Curiously Creative: The Relationship Between Teacher Verbal Autonomy Support and Convergent Musical Creativity in Primary School Aged Children

by

Alexandra M. Günzel

(S4531051)

E-mail: a.m.gunzel.1@student.rug.nl

A thesis submitted in fulfillment of the requirements for the degree of

Master of Science

Talent Development and Creativity

Faculty of Behavioural and Social Sciences, University of Groningen, The Netherlands

Major Supervisor: Steenbeek, H. Ph.D.

Daily Supervisor: Hendriks, L. H. H.

Second Reviewer: Cox, R. F. A.

Date: 19-01-2022

Abstract

Teacher autonomy support has been linked to a number of positive outcomes in students, but has not been explored in the context of musical education in primary schools. This thesis analysed the development of teacher verbal autonomy support (ASV), the development of student convergent musical thinking and acting (CTA) and the relation between the two as a part of ongoing Curious Minds research in Dutch primary schools. An increase in both variables an a positive relationship was expected. Eight lessons from six teachers were recorded in total, with a 2x4x2 format. Teachers were coached via Video Interaction Coaching on their autonomy support during the four intervention lessons. Monte Carlo analyses revealed an increase in ASV for all teachers during the intervention and an increase or stabilization of ASV during post-measures for four teachers, and a decrease in ASV for two teachers. Variable patterns were visible for student CTA across conditions. Spearman correlations revealed a weak relationship between ASV and CTA for two out of six teachers and an overall moderate positive correlation between ASV and CTA during post-measures. State Space Grids revealed a positive shift in dyads. The Curious Minds intervention was related to an increase in ASV, yet seemed unrelated to student CTA due to mixed results. The relationship between autonomy support and student convergent musical thinking and acting remains unclear. Future research could focus on a longer intervention phase, including, scaffolding and student divergent creativity measures and controlling for demographic factors and singing lessons.

Keywords: Complex Dynamic Systems Theory, Curious Minds, autonomy support, convergent musical creativity

Contents

Educating Creativity: A Theoretical Framework

Within education, children's creativity has been recognized as an essential 21st century skill (Kupers, 2019). Imagining and creating new, unique solutions to complex problems has been regarded as a distinctive human trait. Today's increasingly complex problems demand the development of sophisticated creative solutions by current and future generations (Kupers, 2019). Therefore, the development of programs where children's creativity can be encouraged and stimulated, could provide promising prospects for talent development and future problem solving.

In line with a history of research underpinning the neurocognitive, physical and emotional benefits of music education on the development of children, also dubbed the '*Mozart Effect'* (Collins, 2013; Jaschke et al., 2018), there has been growing governmental interest for the implementation of music education in Dutch primary schools (Schippers, 1996). Despite the interest and funding, teachers often report not feeling capable enough of teaching music in primary schools (Schippers, 1996). This leaves an enormous gap in the Dutch educational system, which could be filled by programs that empower musically inexperienced and/or insecure primary school teachers to teach music and grow together with students in their understanding and expression in music.

The purpose of this Master thesis is to contribute to creativity research in the developmental field by examining the relation between teacher autonomy support and student convergent musical creativity. These variables have been studied real-time and over time by means of analysing video data from six teachers in four different Dutch primary schools as a part of the second year of an ongoing Curious Minds PhD research program (Hendriks et al., 2018).

This section will explore the definition of the concepts of convergent musical creativity and autonomy support and will explore the theoretical relation between both concepts.

Furthermore, the Curious Minds project will be explained and three research questions this thesis will be focusing on will be discussed.

Creativity Defined

Creativity has often been defined as imagining and/or producing something novel or unexpected, something beyond that which is already known at a certain point (Kupers et al., 2018), therefore, involving thinking and/or acting in an original manner. The underlying mental process for originality has been defined as *divergent thinking* (Webster, 2002). However, solely bursting out new ideas does not equal an interesting or valuable creative contribution. Thus, in order for an idea, product, or behaviour to be considered creative it must combine originality and appropriateness in the context of a particular task or activity (Beghetto, 2016). The underlying process for appropriateness has been defined as *convergent thinking*, which involves evaluating various possibilities and choosing the most appropriate solution (Burnard, 2002).

Thus, creativity can be been defined as a dynamic mental process alternating between divergent and convergent thinking, moving in stages over time, enabled by enacted skills (innate and learned), and by certain conditions, all resulting in a final product (Webster, 1990).

Creativity in Music

Divergent thinking in music has been defined as "the exploration of the many possibilities of music expression involving imaginative thought, which might be melodic, rhythmic, harmonic or longer complex patterns, possibly realized on some musical instrument" (Webster, 2002). According to Webster (1990), *convergent thinking* in music can be defined as "the ability to apply rhythmic and tonal patterns and musical syntax and sensitivity to the musical whole" which can broadly be described as learning and

understanding music as a language. Both forms of musical creativity refer to thought processes in constant interplay, which ultimately lead to a musical product. This product has been defined as (a) composition, (b) performance/improvisation and (c) listening and (d) analysis (Webster, 2002). Therefore, divergent musical creativity can specifically be referred to as *divergent musical thinking and acting (DTA),* whilst convergent musical creativity can be defined as *convergent musical thinking and acting (CTA)*.

The Development of Creativity

Historically, there has been a debate whether creativity stems from innate characteristics or is a concept that can be taught and/or practiced (Kupers, 2019). Even though there has been evidence provided for both sides of the spectrum, recently consensus has been reached over the idea that creativity can be a product of both innate and learned constructs. Therefore, studying which environments can create the optimal conditions for the process of creativity to flourish, has been a prominent focus in current creativity research (Sternberg, 2010).

The Value of Education. Traditional teaching methods have often been argued to kill curiosity and creativity: as children move through the years, there seems to be a growing importance in teaching the curriculum and providing the 'right' answers to the teacher's questions (Sahlberg, 2019). This controlling motivational climate has been thought to lower intrinsic motivation which has shown to foster more superficial and less transferable learning, greater dropout and lower wellness (Ryan & Niemiec, 2009). According to the *Self-Determination Theory* (SDT) by Ryan and Deci (2002), there are three innate psychological needs that need to be met to stimulate intrinsic motivation: *competency*, *autonomy* and *relatedness*. According to the SDT, the main value of education lies in its promise for enhancing human freedom and capabilities by providing an autonomy-supportive learning climate (Ryan & Niemiec, 2009). In line with the STD, the *Complex Dynamic Systems* theory provides a theoretical framework of creativity as both a cognitive construct and a physical

product of a spectrum of characteristics that are in interplay with one's environment and the specific task at hand.

Creativity in Education from a Complex Dynamic Systems Perspective

From a Complex Dynamic Systems (CDS) perspective, it is crucial that the teacher and student co-create each state. Rather than creativity being evoked within the child through the teacher's instructions and/or feedback, creativity is seen as a process that is embedded in dyadic interactions between the child, teacher and the task at hand. Therefore, creative development is a social, self-organizing process, where children learn by doing (Kupers et al., 2019). This process is iterative, as each state of the interaction is dependent on the previous state of the system and serves as direct input for the next state, where an upward spiral can be created (Beghetto, 2006). An upward spiral is characterized by a high level of enthusiasm, knowledge and rapid interactions (Steenbeek & van Geert, 2013).

For this process to spontaneously occur, the teacher should adopt an autonomy supportive, stimulating and curious attitude regarding the child's responses (Kupers et al., 2019). By means of open questions, investigation and scaffolding, the teacher and child will create a new state of knowledge which serves as a prerequisite for the next state (Steenbeek & van Geert, 2013). Scaffolding is an adaptive support technique where the teacher investigates the student's current performance level and poses thought provoking student-centred questions to help the student reach a new level of knowledge on their own (Kupers et al., 2017; van Vondel et al., 2017). Scaffolding and autonomy support are thought to effectively guide students in their learning process by the central notion of agency; where students are seen as active participants in their own learning and development (Kupers et al., 2017).

Autonomy Support Defined

Autonomy has been defined as the need for the regulation of one's own activities and experiences (Ryan & Deci, 2017). In an autonomous-supportive learning context, teachers nurture the inner motivational resources of students by creating a pressure free atmosphere, focused on the student (Deci et al., 1994). More specifically; teachers provide learners with (1) positive feedback, (2) meaningful explanations using noncontrolling language, (3) choice, (4) acknowledgement of their feelings, perspective and initiative, and (5) confidence in their ability (Gagne, 2003). Observation and active listening have been found to be important teaching skills in nurturing autonomy in a responsive and an adaptive way (Green, 2006; Reeve, 2006). Therefore, autonomy support can include a verbal and a non-verbal or behavioural component (Young-Jones et al., 2014).

In line with the main research this thesis is part of (Hendriks et al., 2018), *non-verbal autonomy support* (ASNV) will be defined as: teacher gestures, movement, modelling, postures and facial expressions used to foster autonomous student exploration. *Verbal autonomy support* (ASV), will be defined as teacher verbal expressions used, such as student-centred open questions, to foster autonomous student exploration.

Autonomy Support and Creativity

An autonomy-supportive teaching style has been linked to a range of positive student outcomes, such as intrinsic motivation, engagement and autonomous behaviour in learning (Kupers et al., 2017). Giving students the option to be involved in planning and choosing which tasks and skills they would like to emphasize, has been found to be a strong motivational technique (Guthrie & Davis, 2003; Nolen & Nicholls, 1994). Students have shown a significant increase in engagement in learning when they assume responsibility and therefore, control over the process or product (Roe, 1997; Teel, Debruin-Parecki, &

Covington, 1998), which refers back to the central notion of agency, posed by the CDS theory.

