively investigated, our meta-analysis was on the effects of the formative approach components on learning and motivation. The major strength of the meta-analysis was the finding of positive effects of the formative approach on learning and motivation. The meta-analysis resulted in significant medium size effects of the formative approach components on learning and motivation which evidenced the components in the theoretical model. Additionally, among all the moderators that were identified, the stage of schooling (high school and tertiary) and the educational needs of students (regular and special needs) were found to be the most relevant moderators and uniquely moderating the effect of formative approach on learning. What remained unaccomplished in the meta-analysis was, whether the formative approach components had differential effects on learning and motivation respectively. This being an important question to answer, led to an experimental study on the effects of the formative approach components on learning and motivation.

The third part of this dissertation comprised the experimental study on formative approach. By means of an online experimental study, we investigated whether the formative approach components have different effects on learning and motivation. As field studies bring along challenges regarding the implementation fidelity, we conducted an online study. The online experiment not only enabled precise implementation of the components but also controlled and

eliminated the confounding variables. The puzzle of differential effects of the formative approach components was finally solved by the online experimental study. The components indeed have differential effects on learning with the adaptive instruction inducing strongest effects on learning followed by the comprehensive feedback and finally the diagnostic assessment component. Consequently, these results confirm the theoretical model of the formative approach that was presented in the first part of this dissertation.

Current Model versus other Formative Approach Models

The goal of the current research was to conceptualize formative approach clearly and determine the effects of the approach on students' learning and motivation. In this section we illustrate how the current model of the formative approach serves its purpose rightly, above and beyond the existing models. To illustrate this precisely, we provide an overview of the existing models and exemplify the strengths and the gaps in each of the conceptualizations.

Cowie and Bell (1999) defined formative approach as a process of enhancing students' learning during the learning by identifying and addressing the gaps in their learning. They classified formative approach as planned and interactive formative approach and elucidated the processes involved in each of the categories. With the model of Cowie and Bell (1999) as the base, Ruiz-Primo and Furtak (2004) developed their model of formative approach wherein they classify the approach as formal and informal formative approach. The above two models focused predominantly on the process of formative approach and shed light on the individual steps and procedures involved in it. The two models compare and contrast the processes involved in planned/ formal versus interactive/ informal formative approach. These models were indeed helpful in terms of understanding and implementing the approach. The missing aspect in these models was the

elements involved in their formative approach steps and procedures, which actually facilitate learning. Both models, although helpful, were limited to procedural aspects.

Sadler's (1989) model of formative approach emphasizes the role of formative approach in developing learners' expertise. In that regard, Sadler (1989) highlights the role of feedback in enhancing learners' expertise and presents three preconditions necessary for effective feedback. The three preconditions are that the learner is aware of the goal, is able to compare current performance with the goal and carry out steps to close the gaps identified. These conditions place the onus primarily on the students with teachers playing more of a facilitating role. This theory is based on the definition of feedback by Ramaprasad (1983), which says *"Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way."* Though this definition of feedback was proposed in the management theory, it can be applied in educational contexts. Sadler (1989) aptly incorporated Ramaprasad's definition of feedback into the model of formative approach and focused on teachers enabling the students fulfill the three conditions. In comparison to Sadler's model, the current model of the formative approach focuses on teacher implemented formative approach and elucidates the aspects that are necessary for these conditions to be accomplished. The diagnostic assessment component is essential to identify the current level of a learner, the comprehensive feedback component helps in recognizing the gap between the current level and the reference level and modifying the strategies, the adaptive instruction component caters to the content aspects of closing the learning gap.

Shute's model of formative approach (2008) focused predominantly on formative feedback. Shute presented a list of dos and don'ts for formative feedback based on the four models of feedback, namely, Kluger and DeNisi (1996), Bangert-Drowns, Kulik, Kulik, and Morgan

(1991), Narciss and Huth (2004), and Mason and Bruning (2001). Shute's guidelines for formative feedback is a wealth of information for the researchers and practitioners of the formative approach. Each of the four models discussed in Shute's article focused on one or more types and forms of feedback. However, those conceptualizations of feedback included instructional aspects, but, under the label of feedback. Throughout the literature, there is a lack of differentiation between feedback and instruction. This is problematic when we are examining the effects of feedback and instructional practices on learning and motivation and making attributions to one or both without discerning them. Having drawn upon these four models and providing concrete guidelines for formative feedback and making a valuable addition to the literature, Shute's model is not spared from shortcomings. The major limitation lies in the ambiguous definitions of feedback and instruction. Although Shute referred to her approach throughout as formative feedback, the approach is not free from content related instructional aspects and therefore the attributions are not clearly for formative feedback.

Black and Wiliam's (2009) theory of formative assessment conceptualized formative approach as comprising of five key strategies that involve teachers, learners and peers. The strategies include clarifying the criteria for success, designing activities to elicit evidence of student learning, providing feedback, activating students as instructional resources and activating students as owners of their learning. Additionally, Black and Wiliam (2009) incorporated the definition of feedback by Ramaprasad (1983) into their model, which emphasizes on three aspects, namely, the reference level which is the learning goal, the current level of the learner and thirdly the gap between the two levels. Based on the third aspect of gap between the two levels, strategies are designed to bridge the gaps. To summarize this model, the authors illustrate the role of teachers, learners and peers amidst the five strategies under the three phases of establishing where the learner

is going, where the learner is currently and how the learner can get there. The strengths of this model are that, it sheds light on all the essential facets of the formative approach, it includes all agents involved in the approach and their respective roles, it breaks down the whole approach into three phases thereby providing an overview. Nevertheless, the model has its limitations. The model is overly focused on processes and strategies which leaves the core aspect of the formative approach unaddressed. Although the authors started out with the aim of bringing together the diverse practices and arriving at a unifying basis, the aim remained unaccomplished in this model.

In comparison to the existing models of the formative approach, the current model is comprehensive, construct oriented and offers a broader scope. So far, all models of the formative approach have been process oriented or strategy oriented. Those approaches have certainly been helpful and at the same time have led to more ambiguity. The current model serves as a paradigm model which unifies all the processes, strategies and practices of the formative approach under the three major constructs, viz. the diagnostic assessment, the comprehensive feedback and the adaptive instruction. This helps preserve an overview amidst the innumerable practices.

Implications

In terms of theoretical implications, our model contributes to a conceptual clarity of the formative approach. In the previous models, there have been attempts to identify the elements and the processes of the formative approach (Black & Wiliam, 2009; Cowie & Bell, 1999; Ruiz-Primo & Furtak, 2004). Despite all the attempts, a clear conceptualization failed which posed a challenge for both theory as well as practice (Bennett, 2011; Black, 2015). In our model, we have included every individual component of the formative approach. After identifying the three components, diagnostic assessment, comprehensive feedback and adaptive instruction, each of those was defined and confirmed empirically. The major advantage of our theoretical model is that, it is

simple, straightforward and complete. As missing elements of the model one could consider the persons involved in implementing the formative approach and the possible outcomes of the approach. However, it is more important to avoid the risk of overgeneralizing the model or losing the overview. Therefore, the current model of the formative approach being simple and direct, serves its purpose aptly.

The meta-analysis of the effects of the formative approach based on our theoretical model makes a significant scientific contribution by being more informative and elucidating the effects of the individual components on learning and motivation. This clarity was lacking in the earlier meta-analyses where the reported effects could only be attributed to a collection of practices, without the awareness of what the collection entailed. Our meta-analytic results confirmed the theoretical distinction of the formative approach components, but a statistical distinction from the effect sizes was not possible. Statistical distinction of the effects of the formative approach components in our theoretical model was accomplished in our online experimental study. These results, in addition to confirming the model, have practical implications.

In terms of practical implications, the model enables precise implementation of the formative approach. With the knowledge of the differential effects of the components, it is possible to channelize the available resources optimally. While the diagnostic assessment had only short-term learning effects, comprehensive feedback and adaptive instruction had learning effects also in the long-term. Here again, it should be noted that the inferences stemming from the diagnostic assessment were central to the long-term effects of comprehensive feedback and adaptive instruction. An efficient use of the diagnostic outcomes to adapt instructions have led to highest learning outcomes for the adaptive instructions component. Therefore, the formative approach components can be appropriately applied for optimal learning effects or to yield the desired

learning effects. Making the teachers aware of these differential learning benefits of the formative approach components and providing them ongoing training and support will help channelize their time and resources more efficiently for students' learning outcomes. Designing professional development programs for teachers keeping in mind the findings of this dissertation will further enrich such training programs.

Limitations

In this research work, we endeavored on a novel technique to test our proposed theoretical model. We ventured into conducting a meta-analysis of the studies on formative approach to collect evidence for the proposed model. This method is exceptional and very few studies have so far employed a meta-analytic approach for confirming their theoretical model (Hattie & Timperley, 2007; Kluger & DeNisi, 1996). A unique aspect of the current research is the statistical comparison of the effect sizes of the formative approach components with the aim of finding whether the theoretically distinct components are also statistically distinct. Nevertheless, innovative methods are accompanied by a set of challenges. In this section, we discuss the limitations of the current study and which of the main goals could not be accomplished.

One of the main goals of this dissertation was, finding meta-analytic evidence for the theoretical model and for differential effects of the formative approach components. The meta-analysis yielded useful results and indeed provided empirical evidence for the formative approach components. The formative approach components had medium size effects on learning and motivation. Typically, in meta-analyses the effect sizes are reported as to which variable a stronger effect has based on the strength of the effect size and not by means of statistical comparison of the effect sizes. As one of our main goals was to find statistical evidence for the theoretically distinct

components, we undertook the method of statistically comparing the effect sizes of the different formative approach components. Even though the impact of the formative approach components was of varying strengths, the effect sizes were not significantly different from each other. Despite the components being theoretically distinct and having different functions on learning, a lack of evidence for statistical distinction was puzzling. Given the large amount of studies on the formative approach, it remained a challenge to disentangle the effects of the different components of the formative approach.

As a further step, we took a closer look at the primary studies, which were in the first place, disparate in their operationalization of the components and their implementation. This provided an explanation for the overlapping effects of the different components, but also led to the initial problem of the lack of clear conceptualization. Different conceptualizations, inadequate understanding, inaccurate operationalization and implementation infidelity of the formative approach components were the main reasons for the overlapping effects. The goal of the current research was to solve the problem of disparate definitions and practices of the formative approach by conceptualizing a comprehensive theoretical model and confirming the model empirically. It turned out, that the existing problems with the conceptualization and implementation of the formative approach were the reasons for the overlapping effects of the different components and the lack of statistical distinction. We ended up running into the issue of circularity.