In line with the CDS theory of creativity development, one would assume that there would be a sufficient basis of research specializing in the relation between autonomy support and creativity in children, but so far, the research regarding this topic has shown to be relatively sparse. Beghetto (2006) found that positive teacher feedback about students' creativity was the strongest unique predictor of the likelihood of students taking the intellectual risks necessary to share their creative insights, interpretations, and ideas. Furthermore, in a study on second grade students performing a painting activity, Koestner et al. (1984) found that informational, noncontrolling teacher language enhanced creativity, while controlling language decreased creativity. Kupers et al. (2015 & 2017) found varying dyadic interactions of teacher autonomy support and student autonomy and musical performance in their multiple case studies during private music lessons. Dyadic teacher-student patterns, or attractor states, seemed to solidify over time. Furthermore, high teacher autonomy and low student autonomy expression seemed to particularly be related to student performance.

In line with the above-mentioned SDT, CDS theory and the anecdotal evidence, one can assume that there may be a relation between teacher autonomy support and student musical creativity, but so far there seems to have not been any research dedicated to (1) increasing teacher autonomy support with an intervention, while (2) studying the relation of teacher autonomy support and the two different forms of student creativity (convergent versus divergent). As enhancing children's musical creativity in the educational system could lead to a number of positive results (Collins, 2013; Jaschke et al., 2018), investigating various ways of increasing teacher autonomy support through an intervention and its relationship with different forms of creativity, could provide more insight regarding this valuable topic.

Current Study

Curious Minds is a Dutch educational cross-curricular program, originally created between 2006 and 2017 in a nationwide collaboration between universities and colleges for the science and technology domain. Curious Minds has been created from a Complex Dynamic Systems perspective and aims to increase the occurrence of upward spirals (also referred to as 'Talent Moments'), where creativity is sparked between teacher and student and a high level of enthusiasm, exchange of ideas and learning is elicited (Menninga et al., 2017). This is practiced by coaching teachers through a Video Feedback Coaching (VFCt) intervention to apply a variety of didactic techniques regarding an adaptive way of increasing student agency by focusing (1) on providing structure by scaffolding and (2) providing room for creative exploration by asking autonomy supportive student-centred open questions. The program has generated successful results in the Dutch science education, in both increasing teacher autonomy support, scaffolding and student performance (van Vondel et al., 2016 & 2017 ; Wetzels et al., 2016). Due to increasing public interest in music education in primary schools, the program has been adapted to other domains, such as music education.

Aim

The aim of this thesis is to provide a quantitative and qualitative analysis of the development of autonomy support, convergent thinking and acting in music and the dynamic interaction between varying levels of a teacher's autonomy support and the student's convergent thinking and acting in music. The following research questions were addressed:

Research questions

(1) How does the teachers' autonomy support develop over the course of the Curious Minds study?

(2) How does the students' convergent musical thinking and acting in musical creativity develop over the course of the Curious Minds study?

(3) Is there a relationship between teacher verbal autonomy support and student convergent musical thinking and acting?

Hypotheses

In line with previous Curious Minds research (van Vondel et al., 2016 & 2017 ; Wetzels et al., 2016), it was expected that (1) teacher autonomy support would increase over the course of the Curious Minds VFCt intervention. From a Complex Dynamic Systems perspective, student creativity can be increased by teacher-student interactions focused on student agency. As the Curious Minds VFCt intervention is specifically geared towards increasing student agency, it was expected that (2) student convergent musical thinking and acting would increase over the course of the intervention. Lastly, in line with the above evidence and the CDS theory (3), a positive relationship between teacher autonomy support and student convergent musical thinking and acting was expected.

Method

Data collection

Data was extracted from the second year of existing longitudinal PhD research within the Curious Minds program. This thesis targeted only the last eight out of ten videotaped lessons (each lasting 30-60 minutes) for six different classes, which provided the last two baseline measures, four intervention measures and two post measures, spaced across four to ten month intervals.

Participants

Six female teachers from four different Dutch primary schools voluntarily admitted themselves to the program. At the start of the data collection they were between 29 and 40

years of age (*M* = 35.8, *SD* = 4.1). They had an average level of eight years of teaching experience (*range* = 4-15, *SD* = 4.5) and an average level of two years of musical teaching experience (*range* = 1-3, *SD* = 1). There was a total amount of 136 children in the six classes, aged six to ten years old (*Mode* = 7) at the time of the video recording. The teachers and parents of the participating students gave informed consent before the start of the study, and all procedures conformed to existing ethical guidelines.

Procedure

Prior to the intervention there was a baseline measurement of classroom interactions of 2x2 lessons, with six weeks in-between the sets. Teachers were asked to commence the music lessons as they normally would. Only the second set of these measures was included in this thesis, as it was thought as the most accurate depiction of the teachers x class current dyad. Subsequently, teachers received an hour and a half of theory and training about the Curious Minds principles regarding 3 main didactic techniques; (1)creating space for the emergence of creative expression by asking child-centred open questions, (2)providing structure by working through the creative cycle and (3)co-constructing a higher level of creativity by adaptive support, or scaffolding (Hendriks et al., 2018). Furthermore, a lesson structure was introduced to provide room for the introduction of an instrument, classical exploration, individual exploration with the formation of focus groups and a classical integration. Teachers received a workbook regarding the mentioned didactic techniques, where personal goals were formulated and reflected upon. Four consecutive lessons were recorded during the intervention and co-constructive Video Feedback CFCt coaching of 45-60ms was provided after each lesson over the course of four to seven weeks. At least four to six weeks of no intervention followed. Finally, two post-measures were recorded, with two- three weeks between both measures.

Irregularities

From the six teachers and classes participating in this research, it was not possible to record the last post-measure lesson, of Teacher and Class 11, due to COVID-19 restrictions. Their results have been included in this thesis, although it has so be noted that their post-measures solely consisted of one lesson. Furthermore, Teacher and Class 10 and Teacher and Class 12 both had late post-measures due to COVID-19 restrictions; instead of the usual four to six week hiatus between intervention measures and post-measures, their post-measures started after nearly five months. Their measures have been included in this thesis and possible implications will be discussed.

Measurements

Four segments per lesson were coded with Media Coder, an online application for coding video footage (Bos & Steenbeek, 2009): three minutes from the beginning and end of the lesson, with two two-minute segments in the middle, totalling to ten minutes per lesson. Coding started from the first *task-related* utterance (e.g. related to a music lesson), and all task-related teacher-student utterances we recoded. Non-task related utterances were excluded from further analysis.

Verbal Autonomy Support (ASV)

ASV was coded on a scale of 1-8, where 1 indicated restriction of autonomy (“stop!”, “it’s not your turn yet”) and 8 indicated encouragement (“beautiful”, “please continue”). No teacher utterances were coded with a 0. Level 1 and 2 were considered as low autonomy support, level 3 -5 were considered as medium autonomy support, centring mainly around explanations and teacher centred questions, and level 6 – 8 were considered as high autonomy support, centring mainly around creativity evoking questions and remarks.

This scale was based on recent autonomy literature (Oliveira, 2010; Meindertsma et al., 2014, Van Vondel et al., 2017; Kupers et al., 2017).

Convergent Musical Thinking and Acting (CTA)

CTA was coded on a scale of 1-9 where 1 indicated single sensory motor actions (inspection of the instrument, producing a single sound) and 9 indicated a systems level of musical abstraction (combining various musical concepts, such as rhythm and sound, into a coherent longer musical sentence). No student utterances were coded with a 0. Level 1 -3 were considered as "Sensory Motor Actions (irregular, exploration behaviour regarding a single sound) level 4 -6 were considered as "Representations" (rhythm or sound are hesitantly combined into a musical pattern) and level 7 – 9 were considered as "Abstractions" (rhythm, melody and harmony arise into consistent musical patterns). This scale was based on the dynamic skill theory (Fischer & Bidell, 2007).

This scale has been approved for task-independent measures in the analysis of student's task-related utterances (Van der Steen et al., 2014, Meindertsma et al., 2012).

Coding Reliability

To establish inter-observer reliability for the application of the coding scheme, the inter-observer agreement was assessed by comparing the codes of the author of the main PhD research project (Hendriks et al., 2018) with the author of this thesis. Nine videos from the first year of the research project (18.8% of recordings used for this thesis) were double coded for teacher ASV, and 7 videos (14.6% of recordings used for this thesis) were double coded for student CTA.

The inter-observer agreement was considered sufficiently high for both measures, with 77% for teacher ASV coding and 82% for student CTA coding. A Cohen's kappa was calculated to determine the consistency of coding among the two observers, with a substantial

kappa of $\kappa = 0.68$ for teacher ASV coding and a substantial kappa of $\kappa = 0.77$ for student CTA coding (Viera & Garett, 2005).

Data Analysis

Quantitative analyses were used to be able to characterize long-term change in teachers and classes across the three research phases (baseline, intervention, post-measure) and qualitative illustrations were used to zoom in on characteristic patterns that played a role in the overall change patterns.

Quantitative

For the quantitative analysis of a global increase or decrease over time of both ASV and CTA, Monte Carlo permutation tests were used, as parametric assumptions could not be met due to the small sample size (Todman & Dugard, 2001). A Monte Carlo analysis is a nonparametric test that evaluates the null hypothesis that the probability of the specific association between the variables under study is based on chance alone (Firestone et al., 1997). The Monte Carlo permutation test provides an estimation of the exact *p* value, which is the probability that the same or a greater difference is found if the null hypothesis is supported (Menninga et al., 2017). Furthermore, significance effect size values were computed using Cohen's d (Cohen. 1992). Based on recent findings (Menninga et al., 2017; van Vondel et al., 2016), an effect size of below .20 and/ or a p value above .05 was considered as non-meaningful, and providing no support for rejecting the null hypothesis, an effect size of between .20 and .50 and/ or a p value below .05 was considered as moderately meaningful, and providing some support for rejecting the null hypothesis, an effect size of above .50 and/ or a p value below .01 was considered as highly meaningful, and providing strong support for rejecting the null hypothesis. All ASV and CTA measures were compared for the intervention versus baseline, post-measures versus intervention and post-measures versus baseline, to obtain an integrated idea of their trajectories.

For the quantitative analysis of the relation between ASV and CTA over time, student CTA codes were paired with the highest teacher ASV codes within the appropriate timeframe. The teacher ASV utterances that were not followed by a student CTA code, were left out of this analysis. Subsequently, a Spearman's Rho correlation was used to calculate the correlation between the two variables overall and within conditions for each teacher x class dyad. A Spearman's $r_s \geq .70$ indicated a very strong (positive or negative) relationship. A Spearman's r_s between .40 and .69 indicated a strong relationship, a r_s between .30 and .39 indicated a moderate relationship, a r_s between .20 and .29 indicated a weak relationship and a $r_s < 20$ indicated a negligible or non-existing relationship (Dancey & Reidy, 2007).

When all teachers and/or classes displayed similar quantitative results (e.g. moderately or highly meaningful results) between conditions, this provided strong support for an overall rejection of the null hypothesis. If this pattern was visible during either the intervention measures or the post-measures, the hypothesis had to be party accepted; for one condition and not the other. If the majority of the teachers or classes ($N > 3$) displayed a similar quantitative result pattern, this provided moderate support for an overall rejection of the null hypothesis. If teachers or classes displayed varying result patterns (e.g. increase, no effect, or decrease) the null hypothesis had to be accepted. Teachers and/or classes that displayed unexpected results were compared with teachers and/ or classes who followed the most meaningful/strongest expected pattern in the qualitative analyses.

Qualitative

For the visual inspection of the development of both teacher ASV and student CTA, time series analyses were performed. No teacher utterances (0) were recoded as a neutral level 5 and the previous ASV levels 5 to 8 were recoded as a 6 to 9. No student utterances (0) were recoded as a neutral CTA level 5 and the previous CTA levels 5 to 9 were recoded as a 6 to 10. This choice was made based on the assumption that leaving no teacher and no student

utterances at ASV or CTA level 0, would distort the visual data as silence would simply categorized as the lowest autonomy or creativity level. Assuming the neutral value of silence (no utterances), the value of 5 was therefore used.

As the time series graphs provided a highly detailed, yet dense display of the individual trajectories, a Moving Maximum graph of 51 seconds (van Geert & van Dijk, 2002) was computed to attain a more visually comprehensible image. For ASV, the graphs of the teacher with the largest ASV increase, and the teacher with the smallest ASV increase or largest ASV decrease, were compared with each other by inspecting the overall pattern and the amount of high ASV peaks (level 8-9) between teachers and between conditions. For student CTA, the same method was followed by comparing the overall pattern and high CTA peaks (level 8-10) of the class with the largest CTA increase, with the class with the smallest CTA increase or largest CTA decrease.

For the relationship between teacher ASV and student CTA over time, State Space Grids (Hollenstein & Lewis, 2006) were created per condition for each teacher x student dyad. Again, the teacher x class trajectory that displayed the strongest relationship between both variables was compared to a teacher x class trajectory that showed the smallest relationship or followed an otherwise unexpected pattern.

The State Space Grids displayed the interaction between different teacher- and student levels by the amount and size of the dots in different regions marked; indicating low, medium and high ASV and CTA. A larger dot size and frequency in a specific region indicated an *attractor state* or a behavioural self-sustaining pattern between teachers and students (Hollenstein & Lewis, 2006). Attractor states moving towards the right upper corner indicated a positive relationial pattern between teacher ASV and student CTA.

The purpose of these qualitative analyses was to provide (1) further insight and/or (2) nuance for the quantitative results found and/or to (3) provide a possible explanation of

unexpected patterns. Together with the quantitative results, a final conclusion was summarized for the overall acceptation or rejection of the null hypothesis for each research question.

Results

The Development of Teacher Autonomy Support (ASV) over the Course of the Study

Quantitative

Table 1 depicts the teacher autonomy support (ASV) results for the three intervention conditions. Teacher results were ordered by means of their Monte Carlo significance and effect size across the three intervention conditions.

Table 1

Summary of Baseline, Intervention and Post- Measures of Teachers' ASV 1-8

	Baseline		Intervention		Post-Measure		Intervention - Baseline		Post-Measure - Baseline		Post-Measure-Intervention	
Teacher	*M*	*SD*	*M*	*SD*	*M*	*SD*	*Mean diff.*	*Cohen's d*	*Mean diff.*	*Cohen's d*	*Mean diff.*	*Cohen's d*
9	3.64	1.29	4.41	1.58	5.64	1.82	.76**	.53	1.99**	1.26	1.23**	.72
8	3.39	1.14	4.23	1.64	4.31	1.94	.84**	.59	.58**	.58	.08	.05
10	3.91	1.37	4.89	1.67	4.15	1.33	.98**	.64	.24	.18	-.74*	.49
7	3.85	1.64	4.18	1.68	4.17	1.66	.34*	.20	.32*	.19	-.03	.01
11	4.47	1.80	4.94	1.66	4.85	1.70	.47*	.27	.38*	.23	-.09	.05
12	4.36	1.38	4.76	1.75	3.75	1.23	.39*	.25	-.61*	.47	-1.01**	.66

Note. Monte Carlo analysis *p <.05, **p <.001. Teacher 11 was missing one post-measure. Teacher 10 and 12 had late post-measures.

As is depicted in Table 1, all teachers showed a significant increase in ASV levels between intervention and baseline measures, with small to average effect sizes, ranging from d =.25 to d = .64. This implies a positive, moderate to highly meaningful ASV increase when the intervention was introduced.

In general, teachers showed mixed results when comparing post-measures and intervention measures. That is, one teacher (Teacher 9) showed a continued significant and highly meaningful ASV increase in the post measures, compared to the intervention measures ($M_{\text{postmeasure Teacher9}}$ = 5.64, $M_{\text{intervention Teacher9}}$ = 4.41, $p < .001$, d = .72). Two teachers (Teacher 10 & 12) showed a significant, moderate to highly meaningful ASV decrease in the post measures, compared to the intervention measures and three teachers (Teacher 7, 8 & 11) did not show a significant difference between both conditions, indicating an overall stabilization of ASV.

As ASV levels mainly stabilized after the intervention, it could be assumed that most teachers would still display significantly higher ASV in the post-measures compared to the baseline measures. Indeed, four teachers (Teacher 7, 8, 9 & 11) showed a significant, moderate to highly meaningful ASV increase in the post-measures, while one teacher (Teacher 12) showed a significant, highly meaningful ASV decrease in the post measures, compared to the baseline measures ($M_{\text{postmeasure teacher12}}$ = 3.75, $M_{\text{baseline teacher12}}$ = 4.36, $p < .05$, d = .25) and one teacher (Teacher 10) did not show a significant difference between both conditions. Interestingly, Teacher 10 and 12 were the teachers with late post-measures.

These results imply that ASV levels mainly stabilized or dropped after the coaching intervention. One teacher (Teacher 9) was able to sustain an overall ASV increase, displaying a significant increase between all conditions. Four out of six teachers still displayed significantly higher ASV during post-measures than during baseline measures. The other two had late (e.g. about five months later than the others) post-measures, which may indicate that

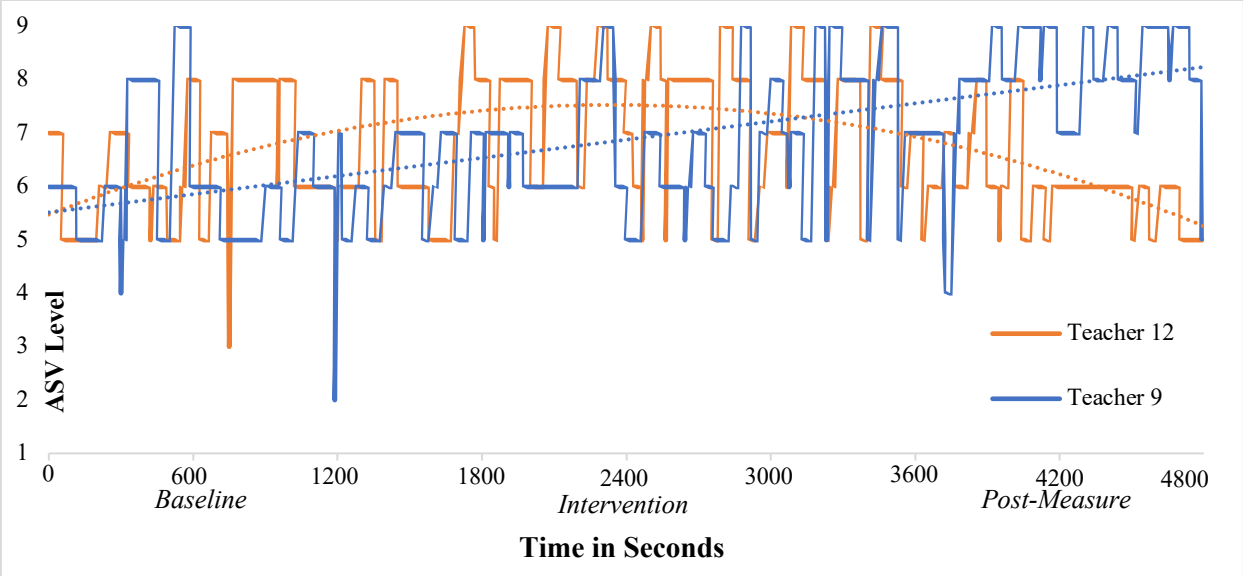
the coaching intervention was associated with a temporal ASV increase, but further intervention may be required to sustain ASV long term.