With an example from the primary studies, we shed light on the problems with the primary studies. In the study by Yin et al. (2008), the authors sought out to investigate the effect of the diagnostic assessment component of the formative approach. However, the experimental group's unplanned implementation of other components and the control group's unintended implementation of the formative approach components resulted in contrarian outcomes. The

problems in the study by Yin et al. (2008) include lack of clear conceptualization, inaccurate operationalization and implementation. Problems in the primary studies led to the problem of circular explanations for the current meta-analysis. Despite approaching every step in the meta-analysis systematically and meticulously coding the primary studies for the formative approach components on clearly defined categories, we couldn't completely avoid the problem of circularity and it posed to be a limitation to our approach.

As such, testing a theoretical model and verifying the model is a challenging task. In this research, we strived to overcome the challenges and accomplish our goal of confirming our theoretical model. To address the diversity and disparity in the conceptualization of the formative approach, we arrived at a theoretical model of the approach. The very same problems of diversity and disparity in the conceptualization proved to be a challenge for finding evidence for the model. Research limitations, when seen in positive light are helpful in guiding future research. The current limitation helped in designing an original study to test the effects of the formative approach components. In an online experimental study, we found evidence for the differential effects of the formative approach components. In the online experiment, we were able to operationalize and implement the formative approach components accurately and prove the statistical distinction of the formative approach components. In actual learning contexts there are confounding variables which were eliminated in the online experiment. This restricts the generalizability of the results and posed to be a limitation of the online experimental study. However, it should be noted that the limitations of the current research are helpful in guiding future research.

To summarize, in order to understand the formative approach, it would be beneficial if we consciously rethink and reevaluate our theories and challenge the outcomes we obtain. Throughout the literature, there are ample arguments for the advantages of the formative approach (Fuchs

& Fuchs, 1986; Gersten et al., 2009; Graham et al., 2015) and numerous attempts to confirm our conviction (Chen, Jiao, & Hu, 2021; Leenknecht et al., 2021). When confronted with disconfirming evidence we are baffled and reluctant to rethink our theories. We quickly move on to designing the next experiment and seldom have we pondered over the contrarian outcomes. Our scientific reasoning would be of high caliber if we exhibit cognitive flexibility in terms of interpreting the outcomes, show confident humility and be willing to modify our theories. In the current study, despite all the limitations, the complications involved in theorizing and investigating the formative approach is the most important learning for scientific processes. Given the constraints, this research project is definitely a leap in the right direction.

Future Research

The current results clearly indicate the differential effects of the formative approach components. Future research can investigate whether the effects of the components are different for different levels of education. Whether the varying skill levels and competencies of students moderate the effects of formative approach on learning would be valuable information for the practice.

In the current experiment, the formative approach components were operationalized in their simplest and purest forms. In future experiments, the components can be operationalized in more complex forms to investigate their effectiveness on learning. As the online experiment facilitates precise implementation, future research could also follow the same format for examining the effectiveness of different forms or varying levels of each of the formative approach components.

Although the experimental study in this dissertation was an online study, which is not a limitation for the aim of this dissertation and in fact enabled clearer outcomes than is possible in field studies, future research could try to replicate the online experimental results in classroom

settings. While the online experiment enables precise implementation and investigation, the advantages of field studies cannot be ignored. Field studies provide the possibility of identifying and addressing the challenges in practice which could have been overlooked in online studies. Therefore, replicating the online study in classrooms would be beneficial in terms of identifying prospects to further enhancing the implementation of formative approach.

To summarize, this dissertation, starting from the theoretical model to the experimental results makes a significant advance to the existing literature. While the precise conceptualization of the formative approach is a valuable addition to the scientific literature, the meta-analytic results and the experimental outcomes confirm the theoretical model and shed light on the effectiveness of the formative approach components and provide guidelines for effective practical implementation of the approach. Having started out in a chaotic mass of literature on formative approach and unclear diverse conceptualizations, the research work in this dissertation has provided a clear conceptualization of the formative approach with strong supporting evidence, thereby contributing significantly to scientific progress.

References

Acuna, E., & Rodriguez, C. (2004). A meta analysis study of outlier detection methods in classification. *Technical Paper, Department of Mathematics, University of Puerto Rico at Mayaguez*, 1–25.

Assink, M., & Wibbelink, C. J. M. (2016). Fitting three-level meta-analytic models in R: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(3), 154–174. https://doi.org/10.20982/tqmp.12.3.p154

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, *64*(6p1), 359. https://doi.org/10.1037/14156-004

Ayala, C. C., Shavelson, R. J., Araceli Ruiz-Primo, M., Brandon, P. R., Yin, Y., Furtak, E. M., . . . Tomita, M. K. (2008). From formal embedded assessments to reflective lessons: The development of formative assessment studies. *Applied Measurement in Education*, *21*(4), 315–334. https://doi.org/10.1080/08957340802347787

Azzi, A. J., Ramnanan, C. J., Smith, J., Dionne, É., & Jalali, A. (2015). To quiz or not to quiz: Formative tests help detect students at risk of failing the clinical anatomy course. *Anatomical Sciences Education*, *8*(5), 413–420. https://doi.org/10.1002/ase.1488

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191. https://doi.org/10.1037/0033-295x.84.2.191

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*(2), 213–238. https://doi.org/10.3102/00346543061002213

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, *85*(5), 536–553. https://doi.org/10.1002/sce.1022

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, *18*(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Bereby-Meyer, Y., Meyer, J., & Flascher, O. M. (2002). Prospect theory analysis of guessing in multiple choice tests. *Journal of Behavioral Decision Making*, *15*(4), 313–327. https://doi.org/10.1002/bdm.417

Bilotta, D. E. (2012). *Using formative assessment to help nurses learn about limited health literacy*. University of Northern Colorado.

Black, P. (2015). Formative assessment–an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 161–177. https://doi.org/10.1080/0969594X.2014.999643

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (Formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 5. https://doi.org/10.1007/s11092-008-9068-5

Blank, R. K., & Las Alas, N. de (2009). The Effects of Teacher Professional Development on Gains in Student Achievement: How Meta Analysis Provides Scientific Evidence Useful to Education Leaders.

Bloom (1968). Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. *Evaluation Comment*, *1*(2), n2.

Bloom, Hastings, J. T., & Maduas, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, *54*(2), 199–231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. John Wiley & Sons.

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2011). Classroom emotional climate, teacher affiliation, and student conduct. *The Journal of Classroom Interaction*, 27–36.

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yin, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, *31*(4), 13–17. https://doi.org/10.1111/j.1745-3992.2012.00251.x

Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). Assessment for learning: 10 principles. *Cambridge, UK: University of Cambridge School of Education*.

Brookhart, S. M., Moss, C. M., & Long, B. A. (2009). Promoting student ownership of learning through high-impact formative assessment practices. *Journal of MultiDisciplinary Evaluation*, *6*(12), 52–67.

Bull, J., & McKenna, C. (2004). *Blueprint for computer-assisted assessment*. Psychology Press.

Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, *21*(5), 279–283. https://doi.org/10.1177/0963721412452728

Chen, Z., Jiao, J., & Hu, K. (2021). Formative assessment as an online instruction intervention: Student engagement, outcomes, and perceptions. *International Journal of Distance Education Technologies (IJDET)*, *19*(1), 50–65. https://doi.org/10.4018/IJDET.20210101.oa1

Clarke, S. (2001). *Unlocking formative assessment: Practical strategies for enhancing pupils' learning in the primary classroom*. Hodder & Stoughton Educational.

Cornelius, K. E. (2013). Formative assessment made easy: Templates for collecting daily data in inclusive classrooms. *Teaching Exceptional Children*, *45*(5), 14–21. https://doi.org/10.1177/0040059914553204

Cowie, B., & Bell, B. (1999). A model of formative assessment in science education. *Assessment in Education: Principles, Policy & Practice*, *6*(1), 101–116. https://doi.org/10.1080/09695949993026

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, *58*(4), 438–481. https://doi.org/10.3102/00346543058004438

Cumming, G. (2009). Inference by eye: reading the overlap of independent confidence intervals. *Statistics in Medicine*, *28*(2), 205–220. https://doi.org/10.1002/sim.3471

Delandshere, G. (Ed.) (1990). *The International Encyclopedia of Educational Evaluation*. *The International Encyclopedia of Educational Evaluation: Diagnostic Assessment Procedures*. Oxford: Pergamon Press.

Delgado, A., & Prieto, G. (1998). Further Evidence Favoring Three-Option Items in Multiple-Choice Tests. *European Journal of Psychological Assessment*, *14*, 197–201. https://doi.org/10.1027/1015-5759.14.3.197

Department of Education and Science (2004). *Education of Persons with Special Educational Needs Act*. The Stationary Office. Dublin: The Stationary Office. Dublin.

Dreyfus, S. E., & Dreyfus, H. L. (1980). A five-stage model of the mental activities involved in directed skill acquisition. *University of California, Berkeley.* Advance online publication. https://doi.org/10.21236/ada084551

D'Sa, J. L., & Visbal-Dionaldo, M. L. (2017). Analysis of multiple choice questions: item difficulty, discrimination index and distractor efficiency. *International Journal of Nursing Education*, *9*(3), 109–114. https://doi.org/10.5958/0974-9357.2017.00079.4

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341x.2000.00455.x

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*(10), 1040. https://doi.org/10.1037/0003-066x.41.10.1040

Dweck, C. S., & Elliot, E. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, *54*(1), 5–12. https://doi.org/10.1037/0022-3514.54.1.5

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. London: Sage. https://doi.org/10.1024/1012-5302/a000397

Flanagan, J. C. (1969). Program for learning in accordance with needs. *Psychology in the Schools*, *6*(2), 133–136. https://doi.org/10.1002/1520-6807(196904)6:2<133::aid-pits2310060205>3.0.co;2-o

Förster, N., & Souvignier, E. (2014). Learning progress assessment and goal setting: Effects on reading achievement, reading motivation and reading self-concept. *Learning and Instruction*, *32*, 91–100. https://doi.org/10.1016/j.learninstruc.2014.02.002

Fuchs, L. S., & Fuchs, D. (1986). Effects of Systematic Formative Evaluation: A Meta-Analysis. *Exceptional Children*, *53*(3), 199–208. https://doi.org/10.1177/001440298605300301

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R. J., & Yin, Y. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. *Applied Measurement in Education*, *21*(4), 360–389. https://doi.org/10.1080/08957340802347852

Gamer, M., Lemon, J., Robinson, A., & Kendall's, W. (2012). Package 'irr'. *Various Coefficients of Interrater Reliability and Agreement. R Package Version 0.84.1*.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*(3), 1202–1242. https://doi.org/10.3102/0034654309334431

Gipps, C., & Lawton, D. (1990). *Assessment: A teacher's guide to the issues*. Hodder & Stoughton London.

Glaser, R., & Nitko, A. J. (1970). Measurement in Learning and Instruction.

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, *115*(4), 523–547. https://doi.org/10.1086/681947

Griffin, C. P., & Murtagh, L. (2015). Increasing the sight vocabulary and reading fluency of children requiring reading support: The use of a Precision Teaching approach. *Educational Psychology in Practice*, *31*(2), 186–209. https://doi.org/10.1080/02667363.2015.1022818

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, *15*(3), 309–333. https://doi.org/10.1207/s15324818ame1503_5

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hebbecker, K., & Souvignier, E. (2018). Formatives Assessment im Leseunterricht der Grundschule–Implementation und Wirksamkeit eines modularen, materialgestützten Konzepts. *Zeitschrift Für Erziehungswissenschaft*, *21*(4), 735–765. https://doi.org/10.1007/s11618-018-0834-y

Hedges, L., & Olkin, I. (2014). *Statistical methods for meta-analysis*. Academic press.

Herman, J., Osmundson, E., Dai, Y., Ringstaff, C., & Timms, M. (2015). Investigating the dynamics of formative assessment: Relationships between teacher knowledge, assessment

practice and learning. *Assessment in Education: Principles, Policy & Practice*, *22*(3), 344–367. https://doi.org/10.1080/0969594x.2015.1006521

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge. https://doi.org/10.4324/9781315650982

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods*, *11*(2), 193–206. https://doi.org/10.1037/1082-989x.11.2.193

Huhta, A. (2008). Diagnostic and Formative Assessment. In B. Spolsky & F. M. Hult (Eds.), *The Handbook of Educational Linguistics* (pp. 469–482). Oxford: Blackwell Publishing.

Hung, H.-T., Chiu, Y.-C. J., & Yeh, H.-C. (2013). Multimodal assessment of and for learning: A theory-driven design rubric. *British Journal of Educational Technology*, *44*(3), 400–409. https://doi.org/10.1111/j.1467-8535.2012.01337

Hwang, G.-J., & Chang, H.-F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. *Computers & Education*, *56*(4), 1023–1031. https://doi.org/10.1016/j.compedu.2010.12.002

Jonassen, D. H., & Grabowski, B. L. (2012). *Handbook of individual differences, learning, and instruction*. Routledge.

Kibble, J. D., Johnson, T. R., Khalil, M. K., Nelson, L. D., Riggs, G. H., Borrero, J. L., & Payer, A. F. (2011). Insights gained from the analysis of performance and participation in online formative assessment. *Teaching and Learning in Medicine*, *23*(2), 125–129. https://doi.org/10.1080/10401334.2011.561687

Kickmeier-Rust, M. D., Hillemann, E.-C., & Albert, D. (2014). Gamification and smart feedback: Experiences with a primary school level math app. *International Journal of Game-Based Learning (IJGBL)*, *4*(3), 35–46. https://doi.org/10.4018/ijgbl.2014070104

Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, *30*(4), 28–37. https://doi.org/10.1111/j.1745-3992.2011.00220.x

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254–284. https://doi.org/10.1037/0033-2909.119.2.254

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, *1*(4), 279–308. https://doi.org/10.1007/bf01320096

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research*, *86*(2), 602–640. https://doi.org/10.3102/0034654315617832

Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students' motivation. *Assessment & Evaluation in Higher Education*, *46*(2), 236–255. https://doi.org/10.1080/02602938.2020.1765228

Lipnevich, A. A., & Smith, J. K. (2009). Effects of differential feedback on students' examination performance. *Journal of Experimental Psychology: Applied*, *15*(4), 319–333. https://doi.org/10.1037/a0017841

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Livingston, C., & Borko, H. (1989). Expert-novice differences in teaching: A cognitive analysis and implications for teacher education. *Journal of Teacher Education*, *40*(4), 36–42. https://doi.org/10.1177/002248718904000407

Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied, 14*(4), 329–339. https://doi.org/10.1037/a0013835

McMillan, J. H., Venable, J. C., & Varier, D. (2013). Studies of the Effect of Formative Assessment on Student Achievement: So Much More is Needed. *Practical Assessment, Research & Evaluation*, *18*. https://doi.org/10.7275/tmwm-7792

McNeely, S. A., & Cummings, H. H. (1953). TWO BIG JOBS—BETTER HEALTH AND BASIC EDUCATION. *Journal of School Health*, *23*(5), 149–153. https://doi.org/10.1111/j.1746-1561.1953.tb00143.x

McNulty, J. A., Espiritu, B. R., Hoyt, A. E., Ensminger, D. C., & Chandrasekhar, A. J. (2015). Associations between formative practice quizzes and summative examination outcomes in a medical anatomy course. *Anatomical Sciences Education*, *8*(1), 37–44. https://doi.org/10.1002/ase.1442

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *110*(1), 40. https://doi.org/10.1037/0021-843x.110.1.40

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, *11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, *22*(2), 155–175. https://doi.org/10.1207/s15326985ep2202_4

Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, *91*(3), 328–346. https://doi.org/10.1037/0033-295x.91.3.328

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational Statistics*, *8*(2), 157–159. https://doi.org/10.2307/1164923

Ozogul, G., & Sullivan, H. (2009). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development*, *57*(3), 393–410. https://doi.org/10.1007/s11423-007-9052-7

Paiva, R. C., Ferreira, M. S., & Frade, M. M. (2017). Intelligent tutorial system based on personalized system of instruction to teach or remind mathematical concepts. *Journal of Computer Assisted Learning*, *33*(4), 370–381. https://doi.org/10.1111/jcal.12186

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*(7), 710. https://doi.org/10.1037/bul0000151

Parikh, S., Sai, A. B., Nema, P., & Khapra, M. M. (2018). ElimiNet: A Model for Eliminating Options for Reading Comprehension with Multiple Choice Questions. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (*. Retrieved from https://arxiv.org/pdf/1904.02651

Perkins, D. N., & Salomon, G. (1992). Transfer of learning. *International Encyclopedia of Education*, *2*, 6452–6457.

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell Publishing. https://doi.org/10.1080/14733140600986250

Ponce, H. R., Mayer, R. E., Figueroa, V. A., & López, M. J. (2018). Interactive highlighting for just-in-time formative assessment during whole-class instruction: effects on vocabulary

learning and reading comprehension. *Interactive Learning Environments*, *26*(1), 42–60. https://doi.org/10.1080/10494820.2017.1282878

R Core Team (2013). R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing, Vienna, Austria.Vienna, Austria: R Foundation for Statistical Computing, Vienna, Austria.Vienna, Austria. Retrieved from http://www.R-project.org/

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*(1), 4–13. https://doi.org/10.1002/bs.3830280103

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, *24*(2), 3–13. https://doi.org/10.1111/j.1745-3992.2005.00006.x

Roschelle, J., Rafanan, K., Bhanot, R., Estrella, G., Penuel, B., Nussbaum, M., & Claro, S. (2010). Scaffolding group explanation and feedback with handheld technology: Impact on students' mathematics learning. *Educational Technology Research and Development*, *58*(4), 399–419. https://doi.org/10.1007/s11423-009-9142-9

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (Eds.) (2005). *Publication bias in meta-analysis: prevention, assessment and adjustment: The Funnel Plot*. West Sussex, England: John Wiley & Sons. https://doi.org/10.1002/0470870168.ch1

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2006). *Publication bias in meta-analysis: Prevention, assessment and adjustments*. John Wiley & Sons.

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, *37*(1), 15–24. https://doi.org/10.1016/j.stueduc.2011.04.003

Ruiz-Primo, M. A., & Furtak, E. M. (2004). Informal Formative Assessment of Students' Understanding of Scientific Inquiry. CSE Report 639. *Center for Research on Evaluation Standards and Student Testing CRESST*.

Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, *11*(3-4), 237–263. https://doi.org/10.1080/10627197.2006.9652991

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, *44*(1), 57–84. https://doi.org/10.1002/tea.20163

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

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/bf00117714

Salomon, G., & Globerson, T. (1987). Skill may not be enough: The role of mindfulness in learning and transfer. *International Journal of Educational Research*, *11*(6), 623–637. https://doi.org/10.1016/0883-0355(87)90006-1

Scriven, M. S. (1967). *The Methodology of Evaluation in ed. Stokes RE, Curriculum Evaluation, American Educational Research Association Monograph Series on Evaluation No. 1*. Chicago: Rand McNally: Chicago: Rand McNally.

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., . . . Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295–314. https://doi.org/10.1080/08957340802347647

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 4–14.

Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, *28*(3), 32–37.

Shih, J.-L., Ku, D. T., & Hung, S.-H. (2013). Enhancing the learning achievements and attitudes of taiwan vocational school students in accounting with the dynamic assessment system. *Journal of Educational Computing Research*, *49*(3), 345–362. https://doi.org/10.2190/ec.49.3.d

Shute, V. J. (2008a). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189.

Shute, V. J. (2008b). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Siegler, R. S. (1998). *Emerging minds: The process of change in children's thinking*. Oxford University Press.

Simmons, D. C., Kim, M., Kwok, O.-m., Coyne, M. D., Simmons, L. E., Oslund, E., . . . Rawlinson, D.'A. (2015). Examining the effects of linking student performance and progression in a tier 2 kindergarten reading intervention. *Journal of Learning Disabilities*, *48*(3), 255–270. https://doi.org/10.1177/0022219413497097

Spector, T. D., & Thompson, S. G. (1991). The potential and limitations of meta-analysis. *Journal of Epidemiology and Community Health*, *45*(2), 89. https://doi.org/10.1136/jech.45.2.89

Spinath, B. (2012). *SELLMO: Skalen zur Erfassung der Lern-und Leistungsmotivation*. Hogrefe.

Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

Tarrant, M., & Ware, J. (2010). A comparison of the psychometric properties of three-and four-option multiple-choice questions in nursing assessments. *Nurse Education Today*, *30*(6), 539–543. https://doi.org/10.1016/j.nedt.2009.11.002

Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: a descriptive analysis. *BMC Medical Education*, *9*(1), 40. https://doi.org/10.1186/1472-6920-9-40

Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational Psychology Review*, *26*(2), 265–283. https://doi.org/10.1007/s10648-013-9243-1

Tunstall, P., & Gipps, C. (1996). 'How does your teacher help you to make your work better?'Children's understanding of formative assessment. *The Curriculum Journal*, *7*(2), 185–203. https://doi.org/10.1080/0958517960070205

Van den Berg, M., Harskamp, E. G., & Suhre, C. J.M. (2016). Developing classroom formative assessment in Dutch primary mathematics education. *Educational Studies*, *42*(4), 305–322. https://doi.org/10.1080/03055698.2016.1193475

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294. https://doi.org/10.3758/s13428-014-0527-2

VandeWalle, D., Cron, W. L., & Slocum Jr, J. W. (2001). The role of goal orientation following performance feedback. *Journal of Applied Psychology*, *86*(4), 629. https://doi.org/10.1037//0021-9010.86.4.629

Vaus, D. de (2001). *Research design in social research*. Sage.

Vegada, B., Shukla, A., Khilnani, A., Charan, J., & Desai, C. (2016). Comparison between three option, four option and five option multiple choice question tests for quality parameters: A randomized study. *Indian Journal of Pharmacology*, *48*(5), 571. https://doi.org/10.4103/0253-7613.190757

Viechtbauer, W. (2007). Accounting for heterogeneity via random-effects models and moderator analyses in meta-analysis. *Zeitschrift Für Psychologie/Journal of Psychology*, *215*(2), 104–121. https://doi.org/10.1027/0044-3409.215.2.104

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Wehmeyer, M. L. (1999). A functional model of self-determination: Describing development and implementing instruction. *Focus on Autism and Other Developmental Disabilities*, *14*(1), 53–61. https://doi.org/10.1177/108835769901400107

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*, *92*(4), 548–573. https://doi.org/10.1037/0033-295x.92.4.548

Wesson, C. J. (2013). Introducing Patchwork Assessment to a Social Psychology Module: The Utility of Feedback. *Psychology Teaching Review*, *19*(2), 97–105.

Weston, C., McAlpine, L., & Bordonaro, T. (1995). A model for understanding formative evaluation in instructional design. *Educational Technology Research and Development*, *43*(3), 29–48. https://doi.org/10.1007/bf02300454

Weurlander, M., Söderberg, M., Scheja, M., Hult, H., & Wernerson, A. (2012). Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assessment & Evaluation in Higher Education*, *37*(6), 747–760. https://doi.org/10.1080/02602938.2011.572153

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1016/j.jclinepi.2013.03.017

Wiliam, D. (2000). Formative Assessment in Mathematics Part 3: The Learner's Role. *Equals: Mathematics and Special Educational Needs*, *6*(1), 19–22.

Wiliam, D. (2011). *Embedded formative assessment*. Solution Tree Press.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, *11*(1), 49–65. https://doi.org/10.1080/0969594042000208994

Wiliam, D., & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? In. Routledge.

Wongwatkit, C., Srisawasdi, N., Hwang, G.-J., & Panjaburee, P. (2017). Influence of an integrated learning diagnosis and formative assessment-based personalized web learning approach on students learning performances and perceptions. *Interactive Learning Environments*, *25*(7), 889–903. https://doi.org/10.1080/10494820.2016.1224255

Yin, Y., Shavelson, R. J., Ayala, C. C., Ruiz-Primo, M. A., Brandon, P. R., Furtak, E. M., . . . Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, *21*(4), 335–359. https://doi.org/10.1080/08957340802347845

Young, J. E. J., & Jackman, M. G.-A. (2014). Formative assessment in the Grenadian lower secondary school: Teachers' perceptions, attitudes and practices. *Assessment in Education: Principles, Policy & Practice*, *21*(4), 398–411. https://doi.org/10.1080/0969594X.2014.919248

Appendix A: A List of all the Studies included in the Systematic Review and Meta-Analysis

Studies included in the meta-analysis for learning outcomes

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Ali, Sultana & Marwat (2010)	J	60	1	DA+AI	0.8	Sec	Reg	-	-	F	other	grp	facetoface
Anderson (2017)	D	69	4	DA	-0.12 - 0.18	Elm	Reg	-	-	F	math & read	grp	comp
Andersson & Palm (2017)	J	45	2	DA+CF+AI	0.19 - 0.33	Elm	Reg	yes	-	F	math	grp	facetoface
Attali (2015)	J	197 - 215	6	DA+CF	-0.04 - 0.34	none	-	-	-	F	math	ind	comp
Azzi et al. (2015)	J	9 - 164	8	DA+CF+AI	-0.47 - 2.28	Ter	Reg & below avg	no	-	F	med	grp	facetoface
Belcher (2016)	D	337	1	DA+CF+AI	-0.03	Sec	Reg	yes	-	F	math	grp	facetoface
Bijol et al. (2015)	J	161	2	DA+CF	0.47 - 0.69	Ter	Reg	-	-	F	other	ind	comp
Bilotta (2014)	D	78 - 104	5	DA+AI	0.19 - 0.62	nurses	Reg	yes	-	F	other	grp	facetoface

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Motivation Construct	Stage of Schooling	Ed. Needs	Teacher Training	Teacher Exp	Sub	Ind or Grp	Mode
Boulet, Simard & Demelo (2012)	J	50 - 57	2	DA+CF	-0.02 - 0.19	Ter	Reg	no	-	F	other	ind	paper pencil
Bowens (2016)	D	248 - 1007	4	DA+CF+AI	-0.36 - 1.13	Sec	Reg	yes	-	F	math	grp	facetoface
Boyd (2018)	D	16	1	DA	0.63	Sec	Reg	-	-	F	math	ind	paper pencil
Brinkman (2016)	D	50	1	DA+CF	-0.1	Sec	Reg	-	-	F	math	ind	facetoface
Brookhart, Moss & Long (2010)	J	116 - 151	2	DA+CF+AI	0.07 - 0.43	KG & Elm	SpEd	yes	-	F	read	ind	facetoface
Butler (2017)	D	19 - 39	8	DA+AI	-1.52 - 1.26	Sec	Reg	yes	-	F	math	grp	facetoface
Campbell (2013)	D	41 - 45	4	DA+CF	-0.04 - 0.56	Ter	Reg	-	-	F	other	ind	comp
Chauncey (2010)	D	99	2	DA+CF+AI	0.53 - 0.77	Sec	Reg	no	-	F	sci	grp	facetoface
Chen (2016)	D	40 - 79	10	DA+CF	-0.15 - 0.84	Sec	Reg	yes	low	Inf	sci	grp	facetoface

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Chu (2014)	J	64	1	DA+AI	-0.83	Elm	Reg	-	-	F	other	ind	comp
Comes (2016)	D	60	3	DA	0.20 - 0.50	Sec	Reg	no	high	Inf	sci	grp	paper pencil
Conejo et al. (2016)	J	16 - 40	4	DA+CF+AI	0.15 - 0.78	Ter	Reg & below avg	no	-	F	sci	ind	facetoface
Cotos (2011)	D	16 - 105	9	DA+CF+AI	0.67 - 2.81	Ter	Reg	-	-	F	wri	ind	comp
DeLeon (2015)	D	386	2	DA+CF+AI	0.17 - 0.39	HSc	Reg	no	-	F	other	grp	facetoface
**Dibbs (2015)	D	31 - 38	6	DA+CF+AI	-0.62 - 13.07	Ter	Reg	no	-	Inf	math	grp	facetoface
Dorsey (2017)	D	239	3	DA+AI	0.65 - 1.15	Elm	below avg	yes	-	F	read	ind	facetoface
Einig (2013)	J	73 - 94	2	DA+CF	0.48 - 1.27	Ter	Reg	-	-	F	other	ind	comp
Emerson (2011)	D	16 - 27	5	DA	0.20 - 1.26	KG & Elm	SpEd	no	-	F	math	ind	paper pencil

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Faber, Luyten & Visscher (2017)	J	1774	1	DA+CF+AI	0.22	Elm	Reg	yes	-	F	math	ind	comp
Finlay et al. (1998)	J	72 - 137	8	DA+CF+AI	0.04 - 0.65	Ter	Reg	yes	-	F	other	grp	facetoface
Fisher, Cavanagh & Bowles (2011)	J	539	3	DA+CF	0.46 - 0.65	Ter	Reg	no	-	F	other	ind	facetoface
Förster & Souvignier (2014)	J	615	1	DA	0.2	Elm	Reg	-	-	F	read	grp	paper pencil
Förster & Souvignier (2015)	J	618 - 669	4	DA; DA+AI	0.71 - 0.94	Elm	Reg	yes	-	F	read	grp	paper pencil
Förster, Kawohl & Souvignier (2018)	J	619	4	DA+AI	-0.54 - 0.30	Elm	Reg	yes	high	F	read	ind	facetoface
Fox (2013)	D	81 - 152	2	DA+CF+AI	-0.04 - 0.49	Sec	Reg	yes	-	F	wri	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Furtak et al. (2016)	J	14 - 109	18	DA+CF+AI	-0.21 - 0.95	HSc	Reg	yes	high	Inf	sci	grp	facetoface
Gibson (2017)	D	9 - 16	6	DA+AI	0.48 - 1.44	Elm	SpEd	yes	-	F	read	grp	facetoface
Griffin & Murtagh (2015)	J	20 - 40	4	DA+CF+AI	0.14 - 0.84	Elm	SpEd	yes	-	F	read	ind	facetoface
Guiterez-Maldonado et al. (2014)	J	42	1	DA+CF+AI	1.14	Ter	Reg	-	-	F	other	ind	comp
*Hancock (2010)	J	22	1	DA+CF	5.02	Ter	Reg	-	-	F	other	-	-
Hauer et al. (2009)	J	225 - 231	6	DA+CF	0.04 - 0.69	Ter	Reg	yes	-	F	med	ind	comp
Hebbecker & Souvignier (2018)	J	296 - 336	6	DA; DA+CF; DA+CF+AI	0.78 - 0.91	Elm	Reg	yes	high	F	read	ind	comp
Henderson (2015)	D	103	2	DA+AI	0.12 - 0.93	Elm	Reg	yes	-	F	read	ind	facetoface