Qualitative

Appendix A displays teacher frequencies and percentages per ASV level and intervention condition in **Table 2**. Teacher 9, who showed the largest overall ASV level increase, displayed a decrease of low- and medium ASV levels and an increase of high ASV levels across the study. Teacher 12, who showed the smallest ASV level increase between the experimental and baseline measures, decreased her high ASV levels and increased her medium ASV levels in the post-measures compared to the experimental measures. Both of their moving maximum graphs are depicted in **Figure 1** below, displaying the differences between their trajectories.

Figure 1.

Moving Maximum Graph with a Window of 51 Seconds Displaying ASV Teacher Trajectories over the Course of the Study



Note: Each lesson contained 600 seconds. The intervention phase started at 1200 seconds and ended at 3600 seconds. Here, ASV level 5 indicates no utterances and levels 6-9 indicate the higher ASV levels 5-8.

Figure 1 displays a variable pattern of mid- to high ASV peaks for Teacher 9 during baseline measures. During the first two lessons of the intervention (up to 2400 seconds), there are predominantly level 5-7 ASV peaks visible. During the last five lessons, high ASV peaks become more frequent, building up to almost solely level 8 -9 peaks during the post-measures. Teacher 12 also displays a variable mid- to high ASV pattern during baseline measures. After the first lesson intervention (1800 seconds) until the last lesson of the intervention (3600 seconds), there is a stable pattern visible of first a level 9 peak, followed by a level 8 and a level 6 or 5 peak. During post-measures however, there are solely two high ASV peaks visible, followed by constant medium ASV peaks.

Interestingly, teacher 9 showed fewer high ASV peaks than teacher 12 during the baseline and intervention measures. However, during the post measures, Teacher 9 almost solely displayed high ASV peaks, while Teacher 12 mainly displayed medium ASV level peaks. Therefore, the trajectory of Teacher 9 is showing an almost linear pattern, while the trajectory pattern of Teacher 12 is more reminiscent of a downward opening parabola, explaining their overall Monte Carlo results, of continued increase, or increase and subsequently decrease.

This implies that each teacher showed a different learning trajectory during the entire study; some teachers were quick adapters but reverted back to old behaviour when the intervention ended. Others showed a slow, but steady increase of learning and adapting higher levels of autonomy in the classroom long-term.

Summary

These results indicate that teacher ASV increased during the intervention phase compared to the baseline phase, three out of six teachers showed moderate support for the rejection of the null hypothesis and the other three showed strong evidence for the rejection of the null hypothesis. Furthermore, based on this data, there is not enough evidence for a continued increase of teacher ASV during post-measures, as one teacher showed a strong ASV increase, three teachers showed an ASV stabilization and two teachers showed a strong ASV decrease. However, it has to be noted that the two teachers with a strong ASV decrease, were the teachers with late post-measures. Despite this, the null hypothesis has to be accepted for the post-measures, as continued ASV increase was expected.

Based on the qualitative analyses of two teachers, it can be assumed that teacher ASV developed in an idiosyncratic manner over the course of the study; pointing towards different adaptive learning behaviours. The stabilization or decline in ASV during post-measures after an increase during the intervention, may indicate that teachers benefitted from the intervention, but may have needed a longer intervention phase to sustain high ASV long-term. Therefore, the hypothesis that teacher ASV would increase over the course of the study can be partly accepted, with the nuance that ASV stabilization, instead of continued ASV increase during post-measures should be expected for future reference.

The Development of Student Convergent Musical Thinking and Acting (CTA) over the Course of the Study

Quantitative

Table 3 depicts student Convergent Thinking and Acting (CTA) results for the three intervention conditions. Student CTA results were ordered by means of their Monte Carlo significance and effect size across the three intervention conditions.

Table 3

Summary of Baseline, Intervention and Post-Measures of Student CTA 1-9

	Baseline		Intervention		Post-Measure		Intervention - Baseline		Post-Measure - Baseline		Post-Measure-Intervention	
Class	*M*	*SD*	*M*	*SD*	*M*	*SD*	*Mean diff.*	*Cohen's d*	*Mean diff.*	*Cohen's d*	*Mean diff.*	*Cohen's d*
10	5.01	1.12	5.59	1.64	6.23	1.11	.58**	.41	1.22**	1.10	.65**	.46
9	6.19	1.93	5.82	2.37	7.16	1.53	-.38	.17	.97**	.56	1.34**	.67
11	4.14	2.13	4.57	1.70	4.27	1.05	.43**	.22	.13	.08	-.30	.21
7	6.21	1.92	5.80	1.92	6.31	1.56	-.41*	.22	.1	.06	.51*	.29
8	5.37	1.77	4.99	1.21	4.67	0.95	-.38*	.25	-.71**	.50	-.33*	.30
12	6.04	1.74	6.12	1.62	5.88	2.11	.07	.12	-.16	.09	-.23	.12

Note. Monte Carlo analysis *p <.05, **p <.001. Class 11 was missing one post-measure. Class 10 and 12 had late post-measures.

As is depicted in **Table 3**, students showed mixed CTA results across conditions. Two classes (Class 10 and 11) showed a significant, moderately meaningful CTA increase in the intervention measures compared to the baseline measures and two classes (Class 7 and 8) showed a significant, moderately meaningful CTA decrease in the intervention measures compared to the baseline measures. Three classes (Class, 7, 9 & 10) showed significant, moderate to highly meaningful CTA increase in the post-measures compared to the intervention measures and one class (Class 8) showed a significant, moderately meaningful CTA decrease in the post-measures compared to the intervention measures. Two classes (Class 9 & 10) showed significant, highly meaningful CTA increase in the post-measures compared to the baseline measures and one class (Class 8) showed a significant, moderately meaningful CTA decrease in the post-measures compared to the baseline measures. Unlike the teachers, Class 10 and 12 did not show a clear distinction from the other classes during post-measures.

Overall, one class (Class 10) seemed to show a highly meaningful CTA increase across all conditions, and one class (Class 8) seemed to show a moderately meaningful CTA decrease across all conditions. Two classes (Class 9 and 7) showed a moderate to highly meaningful CTA decrease during the intervention measures compared to the baseline measures, and either restored their baseline CTA levels during post-measures (Class 7) or strongly increased their CTA levels compared to their baseline CTA (Class 9). One class (Class 11) moderately increased their CTA levels during intervention measures and stabilized their CTA during post-measures. One class (Class 12) did not show any change in CTA.

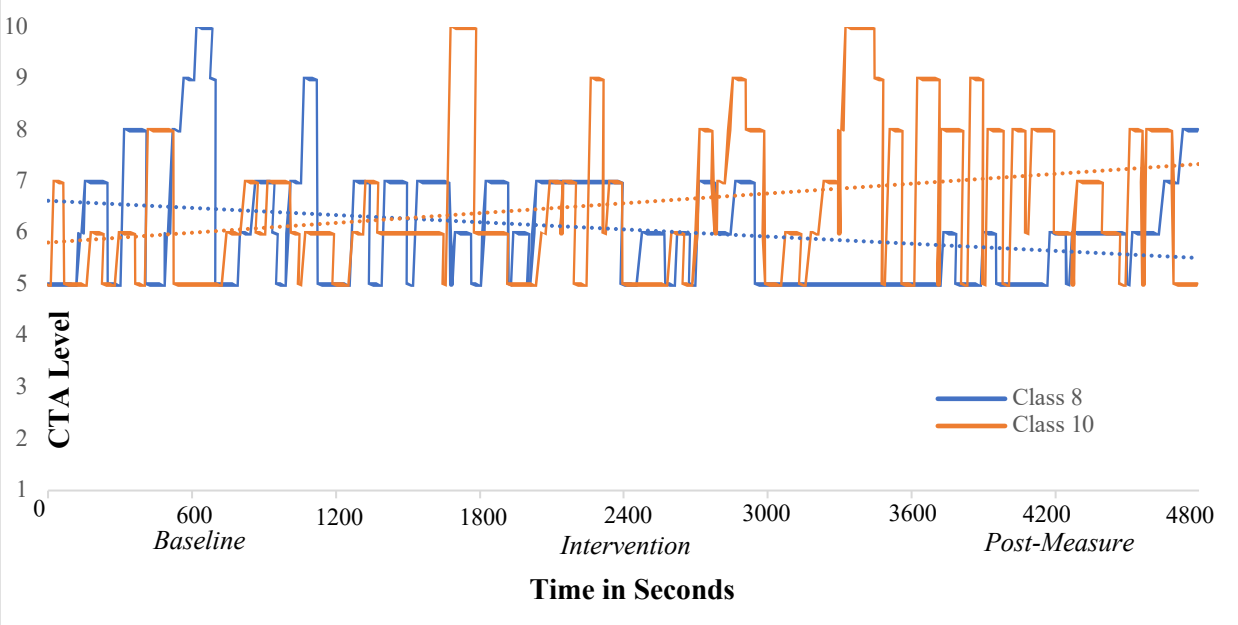
Qualitative

Table 4 in Appendix B displays the frequencies and percentages of class CTA per level and condition. Class 10, who showed a highly meaningful overall CTA level increase, displayed a decrease of low- and medium CTA levels and an increase of high CTA levels across the intervention. Class 8, who showed a moderately meaningful overall CTA level decrease, displayed an increase of medium CTA levels and a decrease of high CTA levels across the intervention. Both of their moving maximum graphs are depicted in **Figure 2** below, displaying the differences between their trajectories.