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Herman et al (2015)	J	803 - 818	2	DA	0.21 - 0.22	Elm	Reg	yes	low	F	sci	ind	paper pencil
Herppich et al. (2014)	J	45	2	DA+CF+AI	0.47 - 0.64	Sec	Reg	yes	-	F	sci	ind	facetoface
Hooley & Thorpe (2017)	J	103	1	DA+CF	0.67	HSc	Reg	no	-	F	read	ind	comp
Hooshyar et al. (2016)	J	58	1	DA+AI	0.58	Ter	Reg	-	low	F	other	ind	comp
Hung, Chiu & Yeh (2013)	J	18	1	DA+CF+AI	1.84	Ter	Reg	no	-	F	other	ind	comp
Hwang & Chang (2011)	J	61	1	DA+AI	0.64	Elm	Reg	-	-	F	other	ind	facetoface
Karuza (2015)	D	502 - 802	7	DA+CF+AI	0.23 - 0.72	Sec	Reg	yes	-	F	math	grp	facetoface
Kellogg et al. (2010)	J	37	2	DA+CF	0.03 - 0.29	Ter	Reg	-	-	F	wri	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Kemp-Murray (2016)	D	122 - 184	4	DA	-0.37 - 0.50	Elm	Reg	no	-	F	read	ind	comp
Khan & Iqbal (2011)	J	80	1	DA	0.76	Sec	Reg	no	-	Inf	other	ind	paper pencil
Kibble et al. (2011)	J	41	2	DA	0.72 - 1.13	Ter	Reg	-	-	F	med	ind	comp
Kickmeier-Rust et al. (2014)	J	40	4	DA+CF	0.07 - 0.46	Elm	Reg	-	-	F	math	ind	comp
King (2003)	D	11 - 29	3	DA+CF+AI	-0.82 - 0.61	Elm	Gifted, Reg & SpEd	yes	low	F	sci	grp	facetoface
Klecker (2007)	J	67	1	DA+CF	0.52	Ter	Reg	-	-	F	sci	ind	comp
Ko (2013)	D	55	1	DA+CF	1.61	Ter	Reg	-	-	F	other	grp	comp
Koedinger, McLaughlin & Heffernan (2010)	J	985	2	DA+CF+AI	0.08 - 0.49	Sec	Reg	-	-	F	math	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Koukounas (2017)	D	92	2	DA+AI	-0.11 - 1.17	Ter	Reg	yes	-	F	math	grp	facetoface
Krasne et al. (2006)	J	146	1	DA+CF	0.68	Ter	Reg	no	-	F	med	ind	comp
Lipnevich & Smith (2009)	J	304 - 314	2	DA+CF	0.65 - 0.85	Ter	Reg	no	-	F	wri	ind	comp
Llorens et al (2016)	J	75 - 101	6	DA+CF	0.03 - 0.88	Sec	Reg	-	-	F	read	ind	comp
Lookadoo et al. (2017)	J	64 - 78	6	DA+CF	0.14 - 0.39	Ter	Reg	-	-	F	sci	ind	comp
Lopez (2016)	J	57 - 124	3	DA	0.32 - 1.67	Elm	Reg	-	-	F	read	-	-
Lopuch (2017)	D	49 - 51	6	DA+AI	-0.23 - 0.13	Elm	SpEd	yes	high	F	read	ind	facetoface
Manuel (2016)	D	53	1	DA+CF	3.38	HSc	Reg	yes	-	F	math	ind	comp
McNulty et al. (2014)	J	140	2	DA+CF	-0.12 - 0.22	Ter	Reg	no	-	Inf	other	grp	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Moylan (2009)	D	22	4	DA+AI	-0.17 - 1.77	Ter	Reg	-	-	F	math	ind	facetoface
Munoz (2011)	D	28	2	DA+AI	0.43 - 0.51	Sec	SpEd	yes	-	F	read	grp	facetoface
Okamoto (2016)	D	22	4	DA+AI	-0.42 - 1.14	Sec	Reg	no	-	Inf	math	grp	facetoface
Ozogul & Sullivan (2009)	J	46	2	DA+CF	0.72 - 1.06	Ter	Reg	-	-	F	-	ind	paperpencil
Paiva et al (2017)	J	72	5	DA+CF+AI	0.55 - 1.42	Ter	Reg	-	-	F	math	ind	comp
Palucci (2011)	D	11 - 32	14	DA+AI	-0.25 - 1.07	Elm	SpEd	yes	-	F	math & read	ind	facetoface
Pemberton (2018)	D	83	3	DA+CF	-0.26 - 0.47	Sec	Reg	-	-	F	math	grp	facetoface
Phelan et al. (2011)	J	1475 - 2616	2	DA+AI	0.03 - 0.09	Sec	Reg	yes	-	F	math	grp	facetoface
Plybour (2016)	D	31 - 37	3	DA+CF	0.65 - 1.52	Ter	Reg	yes	-	Inf	sci	ind	paper pencil

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Ponce et al (2017)	J	22 - 26	4	DA+AI	0.62 - 1.97	HSc & Ter	Reg	yes	-	F	other	ind	facetoface
Radford (2015)	D	128	1	DA+CF	0.28	Ter	Reg	-	-	F	other	ind	comp
Reeves et al. (2017)	J	28	4	DA+AI	-0.31 - 1.05	KG	Reg	-	-	F	read	ind	comp
Resendes et al (2015)	J	42	12	DA+CF	-0.24 - 1.40	Elm	Reg	yes	-	F	sci	grp	comp
Robertson & Kingsley (2015)	J	171 - 203	7	DA+CF	0.02 - 0.50	Ter	Reg	no	-	Inf	other	grp	facetoface
*Rodrigues & Oliveira (2014)	J	723	1	DA+CF	3.97	HSc	Reg	-	-	F	other	-	-
Roschelle et al. (2009)	J	52 - 56	3	DA+CF+AI	0.14 - 0.43	Elm	Reg	yes	-	F	math	grp	facetoface
Segler Zender (2013)	D	48 - 191	6	DA+AI	0.70 - 1.96	HSc	Reg	yes	-	Inf	math & read	grp	facetoface
Shih, Ku & Hung (2013)	J	12	3	DA+CF	0.93 - 2.27	HSc	Reg	-	-	F	math	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Shute, Ventura & Kim (2013)	J	47 - 55	3	DA+CF+AI	0.02 - 0.25	Sec	Reg	-	-	F	sci	ind	comp
Simmons et al. (2015)	J	9 - 156	40	DA+CF+AI	-0.62 - 1.26	KG	SpEd	yes	-	Inf	read	ind	facetoface
Son & Rivas (2016)	J	209	2	DA	-0.18 - 0.23	Ter	Reg	no	-	F	other	ind	comp
Souvignier & Förster (2011)	J	144	1	DA	0.45	Elm	SpEd	-	-	F	read	ind	comp
Stevens (2015)	D	549	2	DA+AI	-0.02 - 0.04	Elm	Reg	no	-	Inf	math	grp	facetoface
Stull et al. (2011)	J	96	1	DA+CF+AI	-0.2	Ter	Reg	-	-	F	other	grp	facetoface
Sturges (2010)	D	41 - 45	8	DA+CF	0.28 - 0.53	Ter	Reg	no	-	F	math	ind	paper pencil
Takacs (2010)	D	105	1	DA+AI	0.19	Sec	Reg	yes	high	F	math	ind	facetoface
Terry (2010)	D	297 - 401	6	DA+CF+AI	0.27 - 0.58	Sec	Reg	yes	-	F	math	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Thomas-Browne (2011)	D	20	1	DA+CF+AI	0.61	Ter	Reg	no	-	F	math	grp	facetoface
Timmers et al. (2015)	J	24 - 26	2	DA+CF	0.14 - 0.74	Sec	Reg	-	-	F	other	ind	comp
Todd & McIlroy (2014)	J	33	5	DA+CF	0.16 - 0.69	Ter	Reg	-	-	F	wri	ind	facetoface
Tomita (2008)	D	46 - 51	12	DA+CF+AI	-0.24 - 0.70	Sec	Reg	yes	-	F	sci	grp	facetoface
Topping (2003)	J	294 - 559	2	DA+CF	0.17 - 0.41	-	Reg	yes	-	F	read	ind	comp
Trumpower & Sarwar (2010)	J	24	2	DA+CF+AI	0.53 - 1.26	HSc	Reg	no	-	F	sci	ind	comp
Tuominen (2008)	D	456 - 595	2	DA+CF+AI	-0.01 - 0.07	Sec	Reg	yes	low	F	math	grp	facetoface
Tyler (2010)	D	112 - 134	8	DA+AI	-0.52 - 0.70	KG	Reg	yes	high	F	read	grp	facetoface
Valle (2016)	D	222 - 499	11	DA+CF+AI	-0.95 - 1.00	Elm, Sec & HSc	Reg	yes	-	Inf	other	grp	facetoface

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Van den Berg et al. (2016)	J	55	2	DA+CF+AI	0.09 - 0.33	Elm	Reg	yes	high	F	math	grp	facetoface
Vogelzang & Admiraal (2017)	J	61	2	DA+CF	0.92 - 1.96	HSc	Reg	yes	-	Inf	sci	grp	facetoface
Walters (2013)	D	95	3	DA+CF+AI	-0.18 - 0.40	HSc	Reg	yes	-	Inf	math	grp	facetoface
Wang (2007)	J	172 - 182	2	DA+CF	0.28 - 0.51	Sec	Reg	-	-	F	sci	ind	comp
Wang & Young (2015)	J	16 - 18	2	DA+CF	0.37 - 0.63	Sec & Ter	Reg	-	-	F	other	ind	comp
Wang et al. (2006)	J	313	1	DA+CF	0.26	Sec	Reg	-	-	F	sci	ind	comp
Wesson (2013)	J	121	1	DA+CF	0.79	Ter	Reg	-	-	F	other	-	-
Williams (2010)	D	375	1	DA+AI	0.15	Elm	Reg	yes	-	F	read	grp	facetoface
Wilson, M. (2016)	D	101	1	DA+CF+AI	0.55	Ter	Reg	-	-	F	med	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Stage of Schooling	Ed. Needs	Teacher Training	Teachg Exp	Formal or Informal	Sub	Ind or Grp	Mode
Wilson, R.B. (2010)	D	294 - 683	2	DA+CF+AI	0.78 - 1.00	Sec	Reg	yes	-	Inf	math	grp	facetoface
Wininger (2005)	J	71	1	DA+CF	0.71	Ter	Reg	yes	-	F	other	grp	facetoface
Witmer et al. (2014)	J	115 - 133	5	DA+AI	-0.09 - 0.59	Elm	Reg	yes	high	Inf	read	ind	facetoface
Wongwatkit et al. (2017)	J	30 - 33	2	DA	0.32 - 0.70	Sec	Reg	-	-	F	math	ind	comp
Yin (2008)	J	214 - 254	7	DA	-0.33 - -0.04	Sec	Reg	yes	-	F	sci	ind	paper pencil
Zyburt (2011)	D	134 - 185	9	DA+AI	-0.36 - 1.18	Elm	Reg	yes	high	F	read	grp	facetoface