Class 10 displays an almost linear pattern of medium CTA peaks at the baseline measures, with high peaks increasing from the intervention phase until the end of the post-measures. Class 8 displays mid- to high CTA peaks at the baseline measures, which decrease throughout the intervention with often level 5, no utterances, as highest CTA peak. At the end of the last lesson, CTA peaks seem to increase again. These trajectories show an increase for Class 10, and a decrease for Class 8.

Figure 2

Moving Maximum Graph with a Window of 51 Seconds Displaying CTA Class Trajectories over the Course of the Study



Note: Each lesson contained 600 seconds. The intervention phase started at 1200 seconds and ended at 3600 seconds. Here, CTA level 5 indicates no utterances and levels 6-10 indicate the higher CTA levels 5-9.

Summary

Both the quantitative and qualitative analyses indicate that student CTA developed itself in different directions. Two out of six classes showed moderate to strong evidence for the rejection of the null hypothesis during the intervention phase, while four out of six showed patterns of stabilization or decrease, which indicated the acceptance of the null hypothesis during the intervention phase. Furthermore, three out of six classes showed moderate to strong evidence for the rejection of the null hypothesis during the post-measures, while three out of six showed a weak pattern of decrease, indicating the overall acceptance of the null hypothesis during post-measures. Therefore, the hypothesis that student CTA would increase over the course of the study, has to be rejected.

The Relation Between ASV and CTA over the Course of the Study

Quantitative

Table 5 depicts Teacher Verbal Autonomy Support (ASV) x student Convergent Thinking and Acting (CTA) correlations for the three intervention conditions. Results were ordered by means of their significance and correlation strength across the three intervention conditions.

Table 5 *Summary of Spearman correlations between Teacher ASV and Class' CTA*

	Baseline	Intervention	Post-Measure	Overall
Measure	*Correlation r_s /(N)*	*Correlation r_s /(N)*	*Correlation r_s /(N)*	*Correlation r_s /(N)*
ASV x CTA 8	.05 *(31)*	*.25 (83)**	.48*(49)***	.19 *(163)**
ASV x CTA 9	-.42 *(18)*	.07 *(95)*	.37 *(45)**	.25 *(158)***
ASV x CTA 10	-.23 *(24)*	.10 *(71)*	.07 *(28)*	.04 *(123)*
ASV x CTA 11	.26 *(23)*	-.08 *(71)*	-.39 *(13)*	-.05 *(107)*
ASV x CTA 12	-.01 *(27)*	.16 *(60)*	.17 *(26)*	.13 *(123)*
ASV x CTA 7	-.04 *(54)*	.02 (*91*)	.12(*62)*	.04 (*207*)
ASV x CTA ALL	-.01 (177)	.12 (471)*	.35 (233)**	.12 (881)*

Note. Spearman Correlation analysis * p <.05, **p <.001. Teacher x Class 11 were missing one post-measure. Teacher x Class 10 and 12 had late post-measures.

As depicted in Table 5, four out of six teacher-class dyads displayed non-significant correlations. Teacher x Class 8 showed a significant, but weak overall positive correlation between ASV and CTA, $r_s = .19$, $p < .05$, $N = 163$. ASV and CTA were also weakly correlated within the intervention phase, $r_s = .25$, $p < .05$, $N = 83$ and strongly correlated during the post-measure condition, $r_s = .48$, $p < .001$, $N = 49$. Teacher x Class 9 also showed a significant, but weak overall correlation between ASV and CTA, $r_s = .25$, $p < .001$, $N = 158$. There were no significant correlations between ASV and CTA within the intervention

phase for Teacher x Class 9. ASV and CTA were moderately correlated in the post-measure condition $r_s = .37$, $p < .05$, $N = 45$.

As ASV x CTA interactions showed relatively infrequent during baseline and post-measures per teacher x class dyad, it was subsequently decided to include all teacher x class measures into a group-level analysis. Overall, there was a significant, but negligible correlation found between ASV and CTA, $r_s = .12$, $\rho < .05$, $N = 881$. During baseline measures, there was not a correlation found. For the intervention phase, a significant, but negligible correlation between ASV and CTA was found, $r_s = .12$, $\rho < .05$, $N = 471$. ASV and CTA were moderately correlated during the post-measures, $r_s = .35$, $\rho < .001$, $N = 233$.

These results may indicate a delayed intervention effect; as higher ASV levels were manipulated during the intervention, and mostly stabilized during post-measures, teachers and their classes may have developed a sustainable dyad where ASV x CTA interactions showed more prominently.

Qualitative

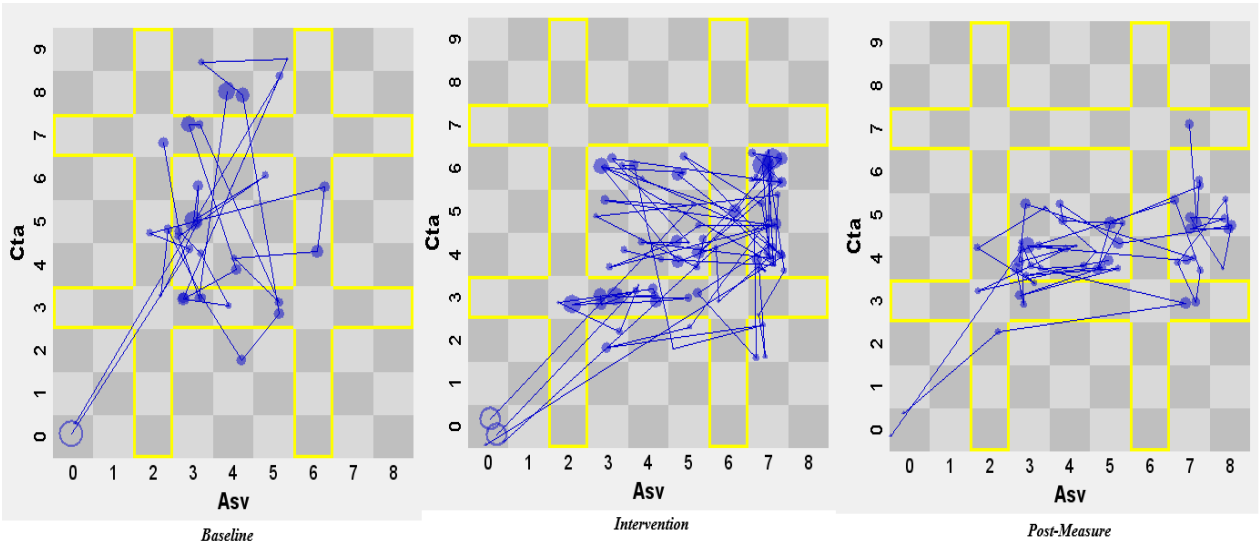
Upon inspection of the individual results of teacher ASV and class CTA, it was expected that Teacher x Class 10 would display a positive correlation, as both of their measures significantly increased over the intervention. Class 8 displayed a significant CTA decrease with a significant Teacher 8 ASV increase, therefore a positive correlation for this teacher x student dyad was not necessarily expected. Their calculated correlations however, displayed an unexpected pattern. Therefore, their relationships were further assessed with State Space Grids, displayed in **Figure 3** and **Figure 4** below.

Figure 3 displays a baseline dyad of mostly medium teacher ASV and low to medium CTA interactions with attractor states in the medium ASV and medium CTA region, visited 26.3% of the time. One attractor state is based in the medium ASV and high CTA region, visited 25.3% of the time.

Interestingly, the Intervention grid shows a shift from medium to high teacher ASV, but paired with predominantly low to medium student CTA levels, with medium ASV x medium CTA attractor states, visited 25% of the time and high ASV x medium CTA attractor states, visited 39.3% of the time. The Post-Measure grid shows a similar dyad, with medium ASV x medium CTA attractor states visited 51.9% of the time and high ASV x medium CTA attractor states, visited 30.9% of the time.

Figure 3

State Space Grid (SSG) of ASV Teacher 8 x CTA Class 8 Interactions over the Course of the Intervention



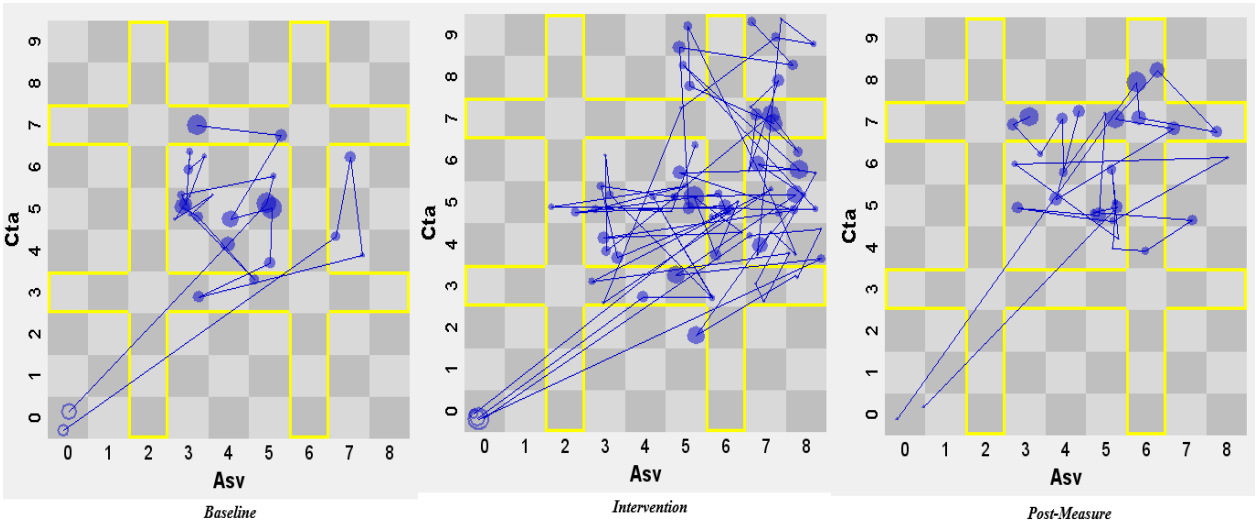
Note: Only teacher ASV x class CTA interactions were taken into account; ASV measures without a CTA response were left out. The intervention phase contained four lessons, while the baseline and post-measure phases contained two lessons each. The regions marked depict low (0-2), medium (3-5) and high (6-8) ASV and low (1-3), medium (4-6) and high (7-9) CTA.

The grids of Teacher x Class 10, who both showed significant positive Monte Carlo results for their respective ASV and CTA measures, but showed weak and negative correlations between the variables, are depicted in **Figure 4** below. The baseline grid also

shows a medium teacher ASV and medium CTA attractor state, visited 64.1% of the time. A smaller attractor state is based in the medium ASV and high CTA region, visited 14% of the time.

Figure 4

State Space Grid Showing ASV Teacher 10 x CTA Class 10 Interactions over the Course of the Intervention



Note: Only teacher ASV x class CTA interactions were taken into account; ASV measures without a CTA response were left out. The intervention phase contained four lessons, while the baseline and post-measure phases contained two lessons each. The regions marked depict low (0-2), medium (3-5) and high (6-8) ASV and low (1-3), medium (4-6) and high (7-9) CTA.

The Intervention grid displays a shift from medium to high teacher ASV and medium to high student CTA levels, with medium ASV x medium CTA attractor states visited 22.2% of the time and high ASV x medium CTA attractor states, visited 32.5% of the time and high ASV x high CTA attractor states, visited 17.5% of the time. The Post-Measure grid displays a predominantly medium teacher ASV and mid to high student CTA levels, with medium ASV x medium CTA attractor states visited 31.2% of the time and medium ASV x high CTA

attractor states, visited 31.4% of the time and high ASV x high CTA attractor states, visited 31.6% of the time.

Interestingly, the dyad of Teacher x Class 10 seems to depict a greater shift towards higher ASV and CTA levels during the intervention measures and post-measures than the dyad of Teacher x Class 8; which corresponds with the Monte Carlo results of the variables individually. However, the dyad of Teacher x Class 8 does seem to show more consistency with the size and frequency of the attractor states than the dyad of Teacher x Class 10; with prominent medium ASV x CTA attractor states and high ASV x medium CTA attractor states during the intervention measures, compared to the mid- to high ASV x CTA attractor states of Teacher x Class 10, displaying a larger variety of interaction.

Summary

These results indicate that there are there is not sufficient evidence for a positive relation between ASV and CTA, but nuances are there to be found. When looked at teacher x class results individually, barely any relations where found, perhaps also due to lack of statistical power. Two out of six teachers did display moderate- to strong positive correlations between ASV and CTA during post-measures. This may have accounted for the overall moderately strong positive correlation found between ASV and CTA for post-measures. As all teacher x student dyads did not display any correlations between ASV and CTA during baseline measures, and barely any during the interventions, this may indicate a slow, but significant shift in teacher-student dyads.

State Space Grids provided the additional insight that even though attractor states can shift towards higher regions, indicating a positive relationship, overall correlations may have shown to be weak due to lower ASV paired with high CTA and vice versa. Based on these results, the hypothesis of a positive relationship between ASV and CTA has to be rejected.

Discussion

The classroom forms an interesting playground for unique ways of interacting and learning, for students and teachers. The current Curious Minds study provided the opportunity for studying humans in their natural environment while observing many interacting variables. This thesis singled out the variables of teacher autonomy support and student musical convergent thinking and acting and analysed their development and possible relationship. In line with previous Curious Minds research (van Vondel et al., 2016 & 2017 ; Wetzels et al., 2016) and previous creativity research (Beghetto, 2006; Koestner et al.,1984; Kupers et al.,2015 & 2017) it was hypothesized that an intervention geared towards increasing teacher autonomy support, would in turn, generate higher levels of musical creativity amongst students.

On basis of the discussed results, the hypothesis that teacher's autonomy support increased over the course of the study, can be partly accepted. All teachers showed a moderate- to large increase in autonomy support during the intervention compared to the baseline measures. Afterwards, different trajectories of increase, stabilization or decrease became visible, indicating that the four coaching sessions could not be associated with a longer term increase of teacher autonomy support. The hypothesis that student convergent musical creativity increased over the course of the intervention, can be rejected, due to varying results, such as increase, or decrease, across conditions. Despite an overall moderate positive relationship between teacher autonomy support and student convergent creativity in post-measures, the hypothesis for a positive relation can be rejected due to weak or negligible correlations within conditions for four out of six Teacher x Student dyads. Two Teacher x Student dyads did show an increasing positive relationship over the course of the study, which may point towards higher teacher autonomy slowly increasing student convergent musical creativity over time in some cases. State Space Grids provided the nuanced

explanation that despite an increase in both autonomy support and convergent musical creativity in one case, correlations may have shown to be weak due to varying interaction levels between both variables. This does support the Complex Dynamic Systems theory in explaining why variables should be studied in interaction with each other, and not singled out on their own.

Limitations and Future Directions

Studying humans in their own environment generally produces higher external validity, but low internal validity (Steckler & McLeroy, 2008) due to confounding factors. The internal validity of this research could be compromised due to the following factors.

Singing versus Playing

During baseline measures, teachers were free to organize the music lessons according to their taste. Almost all of them performed at least one musical singing lesson. Children were sometimes already familiar with the song, but it can also be argued that children would easier be able to reach higher levels of convergent creativity over a medium that they are familiar with (singing), instead of a medium they are unfamiliar with (a musical instrument).

This may explain why two classes showed an increase and two classes showed a decrease in convergent creativity during the coaching intervention. Based on this assumption however, it should be expected to see an overall increase in convergent creativity during post measures. Yet still, three out of six classes displayed an increase in convergent creativity during post measures compared to their levels during the coaching intervention, indicating that this factor may only be one part of a bigger story.

Future research could either ask teachers to perform a musical lesson with an instrument or an otherwise musical performance, such as body percussion, during the baseline measures instead of singing. Extending the intervention phase could provide more

insight regarding teacher autonomy trajectories, but especially examine whether students showed a delayed convergent creativity effect, due to unfamiliarity with the instruments and lesson format.

Measure Limitations

The measures of three teachers were affected due to COVID-19 restrictions, with two teachers having delayed post-measures and one teacher with a missing final post-measure. Unlike the others, the teachers with delayed post-measures did display a drop in autonomy during this condition, compared to the intervention measures. Furthermore, students were of varying ages (six to ten years), teachers had varying levels of teaching experience (four to fifteen years) and musical teaching experience (one to three years). Age and musical experience could have been interacting factors during this study.

Future research could perform control analyses of demographic factors possibly correlating with teacher autonomy and/or student convergent creativity.

Analysis Limitations

Interactions between teacher autonomy support utterances and student convergent musical thinking and acting utterances showed relatively infrequent. This may have contributed to the weak correlations found, perhaps due to lack of statistical power.

Computing time series from these events and using this data for correlational analysis in the future may be useful, as the length of the interaction between teacher and student(s) could provide valuable information over the nature of teacher x student dyads and add statistical power.

Dynamic Involvement of Other (Manipulated) Factors

Even though this thesis analysed two clearly defined variables, it should be not forgotten that these singled-out variables were part of a larger study. The Curious Minds intervention did not only target teacher autonomy support, it also targeted scaffolding and

providing structure. These three factors were not manipulated on their own during separate lessons, teachers were using them interchangeably during each lesson. Furthermore, in line with Webster's (1990) creativity theory, convergent and divergent creativity are also in constant interplay with each other.

For example, one pattern observed in teacher x student interactions was; teachers providing room for autonomous exploration, students exploring music in a divergent manner, teachers lowering their autonomy support by responding in a more directive or structured manner, students responding in a more convergent musical manner. Therefore, it could be hypothesized that teacher autonomy support might be more specifically related to divergent musical creativity, while, scaffolding and proving structure might be more related to convergent musical creativity. It is therefore not necessarily a limitation of this entire study, but more a limitation of this particular thesis, attempting to single out two variables out of a more complex web of interacting behaviours.