*outliers – removed from analyses

**outliers – four of six effect sizes were removed

DA = Diagnostic Assessment; CF = Comprehensive Feedback; AI = Adaptive Instruction

J = Journal article; D = Dissertation; KG = Kindergarten; Elm = Elementary; Sec= Secondary; HSc = Higher secondary; Ter = Tertiary; Reg = Regular; bel avg = below average; SpEd = Special Education; high = over ten years teaching experience; low = within five years teaching experience; F = Formal; Inf = Informal; ind = Individual; grp = Group

Studies included in the meta-analysis for motivational outcomes

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Motivation Construct	Stage of Schooling	Ed. Needs	Teacher Training	Teacher Exp	Sub	Ind or Grp	Mode
Campbell (2013)	D	41 - 45	4	DA + CF	-0.33 - 0.99	satisfaction, enjoyment, feeling rewarded	Ter	Reg & bel avg	-	-	other	ind	comp
Faber, Luyten & Visscher (2017)	J	1774	1	DA + CF + AI	0.15	interest, enjoyment, usefulness	Elm	Reg	yes	-	math	ind	comp
Förster & Souvignier (2014)	J	615	2	DA	-0.09 - -0.04	intrinsic & extrinsic	Elm	Reg	-	-	read	grp	paper pencil
Hebbecker & Souvignier (2018)	J	296 - 336	6	DA	-0.48 - -0.08	intrinsic & extrinsic	Elm	Reg	yes	high	read	ind	comp
Hondrich et al. (2018)	J	551	1	DA + CF + AI	0.38	intrinsic motivation	Elm	Reg	yes	high	sci	ind	facetof ace
Hung, Chiu & Yeh (2013)	J	18	8	DA + CF	0.19 - 4.04	helpfulness	Ter	Reg	no	-	other	ind	comp
Hwan & Chang (2011)	J	61	2	DA + AI	0.02 - 0.20	learning interest & attitude	Elm	Reg	-	-	other	ind	faceto face
King (2003)	D	11 - 29	3	DA + CF + AI	-0.75 - 0.19	motivational strategies for learning	Elm	Gifted, Reg & SpEd	yes	low	sci	grp	faceto face

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Motivation Construct	Stage of Schooling	Ed. Needs	Teacher Training	Teacher Exp	Sub	Ind or Grp	Mode
Ko (2013)	D	55	1	DA + CF	0.75	self-efficacy for learning & performance	Ter	Reg	-	-	other	grp	comp
Koukounas (2017)	D	92	2	DA + AI	-0.15 - 0.02	self-efficacy for learning & performance	Ter	Reg	yes	-	math	grp	faceto face
Lipnevich & Smith (2009)	J	304 - 314	2	DA + CF	1.59 - 2.23	perceived helpfulness of computer & instructor feedback	Ter	Reg	no	-	wri	ind	comp
Moylan (2009)	D	22	12	DA + CF; DA + AI; DA + CF + AI	-0.42 - 0.79	self-efficacy, mastery approach & performance approach	Ter	Reg	-	-	math	ind	faceto face
Paiva et al (2017)	J	72	1	DA + CF + AI	0.23	motivation	Ter	Reg	-	-	math	ind	comp
Radford (2015)	D	63	1	DA + CF	-0.01	motivation to work harder	-	Reg	-	-	other	ind	comp
Thomas-Browne (2011)	D	20	1	DA + CF + AI	-0.06	math motivation	Ter	Reg	no	-	math	grp	faceto face
Wongwatkit et al. (2017)	J	30 - 33	2	DA	1.42 - 2.01	perceived usefulness	Sec	Reg	-	-	math	ind	comp

Author (Year)	Pub	*N*	No. of Effect sizes	Formative Approach Components	Hedges' g	Motivation Construct	Stage of Schooling	Ed. Needs	Teacher Training	Teacher Exp	Sub	Ind or Grp	Mode
Yin (2008)	J	254	4	DA	-0.25 - -0.01	task goal, task perception, self-efficacy, interest	Sec	Reg	yes	-	sci	ind	paper pencil

DA = Diagnostic Assessment; CF = Comprehensive Feedback; AI = Adaptive Instruction

J = Journal article; D = Dissertation; KG = Kindergarten; Elm = Elementary; Sec= Secondary; HSc = Higher secondary; Ter = Tertiary; Reg = Regular; bel avg = below average; SpEd = Special Education; ; high = over ten years teaching experience; low = within five years teaching experience; F = Formal; Inf = Informal; ind = Individual; grp = Group

Appendix B: Further Results of the Experimental Study on the Formative Approach

Descriptive Results

Scores of the participants on the multiple-choice-questions – Near-transfer Learning

Formative Approach Component	Pre-test			Post-test			Follow-up-test		
	n	M	SD	n	M	SD	n	M	SD
DA	37	15.49	2.28	36	16.33	1.77	29	16.52	1.45
DA + CF	42	14.31	2.42	38	16.45	2.55	34	16.21	2.17
DA + CF + AI	40	14.05	2.62	40	17.00	1.93	35	16.83	1.77
Total	119	14.59	2.51	114	16.61	2.12	98	16.52	1.84

Scores of the participants on the open-ended question

Formative Approach Component	Pre-test			Post-test			Follow-up-test		
	n	M	SD	n	M	SD	n	M	SD
DA	37	3.19	1.33	36	3.14	1.62	29	2.59	1.70
DA + CF	42	2.83	1.58	38	3.47	1.67	34	2.71	1.59
DA + CF + AI	40	2.90	1.66	40	3.43	2.06	35	2.60	1.52
Total	119	2.97	1.53	114	3.35	1.79	98	2.63	1.58

Motivation of the participants across conditions and across time-points

Formative Approach Component	Pre-test			Post-test			Follow-up-test		
	n	M	SD	n	M	SD	n	M	SD
DA	37	14.16	3.58	36	13.58	3.59	29	14.17	3.62
DA + CF	42	15.67	3.01	38	14.79	3.57	34	13.91	3.30
DA + CF + AI	40	15.20	3.78	40	14.38	3.43	35	14.17	4.37
Total	119	15.04	3.49	114	14.26	3.53	98	14.08	3.77

Cronbach's Alpha of the Motivation Items (Learning Goal Orientation)

	Items	Pre-test		Post-test		Follow-up-test	
	n	n	Alpha	n	Alpha	n	Alpha
Learning Goal Orientation	4	119	.745	114	.826	96	.808

Effect of DA on near-transfer short-term Learning

Paired Samples Statistics

	N	*M*	*SD*	*SE*
Pre-test Score	36	15.44	2.30	.38
Post-test Score	36	16.33	1.77	.30

Paired Samples Correlation

	N	*Correlation*	*Sig.*
Pre-test Score & Post-test Score	36	0.657	.000

Paired Differences

	M	*SD*	*t*	*Sig. (2-tailed)*
Pre-test Score - Post-test Score	-.889	1.75	-3.042	.004

Effect of DA on near-transfer long-term Learning

Paired Samples Statistics

	N	*M*	*SD*	*SE*
Pre-test Score	29	15.90	2.26	0.42
Follow-up-test Score	29	16.52	1.45	0.27

Paired Samples Correlation

	N	*Correlation*	*Sig.*
Pre-test Score& Follow-up test Score	29	.42	.024

Paired Differences

	M	*SD*	*t*	*Sig. (2-tailed)*
Pre-test Score – Follow-up-test Score	-0.62	2.11	1.58	.125

Effect of DA+CF on near-transfer short-term Learning

Paired Samples Statistics

	N	*M*	*SD*	*SE*
Pre-test Score	38	14.29	2.44	.40
Post-test Score	38	16.44	2.55	.41

Paired Samples Correlation

	N	*Correlation*	*Sig.*
Pre-test Score & Post-test Score	38	0.673	.000

Paired Differences

	M	*SD*	*t*	*Sig. (2-tailed)*
Pre-test Score - Post-test Score	-2.157	2.02	-6.584	.000

Effect of DA+CF on near-transfer long-term Learning

Paired Samples Statistics

	N	*M*	*SD*	*SE*
Pre-test Score	34	14.50	2.54	.44
Follow-up-test Score	34	16.21	2.17	.37

Paired Samples Correlation

	N	*Correlation*	*Sig.*
Pre-test Score& Follow-up test Score	34	.503	.002

Paired Differences

	M	*SD*	*t*	*Sig. (2-tailed)*
Pre-test Score – Follow-up-test Score	-1.71	2.37	-4.20	.0002

Effect of DA+CF+AI on near-transfer short-term Learning

Paired Samples Statistics

	N	*M*	*SD*	*SE*
Pre-test Score	40	14.05	2.62	.41
Post-test Score	40	17.00	1.93	.31

Paired Samples Correlation

	N	*Correlation*	*Sig.*
Pre-test Score & Post-test Score	40	.693	.000

Paired Differences

	M	*SD*	*t*	*Sig. (2-tailed)*
Pre-test Score - Post-test Score	-2.95	1.89	-9.85	.000

Effect of DA+CF+AI on near-transfer long-term Learning

Paired Samples Statistics

	N	*M*	*SD*	*SE*
Pre-test Score	35	14.31	2.43	.41
Follow-up-test Score	35	16.83	1.77	.30

Paired Samples Correlation

	N	*Correlation*	*Sig.*
Pre-test Score& Follow-up test Score	35	.428	.010

Paired Differences

	M	*SD*	*t*	*Sig. (2-tailed)*
Pre-test Score – Follow-up-test Score	-2.51	2.32	-6.42	.000