Future research could therefore include measures of scaffolding, providing structure and divergent creativity in its' analysis, to obtain a more well-rounded view of the possible relations. As studying all these variables for six Teacher x Student dyads may be too ambitious for the average master thesis author, future graduates could choose to focus on an in-depth analysis of one Teacher x Student dyad instead.

Theoretical and Practical Implications

What this thesis did not show was a clear, positive relation between teacher autonomy and student convergent musical creativity. However, valuable theoretical and practical implications can be drawn from this thesis. As far as the theoretical implications are concerned, this thesis illuminates the importance of studying variables in relation with each other over time, from a mixed method approach. It does support the Complex Dynamic

Systems Theory (Kupers et al., 2017) in revealing specific behavioural patterns between teachers and students when looking at data in relation with each other, instead of solely comparing means. Exploring individual trajectories with Moving Maximum graphs and State Space Grids was a strength of this thesis, as it provided overall data with nuances that could have been overlooked otherwise.

Even though teacher autonomy support was successfully increased during the intervention, these increased autonomy support levels seemed to stagnate or even decline long-term. Further research could focus on understanding under which circumstances teachers and their students can create and maintain an upward spiral. And what characterizes an upward spiral specifically in the context of music education? Does it solely consist of highly autonomy stimulating teacher utterances and high levels of creative student thinking and acting? The inclusion and study of variables hypothesized to facilitate or moderate these effects, could provide more insight in the development of musical creativity amongst children. Furthermore, it could provide more insight regarding different adaptive learning behaviours for teachers and students, which could lead to interventions targeting these specific behaviours.

Conclusion

This thesis may not directly practically impact the implementation of music education in schools, yet it may contribute to more targeted interventions that produce longer lasting effects by the directions and implications provided previously. Furthermore, despite teacher satisfaction not being measured in this thesis, it has to be noted that all teachers and classes were exploring and performing music together in the post-measures. Therefore, the Curious Minds intervention could already be considered a success as it stimulated teachers in moving away from singing lessons to more creative ways of exploring music together with children.

References

Beghetto, R. A. & Kaufman, J.C. (2016). Ever broadening conceptions of creativity in the classroom. In: R.A. Beghetto. & J.C. Kaufmann (Eds.). *Nurturing creativity in the classroom* (pp. 67-85). New York, NY: Cambridge University Press. https://doi-org.proxy-ub.rug.nl/10.1017/9781316212899

Bos, J., & Steenbeek, H. (2009). Mediacoder: software voor het coderen van video-en audio-materialen. *Groningen: Internal publication. IDP Department, University of Groningen*.

Collins, A. (2014). Music education and the brain: what does it take to make a change? *Update: Applications of Research in Music Education, 32*(2), 4–10. https://doi-org.proxy-ub.rug.nl/10.1177/8755123313502346

Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson education.

Deci, E.L., & Ryan, R.M. (1985). Intrinsic Motivation and self-determination in human behavior. New York: Plenum. https://doi.org/10.2307/2070638

Deci, E.L., Eghrari, H., Patrick, B.C., & Leone, D.R. (1994). Facilitating internalization: The self-determination perspective. *Journal of Personality, 62*, 119-142. https://doi-org.proxy-ub.rug.nl/10.1111/j.1467-6494.1994.tb00797.x

Firestone, M., Fenner-Crisp, P., Barry, T., Bennett, D., Chang, S., Callahan, M., ... & Knott, S. M. (1997*). Guiding principles for Monte Carlo analysis*. Washington, DC: US Environmental Protection Agency.

Fischer, K. W., & Bidell, T. R. (2007). Dynamic development of action and thought. In Damon, W. & Lerner, R., M (Ed.), *Handbook of child psychology, 1* (pp. 313 -393). Wiley. https://doi.org/10.1002/9780470147658.chpsy0107

Gagne, M. (2003). The role of autonomy support and autonomy orientation in prosocial behavior engagement. *Motivation and Emotion, 27*(3), 199–223. https://doi.org/10.1023/A:1025007614869

Geert, P. L. C., & Van Dijk, M. W. G. (2011). The dynamics of children's science and technology talents: A conceptual framework for early science education. *Netherlands Journal of Psychology, 66* , 96–109.

Green, L. (2006). Popular music education in and for itself, and for 'other' music: current research in the classroom. *International Journal of Music Education, 24*(2), 101–118. https://doi.org/10.1177/0255761406065471

Guthrie, J.T., & Davis, M.H. (2003). Motivating Struggling Readers in Middle School Through an Engagement Model of Classroom Practice. *Reading & Writing Quarterly, 19*(1), 59-85. https://doi-org.proxy-ub.rug.nl/10.1080/10573560308203

Hendriks, L.H., Steenbeek, H.W., & Bisschop Boele, E. H. (2018). *Curious Minds Muziekeducatie: werkboek voor leerkrachten in de middenbouw van het basisonderwijs: muzikale creativiteit in de klas.* Hanzehogeschool Groningen.

Hollenstein, T., & Lewis, M. D. (2006). A state space analysis of emotion and flexibility in parent-child interactions. Emotion, 6(4), 656. doi:10.1037/1528-3542.6.4.656

Jaschke, A. C., Honing, H., & Scherder, E. J. A. (2018). Longitudinal analysis of music education on executive functions in primary school children. *Frontiers in Neuroscience, 12*, 103–103. https://doi.org/10.3389/fnins.2018.00103

Koestner, R., Ryan, R. M., Bernieri, F., & Holt, K. (1984). Setting limits on children's behavior: the differential effects of controlling vs. informational styles on intrinsic motivation and creativity. *Journal of Personality, 52*(3), 233–248. https://doi.org/10.1111/j.1467-6494.1984.tb00879.

Kupers, E., van Dijk, M., & van Geert, P. (2017). Changing Patterns of

Scaffolding and Autonomy During Individual Music Lessons: A Mixed Methods Approach, *Journal of the Learning Sciences, 26*:1, 131-166. doi: 10.1080/10508406.2016.1259624

Kupers, E., van Dijk, M., van Geert, P., & McPherson, G. E. (2015). A mixed-methods approach to studying co-regulation of student autonomy through teacher–student interactions in music lessons. *Psychology of music, 43*(3), 333-358. https://doi.org/10.1177/0305735613503180

Kupers, E., Lehmann-Wermser, A., McPherson, G., & van Geert, P. (2019). Children's creativity: a theoretical framework and systematic review. *Review of Educational Research, 89* (1), 93-124. https://doi-org.proxy-ub.rug.nl/10.3102/0034654318815707

Meindertsma, H., Van Dijk, M. W. G., Steenbeek, H., & Van Geert, P. L. C. (2012). Application of Skill Theory to compare scientific reasoning of young children in different tasks. *Netherlands Journal of Psychology, 67*(1), 9-19.

Menninga, A., van Dijk, M., Steenbeek, H., & van Geert, P. (2017). Language use in real-time interactions during early elementary science lessons: The bidirectional dynamics of the language complexity of teachers and students. *Language Learning, 67*(2), 284-320. https://doi.org/10.1111/lang.1221

Reeve, J. (2006). Teachers as facilitators: What autonomy-supportive teachers do and why their students benefit. *The elementary school journal, 106*(3), 225-236. https://doi-org.proxy-ub.rug.nl/10.1086/501484

Roe, M.F. (1997). Combining enablement and engagement to assist students who do not read and write well. *Middle School Journal, 28*, 35-41. http://www.jstor.org/stable/23024751

Ryan, R.M. & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Publications. https://ebookcentral.proquest.com/lib/rug/detail.action?docID=4773318

Ryan, R., & Niemiec, C. (2009). Self-determination theory in schools of education: Can

an empirically supported framework also be critical and liberating. *Theory and Research in Education, 7*(2), 263–272. https://doi-org.proxy-ub.rug.nl/10.1177/1477878509104331

Sahlberg, P. (2009). The role of education: potential barriers and enabling factors. In: Villalba, E. (Eds.) *Measuring creativity* (pp. 337-345). Luxembourg: Publications Office of the European Union.

Schippers, H. (1996). Teaching world music in the Netherlands: towards a model for cultural diversity in music education. *International Journal of Music Education*, (1), 16-23. https://doi.org/10.1177/025576149602700103

Steckler, A., & McLeroy, K. R. (2008). The importance of external validity. *American journal of public health, 98*(1), 9-10. https://doi.org/10.2105/AJPH.2007.126847

Steenbeek, H., & van Geert, P. (2013). The emergence of learning-teaching trajectories in education: A complex dynamic systems approach. *Nonlinear dynamics, psychology, and life sciences, 17*(2), 233-267.

Sternberg, R. J. (2010). Teaching for creativity. In: R. A. Beghetto & J. C. Kaufman (Eds.), Nurturing creativity in the classroom (pp. 394–414). Cambridge University Press. https://doi.org/10.1017/CBO9780511781629.020

Teel, K.M., Debruin-Parecki, A., & Covington, M.V. (1998). Teaching strategies that honor and motivate inner-city African-American students: A school/university collaboration. *Teaching & Teacher Education, 14*, 479-495. https://doi.org/10.1016/S0742-051X(98)00001-8

Todman, J. B., & Dugard, P. (2001*). Single-case and Small-n Experimental Designs: A Practical Guide To Randomization Tests (1st ed).* Psychology Press. https://doi.org/10.4324/9781410600943

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam med, 37*(5), 360-363.