Analysis of Motivation scores at the Pre-test

ANOVA

Motivation_Pre-test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	46.030	2	23.015	1.922	.151
Within Groups	1388.760	116	11.972		
Total	1434.790	118			

Contrast Coefficients

	DA	DA+CF	DA+CF+AI
DA	1	-1	0
DA+CF	1	0	-1
DA+CF+AI	0	1	-1

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
Motivation Pre-test	Assume equal variances	DA	-1.5045	.78014	-1.929	116	.056
		DA+CF	-1.0378	.78922	-1.315	116	.191
		DA+CF+AI	.4667	.76443	.610	116	.543
	Does not assume equal variances	DA	-1.5045	.74950	-2.007	70.726	.049
		DA+CF	-1.0378	.83832	-1.238	74.953	.220
		DA+CF+AI	.4667	.75651	.617	74.508	.539

Appendix C: Test Materials

Multiple-Choice-Questions

1. Welche der folgenden Aussagen beinhaltet eine realistische Vorstellung der Lehrprofession?
   a) Jeder ist zur Schule gegangen, deshalb kann jeder unterrichten.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Um unterrichten zu können, braucht man Fach- und pädagogisches Wissen.
   b) Lehrer mit dem AVEM Profil A können schneller an Burnout erkranken.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Lehrer mit dem AVEM Profil A erleben negative Emotionen. Aber die Lehrkräfte die nie wirklich für ihren Beruf „gebrannt" haben, sind diejenigen die nicht „ausbrennen".
   c) Lehrer können sich die einzelnen SchülerInnen, mit denen sie arbeiten, nicht aussuchen.
      DA + CF – Die von Ihnen gewählte Antwort ist richtig.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Die Schulverwaltung kann Lehrern unterschiedliche Klassen aber nicht einzelne SchülerInnen zuteilen.

2. Welche Herausforderung kann es im Lehrerberuf geben?
   a) Lehrer sollten zufrieden sein, weil Lehren ist die beste Profession für eine optimale Work-Life-Balance.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Work-Life-Balance ist nicht abhängig von einer bestimmten Profession.
   b) Lehrer sind einzig und allein für das Lernen und die Leistung ihrer SchülerInnen verantwortlich.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Lehrer können das Lernen und die Leistung ihrer SchülerInnen nur bis zu einem gewissen Grad beeinflussen.
   c) Lehrer haben nicht nur die Lehraufgaben, sondern ein vielfältiges Aufgabenspektrum.
      DA + CF – Die von Ihnen gewählte Antwort ist richtig.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Die Aufgaben von Lehrern sind nicht immer klar definiert und abgrenzbar.

3. Welcher der folgenden ist der zentrale Mechanismus dahinter, dass Lehrkräfte wütend werden und sanktionierende Maßnahmen einsetzen, wenn SchülerInnen Fehlverhalten zeigen?
   a) Lehrer sehen das Fehlverhalten als vorsätzlich und kontrollierbar an.
   DA + CF – Die von Ihnen gewählte Antwort ist richtig.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Diese Interpretation des Fehlverhaltens der SchülerInnen führt zu Wut und Bestrafung. Das Fehlverhalten hat normalerweise Ursachen, welche identifiziert und angesprochen werden müssen.
   b) Lehrer haben Stress im Privatleben.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Emotionale Erschöpfung kann die Fähigkeit des logischen Denkens und des Reflektierens der Situation beeinflussen.
   c) Lehrer üben gerne Macht über ihre SchülerInnen aus.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Das Bedürfnis nach Machtausübung kann durch das Gefühl der Unzulänglichkeit entstehen.

4. Burnout ist…
   a) vor allem eine Lehrerkrankheit.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Burnout ist nicht spezifisch für einen bestimmten Beruf. Burnout ist allerdings berufsbezogen (nicht auf private Lebensbereiche bezogen) und tritt in vor allem in solchen Berufen auf, in denen Menschen „Gebende" sind (z.B. Heilberufe, Beratung, Schule).
   b) keine klinische Diagnose.
   DA + CF – Die von Ihnen gewählte Antwort ist richtig.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Burnout ist keine Krankheit im Sinne einer medizinischen Definition von DSM-5. In ICD-10 wurde Burnout unter Probleme mit Bezug auf Schwierigkeiten bei der Lebensbewältigung (Z73) aufgelistet.
   c) nicht mit Depression zu vergleichen.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Burnout und Depression haben Überschneidungen. Sie teilen sich die Symptome *Emotionale Erschöpfung* und *Reduzierte Leistungsfähigkeit*.

5. Effiziente Klassenführung…
   a) ist präventiv und umfassend.
   DA + CF – Die von Ihnen gewählte Antwort ist richtig.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Etablierte Regeln und Routinen zu allen Aspekten des Unterrichts verhindern Störungen und dienen der effizienten Klassenführung.
   b) ist frei von Störungen.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Kein Klassenzimmer ist frei von störendem Verhalten. Das Vorausplanen reduziert Unterbrechungen und vereinfacht den Umgang damit.
   c) sanktioniert störendes Verhalten.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Sanktionen sollten immer die letzte Instanz sein. Das Vorausplanen reduziert Unterbrechungen und vereinfacht den Umgang damit.

6. Eine effiziente Klassenführung…
   a) kostet viel Zeit.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Es kostet weniger Zeit als der Umgang mit Störungen.
   b) erhöht die aktive Lernzeit.
   DA + CF – Die von Ihnen gewählte Antwort ist richtig.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Lehrkräfte haben mehr Zeit für das Lehren, weil sie alles im Griff haben.
   c) reduziert die Planungszeit.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Es kostet Zeit für die Planung, weil man alle Unterrichtselemente genau durchdenkt und plant.

7. Der Low Profile Ansatz zielt darauf ab, Störungen frühzeitig so zu deeskalieren, dass der Unterrichtsfluss nicht unterbrochen wird. Welche von den folgenden sind Teilaspekte des Low Profile Ansatzes?
   a) Preparation, Action, Reflection
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Er wird in die Teilaspekte Anticipation (im Auge behalten), Deflection (Unterbinden) und Reaction (Ignorieren/undramatisch eingreifen) untergliedert.
   b) Anticipation, Deflection, Reaction
      DA + CF – Die von Ihnen gewählte Antwort ist richtig.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Er wird in die Teilaspekte Anticipation (im Auge behalten), Deflection (Unterbinden) und Reaction (Ignorieren/undramatisch eingreifen) untergliedert.
   c) Introspection, Reflection, Reaction
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Er wird in die Teilaspekte Anticipation (im Auge behalten), Deflection (Unterbinden) und Reaction (Ignorieren/undramatisch eingreifen) untergliedert.

8. Welche der folgenden Antwortmöglichkeiten lässt sich der Abwehr von störendem Verhalten ohne Unterbrechung des Unterrichts zuordnen?
   a) Die störende Schülerin/den störenden Schüler warnen.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Warnung der Schülerin/des Schülers führt zu Aufmerksamkeit und beeinflusst somit den Unterrichtsfluss.
   b) Die störende Schülerin/den störenden Schüler dazu auffordern, den Klassenraum zu verlassen.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Aufforderung an die Schülerin/den Schüler den Klassensaal zu verlassen führt zu Aufmerksamkeit und beeinflusst somit den Unterrichtsfluss.
   c) Blickkontakt mit der störenden Schülerin/dem störenden Schüler aufnehmen.
      DA + CF – Die von Ihnen gewählte Antwort ist richtig.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Blickkontakt mit dem Störenden kommuniziert die Unangemessenheit des Verhaltens und hilft die Aufmerksamkeit auf den Unterricht zu lenken, ohne diesen zu unterbrechen.

9. Wozu dienen Regeln, Routinen und Rituale im Unterricht?
   a) Um eine Arbeitsatmosphäre für SchülerInnen und Lehrkräfte zu schaffen.
   DA + CF – Die von Ihnen gewählte Antwort ist richtig.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Regeln, Routinen und Rituale ermöglichen ein harmonisches Arbeitsumfeld für alle im Klassensaal durch das Verhindern von Unterbrechungen.
   b) Um SchülerInnen zu identifizieren, die Fehlverhalten zeigen und um ihr Verhalten zu korrigieren.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Regeln, Routinen und Rituale helfen Fehlverhalten vorzubeugen, indem sie SchülerInnen im Voraus über Konsequenzen von Fehlverhalten informieren.
   c) Um die Autorität des Lehrers zu demonstrieren.
   DA + CF – Die von Ihnen gewählte Antwort ist falsch.
   DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Regeln, Routinen und Rituale werden gemeinsam von SchülerInnen und Lehrkräften erarbeitet, sodass SchülerInnen für die Konsequenzen beim Brechen einer Regel verantwortlich sind.

10. Welcher der folgenden Strategien würdest du Regeln, Routinen und Rituale zuordnen?
    a) Präventive Strategien.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Durch das Schaffen von Regeln, Routinen und Ritualen können Lehrkräfte effektiv Unterbrechungen im Unterricht vermeiden.
    b) Reaktive Strategien.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Regeln, Routinen und Rituale werden üblicherweise im Voraus aufgestellt und bei Bedarf angepasst.
    c) Kontrollstrategien.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Regeln, Routinen und Rituale werden gemeinsam von SchülerInnen und Lehrkräften eingeführt um einen ungestörten Unterrichtsfluss zu gewährleisten.

11. Welche der folgenden Aussagen trifft auf Direkte Instruktion (DI) zu?
    a) DI erfolgt kleinschrittig und kontrolliert.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Durch das kleinschrittige und kontrollierte Durchführen wird der Erwerb aller deklarativen Wissenselemente und prozeduralen Regeln sichergestellt.
    b) DI ist nicht für SchülerInnen mit geringer Selbstkontrolle geeignet.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. DI ist gut plan- und steuerbar und somit auch bei SchülerInnen mit geringer Selbstkontrolle geeignet.
    c) DI prüft kein Vorwissen.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. DI prüft Vorwissen immer direkt ab.

12. Wie ist der korrekte Ablauf der direkten Instruktion in der Klasse nach Wiechmann (2009)?
    a) Advance Organizer → lehrergeleitete Präsentation des Unterrichtsgegenstandes→ individuelle Übung → gemeinsame und strukturierte Übung mit der gesamten Klasse→ Transferübung
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Lernenden üben den Unterrichtsgegenstand nicht, bevor sie gemeinsam und strukturiert in der gesamten Klasse üben. Stattdessen üben sie individuell, bevor sie die Transferübung absolvieren.
    b) Advance Organizer → lehrergeleitete Präsentation des Unterrichtsgegenstandes → gemeinsame und strukturierte Übung mit der gesamten Klasse → individuelle Übung → Transferübung
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Nach Wiechmann (2009) sieht die direkte Instruktion vor, dass nach der Präsentation eine gemeinsame Übung stattfindet, gefolgt von einer individuellen Übung, die schließlich in einer Transferübung mündet.
    c) Advance Organizer → Lehrergeleitete Präsentation des Unterrichtsgegenstandes→ gemeinsame und strukturierte Übung mit der gesamten Klasse→ Transferübung → individuelle Übung
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Lernenden üben den Unterrichtsgegenstand nicht, nachdem sie eine Transferübung absolviert haben. Stattdessen üben sie individuell, bevor sie die Transferübung absolvieren.

13. Welche Aussage zu den Qualitätsmerkmalen guten Unterrichts stimmt?
    a) Merkmale des Unterrichts stehen eher weniger mit erwünschten Zielen in Zusammenhang.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Merkmale des Unterrichts stehen mit erwünschten Zielen in Zusammenhang.
    b) Selbst kompetenten Lehrkräften fällt es schwer, Qualitätsmerkmale des Unterrichts auf hohem Niveau zu realisieren.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. - Kompetente Lehrkräfte wissen, wie sie Qualitätsmerkmale des Unterrichts auf hohem Niveau realisieren können und handeln entsprechend.
    c) Manchmal müssen bestimmten Merkmalen Prioritäten gegeben werden.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Es sind nicht immer alle Merkmale in optimaler Weise erfüllbar, gemessen an Zielkriterien müssen ggf. bestimmten Merkmalen Prioritäten gegeben werden.

14. Welche von den folgenden ist die Kernkompetenz der Lehrkräfte?
    a) Einen qualitativ guten Unterricht zu planen und durchzuführen.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Die Kernkompetenz von Lehrkräften ist einen qualitativ guten Unterricht zu planen und durchzuführen und alle anderen Kompetenzen tragen dazu bei.
    b) Die Sicherstellung, dass alle SchülerInnen gute Noten erhalten.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Lehrkräfte können lediglich guten Unterricht, jedoch keine guten Noten sicherstellen.
    c) Leistungsstarke SchülerInnen zu identifizieren und diese weiter zu fördern.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Unterstützung und Aufmerksamkeit der Lehrkraft ist essentiell für SchülerInnen, die keine guten Leistungen zeigen. Lehrkräfte sollten sicherstellen, dass diese nicht abgehängt werden.

15. Welche der folgenden Unterrichtsmerkmale kann nicht auf die Lehrkraft zurückgeführt werden?
    a) Leistungsheterogenität.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Das Vorwissen der SchülerInnen kann nicht vollständig auf die Lehrkraft zurückgeführt werden.
    b) Lernförderliches Klima.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Eine Lehrkraft kann ein lernfördrliches Umfeld begünstigen.
    c) Angebotsvielfalt.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Eine Lehrkraft kann eine breites Spektrum an Methoden einführen um Lerngelegenheiten zu schaffen.

16. Welche der folgenden Aspekte kann durch die Lehrkraft unmittelbar beeinflusst werden?
    a) Minimierung/ Vermeidung der Störungen.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Durch eine effiziente Klassenführung kann die Lehrkraft die Störungen minimieren oder vermeiden.
    b) Intelligenz der SchülerInnen.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Eine Lehrkraft kann das Lernen erleichtern, jedoch nicht die intellektuellen Fähigkeiten der SchülerInnen beeinflussen.
    c) Anstrengungsbereitschaft der SchülerInnen.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Eine Lehrkraft kann SchülerInnen lediglich dazu animieren sich stärker zu bemühen, dies jedoch nicht garantieren.

17. Wie können professionelle Selbstregulationsfähigkeiten Lehrkräften helfen, sich den Herausforderungen ihrer Profession kompetent zu stellen?
    a) Sie helfen Lehrkräften alles selbstständig zu bewältigen.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Lehrkräfte können nicht alles selbstständig bewältigen. Sie sollten lernen, dass sie Hilfe suchen können.
    b) Sie helfen Lehrkräften mit Unsicherheiten umzugehen.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Die Selbstregulationsfähigkeit von Lehrkräften hilft ihnen mit Unsicherheiten umzugehen.
    c) Sie helfen Lehrkräften arbeitsintensiv zu arbeiten.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Selbstregulationsfähigkeit von Lehrkräften hilft ihnen zu erkennen, wann sie aufhören und eine Pause machen sollten oder wann sie Hilfe suchen oder Aufgaben an andere delegieren sollen.

18. Wie hängt effiziente Klassenführung mit dem Stresslevel der Lehrkräfte zusammen?
    a) Je besser die Klassenführung, desto geringer das Stresslevel der Lehrkraft.
    DA + CF – Die von Ihnen gewählte Antwort ist richtig.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Effizientere Klassenführung bedeutet weniger Unterbrechungen, wodurch der Lehrer weniger beansprucht wird und somit weniger Stress empfindet.
    b) Je besser die Klassenführung, desto höher das Stresslevel der Lehrkraft.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Effizientere Klassenführung bedeutet mehr Arbeit des Lehrkraft vor der Schulstunde, jedoch weniger Unterbrechungen innerhalb der Schulstunden und somit weniger Beanspruchung des Lehrkraft.
    c) Klassenführung hängt nicht mit dem Stresslevel der Lehrkraft zusammen.
    DA + CF – Die von Ihnen gewählte Antwort ist falsch.
    DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Schlechte Klassenführung bedeutet häufigere Unterbrechungen und somit eine stärkere Beanspruchung der Lehrkraft.

19. Um eine effektive Instruktion zu gewährleisten, muss eine Lehrkraft…
    - **a) die Inhalte und die Abfolge des Unterrichts im Voraus planen.**
      DA + CF – Die von Ihnen gewählte Antwort ist richtig.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Durch das Planen der Details im Voraus sind Lehrkräfte für eine effektive Instruktion vorbereitet.
    - b) die SchülerInnen über den Inhalt und die zeitliche Abfolge entscheiden lassen.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Auch wenn SchülerInnen von Autonomie profitieren, kann zu viel davon nachteilig für den Lernerfolg sein.
    - c) spontan entscheiden, welche Inhalte und zeitliche Abfolge sie wählt.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Auch wenn Flexibilität im Stundenplan hilfreich sein kann, kann starke Spontanität zu Chaos führen und das Lernen beeinträchtigen.

20. Was ist eine realistische Sicht auf die Lehrprofession?
    - **a) Lehren kann belebend und erschöpfend sein.**
      DA + CF – Die von Ihnen gewählte Antwort ist richtig.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist richtig. Lehren ist aufwändig, aber das Ergebnis kann bereichernd sein.
    - b) Lehrer haben meistens Nachmittags frei.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Die Lehrkräfte sind nachmittags meistens mit der Vor- und Nachbereitung ihres Unterrichts beschäftigt, müssen Hausarbeiten korrigieren oder besuchen Konferenzen und Fortbildungen.
    - c) Lehrkräfte haben den größten Einfluss auf das Leben einer Schülerin / eines Schülers.
      DA + CF – Die von Ihnen gewählte Antwort ist falsch.
      DA + CF + AI – Die von Ihnen gewählte Antwort ist falsch. Es gibt viele wichtige Einflussgrößen im Leben einer Schülerin/ eines Schülers.

Open-ended Question on Classroom Management & Coding Scheme

Erläutern Sie kurz (100 - 150 Wörter), was effiziente Klassenführung bedeutet.

Nr.	Sinneinheit	Punkte
a)	Geordneten Rahmen schaffen	1
b)	Maximierung aktiver Lernzeit	1
c)	Regeln für das Miteinander etablieren und durchsetzen, vorbeugen & planen	1
	Weitergehende Erläuterung: • Verbindliche Vereinbarungen für das Verhalten • Festlegung von Konsequenzen für Verstöße und konformes Verhalten • Gemeinsames Einführen der Regeln mit Schüler/-innen • Verbindlichkeit durch Visualisierung (z.B. Plakate) • Commitment schaffen • Regeln konsequent durchsetzen	1
d)	Routinen	1
	Weitergehende Erläuterung: • Festgelegte Abläufe für wiederkehrende Situationen • Können durch festgelegte Signale unterstützt werden	1
e)	Rituale (festgelegte Abläufe für wiederkehrende Situationen)	1
	Funktion von Ritualen: • Klare Erwartungen für Standardsituationen • Entlastung von Lehrern und Schülern • Erzieherische Funktion (Lehrer als Über-Ich)	1
f)	Umgang mit Störungen (Low Profile Ansatz)	1
	Weitergehende Erläuterung: • „den Ball flach halten" • Frühzeitiges, deeskalierendes Eingreifen • Anticipation – im Auge behalten • Deflection – sparsame, non-verbale Aktion, Nähe zeigen • Reaction – nur oberhalb der Schwelle eingreifen, dann undramatisch	1

Items on Motivation – Learning Goal Orientation

Bei der Beantwortung des Fragebogens war es für mich wichtig,
so viel wie möglich über „Psychologie für den Lehrerberuf" zu lernen.

□ stimmt gar nicht	□ stimmt wenig	□ stimmt teils teils	□ stimmt ziemlich	□ stimmt völlig

Bei der Beantwortung des Fragebogens war es für mich wichtig,
wirklich zu begreifen, was ich durch diese Aufgabenbearbeitung lernen kann.

□ stimmt gar nicht	□ stimmt wenig	□ stimmt teils teils	□ stimmt ziemlich	□ stimmt völlig

Bei der Beantwortung des Fragebogens war es für mich wichtig,
die Aufgabe zu nutzen, um eigenes Verbesserungspotential zu erkennen.

□ stimmt gar nicht	□ stimmt wenig	□ stimmt teils teils	□ stimmt ziemlich	□ stimmt völlig

Bei der Beantwortung des Fragebogens war es für mich wichtig,
Erkenntnisse zu gewinnen, wie ich mein psychologisches Wissen für den Lehrerberuf verbessern kann.

□ stimmt gar nicht	□ stimmt wenig	□ stimmt teils teils	□ stimmt ziemlich	□ stimmt völlig