Van Geert, P., & Van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development, 25*(4), 340-374. https://doi.org/10.1016/S0163-6383(02)00140-6

Van der Steen, S., Steenbeek, H.W., van Dijk, M.W.G. & van Geert, P.L.C.(2014). A process approach to children's understanding of scientific concepts: A longitudinal case study. *Learning and Individual Differences,* 8-19. https://doi.org/10.1016/j.lindif.2013.12.004

Van Vondel S., Steenbeek H., van Dijk M., van Geert P. (2016) "Looking at" Educational Interventions: Surplus Value of a Complex Dynamic Systems Approach to Study the Effectiveness of a Science and Technology Educational Intervention. In: Koopmans M., Stamovlasis D. (eds) *Complex Dynamical Systems in Education*. Springer, Cham. https://doi.org/10.1007/978-3-319-27577-2_11

Van Vondel, S., Steenbeek, H., van Dijk, M., & van Geert, P. (2017). Ask, don't tell; A complex dynamic systems approach to improving science education by focusing on the co-construction of scientific understanding. *Teaching and Teacher Education, 63*, 243-253. https://doi.org/10.1016/j.tate.2016.12.012

Webster, P. (1990). Creativity as creative thinking. *Music Educators Journal, 76* (9), 22-28. https://doi.org/10.2307/3401073

Wetzels, A., Steenbeek, H., & Van Geert, P. (2016). A Complexity Approach to Investigating the Effectiveness of an Intervention for Lower Grade Teachers on Teaching Science. *Complicity: An International Journal of Complexity and Education*, *13*(1), 81-104. https://doi.org/10.29173/cmplct23022

Young-Jones, A., Cara, K. C., & Levesque-Bristol, C. (2014). Verbal and behavioral cues: creating an autonomy-supportive classroom. *Teaching in Higher Education, 19*(5), 497–509. https://doi.org/10.1080/13562517.2014.880684

Appendix A

Table 2.

Baseline, intervention and post- frequencies and percentages of teachers ASV 1-8

			Teacher 7	Teacher 8	**Teacher 9**	Teacher 10	Teacher 11	**Teacher 12**
ASV Levels		*Research phase*	*N (%)*	*N (%)*	*N (%)*	*N (%)*	*N (%)*	*N (%)*
	1 Stop	*Baseline*	4 (2.30)	1 (0.85)	-	-	-	-
		Inter.	5 (1.48)	1 (0.36)	-	-	-	2 **(1.02)**
		Post	1 (0.59)	2 (1.36)	-	-	-	6 **(5.17)**
LOW	2 Instruc-tion	*Baseline*	18 (10.35)	22 (18.80)	13 **(14.94)**	1 (1.56)	*5* (5.68)	1 **(1.39)**
		Inter.	30 (8.88)	18 (6.48)	8 **(4.35)**	4 (1.82)	*3* (1.36)	6 **(3.06)**
		Post	17 (10)	11 (7.48)	-	1 (0.92)	1 (1.92)	2 **(1.72)**
	3 Explan-ation	*Baseline*	79 (45.40)	51 (43.59)	40 **(45.98)**	39 (60.94)	36 (40.91)	29 **(40.28)**
		Inter.	125 (36.98)	111 (39.93)	73 **(39.67)**	65 (29.55)	70 (31.82)	66 **(33.67)**
		Post	66 (38.82)	65 (44.22)	21 **(24.71)**	48 (44.04)	19 (36.54)	55 **(47.41)**
MED-IUM	4 Teacher centered question	*Baseline*	27 (15.52)	26 (22.22)	8 **(9.20)**	4 (6.25)	7 (7.95)	8 **(11.11)**
		Inter.	52 (15.39)	57 (20.50)	13 **(7.07)**	19 (8.64)	13 (5.91)	11 **(5.61)**
		Post	22 (9.4)	15 (10.20)	5 **(5.88)**	20 (18.35)	1 (1.92)	12 **(10.35)**
	5	*Baseline*	21 (12.07)	7 (5.92)	21 **(24.14)**	10 (15.63)	19 (21.59)	16 **(22.22)**

	Student centered question	*Inter.*	48 (14.20)	24 (8.63)	44 **(23.91)**	60 (27.27)	50 (22.73)	47 **(23.98)**
		Post	28 (16.47)	19 (12.93)	5 **(5.88)**	24 (22.02)	11 (21.15)	36 **(31.03)**
	6 Cognitive musical question	*Baseline*	5 (2.87)	10 (8.54)	2 **(2.30)**	6 (9.38)	2 (2.27)	13 **(18.06)**
		Inter.	30 (8.88)	11 (3.96)	27 **(14.67)**	26 (11.82)	24 (10.91)	15 **(7.65)**
		Post	7 (4.12)	-	23 **(27.06)**	8 (7.34)	8 (15.39)	3 **(2.59)**
HIGH	7 Creative musical question	*Baseline*	11 (6.32)	-	2 **(2.30)**	3 (4.69)	12 (13.63)	5 **(6.94)**
		Inter.	38 (11.24)	56 (20.14)	10 **(5.43)**	28 (12.73)	56 (25.45)	41 **(20.92)**
		Post	27 (15.88)	20 (13.61)	15 **(17.65)**	6 (5.51)	10 (19.23)	2 **(1.72)**
	8 Encouragement	*Baseline*	9 (5.17)	-	1 **(1.15)**	1 (1.56)	7 (7.95)	-
		Inter.	10 (2.96)	-	9 **(4.89)**	18 (8.18)	4 (1.82)	8 **(4.08)**
		Post	2 (1.18)	15 (10.20)	16 **(18.82)**	2 (1.83)	2 (3.85)	2 **(1.72)**

Note. Teacher results reflected upon in the qualitative analysis were boldened.

Appendix B

Table 4.

Baseline, intervention and post- frequencies and percentages of Class CTA 1-9

			Class 7	**Class 8**	Class 9	**Class 10**	Class 11	Class 12
CTA Levels		*Research phase*	*N (%)*	*N (%)*	*N (%)*	*N (%)*	*N (%)*	*N (%)*
SENSORY MOTOR ACTIONS	1 Single	*Baseline*	-	-	-	-	-	-
		Inter.	-	-	49 (4.89)	-	4 (0.72)	
		Post	-	-	-	-	-	-
	2 Mapping	*Baseline*	12 (2.63)	35 **(7.61)**	8 (3.14)	3 **(0.90)**	75 (35.38)	9 (2.88)
		Inter.	7 (0.97)	53 **(5.56)**	81 (8.08)	17 **(2.98)**	65 (11.82)	18 (2.07)
		Post	-	22 **(4.09)**	-	-	27 (12.98)	26 (7.95)
	3 Systems	*Baseline*	38 (8.33)	38 **(8.26)**	23 (9.02)	20 **(5.97)**	32 (15.09)	23 (7.35)
		Inter.	75 (10.35)	54 **(5.66)**	37 (3.69)	17 **(2.98)**	80 (14.55)	15 (2.07)
		Post	4 (0.87)	38 **(8.26)**	24 (5.57)	-	14 (6.73)	3 (0.92)
RESPRESENTATIONS	4 Single	*Baseline*	20 (4.39)	61 **(13.26)**	15 (5.88)	77 **(22.99)**	3 (1.41)	5 (1.60)
		Inter.	100 (13.79)	221 **(23.17)**	138 (13.77)	94 **(16.46)**	122 (22.18)	21 (2.89)
		Post	35 (7.64)	149 **(27.70)**	2 (0.46)	7 **(1.68)**	43 (20.67)	38 (11.62)
	5 Mapping	*Baseline*	141 (30.92)	109 **(23.70)**	68 (26.67)	161 **(48.06)**	60 (28.30)	94 (30.03)
		Inter.	223	146	179	214	160	306

			(30.76)	**(15.30)**	(17.86)	**(37.48)**	(29.09)	(42.15)
		Post	159	296	32	149	124	113
			(34.72)	**(55.02)**	(7.43)	**(35.73)**	(59.62)	(34.56)
	6	*Baseline*	21	94	17	20	8	78
	Syste		(4.61)	**(20.43)**	(6.67)	**(5.97)**	(3.77)	(24.92)
	ms	*Inter.*	77	480	63	72	18	78
			(10.62)	**(50.31)**	(6.29)	**(12.61)**	(3.27)	(10.74)
		Post	61	27	75	155	-	55
			(13.32)	**(5.02)**	(17.40)	**(37.17)**		(16.82)
	7	*Baseline*	39	61	-	54	-	-
	Single		(8.55)	**(13.26)**		**(16.12)**		
		Inter.	37	-	40	68	67	63
			(5.1)		(3.99)	**(11.9)**	(12.18)	(8.68)
		Post	50	24	39	155	-	-
			(10.92)	**(4.46)**	(9.05)	**(37.17)**		
	8	*Baseline*	153	51	124	-	34	91
ABSTRACT	Mappi		(33.55)	**(11.09)**	(48.63)		(16.04)	(29.07)
IONS	ng	*Inter.*	123	-	335	51	33	197
			(16.97)		(44.43)	**(8.93)**	(6)	(27.14)
		Post	114	-	209	52	-	15
			(24.89)		(48.49)	**(12.47)**		(4.59)
	9	*Baseline*	32	11	-	-	-	13
	Syste		(7.02)	**(2.39)**				(4.15)
	ms	*Inter.*	83	-	80	38	1	28
			(11.45)		(7.98)	**(6.66)**	(0.18)	(3.86)
		Post	35	-	50	-	-	77
			(7.64)		(11.6)			(23.55)

Note. Student results reflected upon in the qualitative analysis were boldened.