The relationship between instructional differentiation, student diversity and academic-engagement: A pilot observation study

Susen R. Smith, Greg Robinson, Michael Arthur-Kelly and Phil Morgan
University of Newcastle, Australia

Introduction

Issues underlying this PhD study involved concerns regarding literacy development, student diversity, instructional differentiation, methods used to observe teaching and learning practices and academic-engagement\(^1\) (Conway, Arthur-Kelly & Pascoe, 2004; Luke, Freebody, & Land, 2000). Even though there has been a focus on literacy in New South Wales (NSW) education, many students still fail to acquire basic literacy skills (Brent, Gough & Robinson, 2001; Masters & Forster, 1997; NSW Department of Education & Training, 1997; Rohl & Milton, 2002).

Instructional differentiation has gained prominence in recent years to address the individual and diverse needs of students within primary classrooms (Tomlinson, 2001). However, concerns regarding differentiated practices have been raised in the literature (Hart, 1992; Westwood, 2001).

Observation of classroom practice has evolved to investigate instruction in a variety of contexts (Logan, Bakeman, & Keefe, 1997; Ysseldyke & Christenson, 1994). This exploratory study entailed observing and examining student behaviour, teacher behaviour and classroom ecology and reports an investigation of differentiated instructional contexts of primary literacy blocks. Specifically, the pilot study examined relationships between classroom ecology and academic-engagement.

\(^1\) KEYWORDS: Literacy, classroom ecology, differentiated instruction, student diversity, academic-engagement.
teacher instruction for students with low-, average- and high-reading ability and students’ academic-engagement. Some observed events included students’ academic or competing behaviours, teacher’s instruction and management and classroom ecology events such as implementation processes and instructional groupings.

The Code for Instructional Structure and Student Academic Response - Mainstream Version was used in this study (MS-CISSAR; Carta, Greenwood, Schulte, Arreaga-Mayer & Terry; 1989). The MS-CISSAR is a widely-used observation tool that was modified to include aspects of differentiated instruction, and field-tested by observing targeted students from three primary school classrooms. A summary of results for all students combined and students with low-, average- and high-reading ability will be provided. Some implications for the main study will also be presented.

Aims of the pilot observational study

The aims were to:

i) amend a published observation instrument to explore observed classroom practice,

ii) field-test, analyse and assess the reliability of the observation instrument,

iii) identify the implications for the research process.

Research questions for the pilot observation study

i) Were there differences in frequency of instruction for students with different levels of ability? If so, what were they?

ii) Were there differences in relationships between the most frequently coded variables and academic-engagement? If so, what were they?

iii) What were the implications for the research process?

The method will now be outlined.
Method
The MS-CISSAR was modified to include aspects that could be used to effectively
differentiate instruction. The observation tool was initially field-tested in an infants class and
then by observing six targeted students in each of three classrooms in one New South
Wale’s primary school.

Observation procedure for the pilot study
The observation procedure for the pilot observation study included:

i) the amendment of the observation tool and process following trial observations
in an infants class,

ii) on-going inter-observer reliability checks,

iii) observations of six randomly selected students with varying reading ability
from each of three classes coded A, B and C – Year 2, Year 4 and Year 6
respectively,

iv) collection of targeted student work samples,

v) analyses of results, which resulted in further refinements to the observation tool
and research process for the main study.

The amendments to the MS-CISSAR observation tool for the pilot study
Following an in-depth literature search and review, the MS-CISSAR was chosen to assess
classroom practice. The researcher developed an observation data checksheet with the three
main categories from the original MS-CISSAR retained: i) student behaviours; ii) teacher
behaviours, and iii) classroom ecology. Student behaviour focused on student behaviours
and responses in relation to the ecological and teacher instructional contexts in which they
were positioned. Teacher behaviour identified who was providing the instruction, their focus
in relation to the target student and how they were instructing, managing or reinforcing
student responses. Classroom ecology included the context, processes, educational
materials and environment within which the students were taught, that influenced teacher
instruction and student behaviours.

Susen R. Smith, PhD candidate, University of Newcastle ©.
After trialling in an infants class, modifications to the MS-CISSAR were made and 99 codes within 13 sub-categories remained for the pilot study. However, some sub-categories and events then differed from the original (refer to Table 1 for the amended sub-categories, codes and descriptions).

Table 1  Categories, sub-categories, descriptions and code examples for the MS-CISSAR pilot observation study

<table>
<thead>
<tr>
<th>Categories &amp; sub-categories</th>
<th>Code or event no.</th>
<th>Description of sub-categories</th>
<th>Codes / events relevant to the pilot study</th>
</tr>
</thead>
<tbody>
<tr>
<td>STUDENT BEHAVIOUR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic responses</td>
<td>6</td>
<td>specific, active response</td>
<td>writing, task participation, reading aloud, reading silently, talk academically, no academic response</td>
</tr>
<tr>
<td>Task management responses</td>
<td>7</td>
<td>prerequisite or enabling response</td>
<td>raise hand, play appropriately, manipulate materials, moving, talk management, attention, no management</td>
</tr>
<tr>
<td>Competing responses</td>
<td>6</td>
<td>response that competes with academic response</td>
<td>aggression, disruption, non-compliance, talk inappropriately, looking around, no inappropriate behaviours</td>
</tr>
<tr>
<td>TEACHER BEHAVIOUR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approval/disapproval</td>
<td>3</td>
<td>approval/disapproval of student behaviour</td>
<td>approval, disapproval, neutral</td>
</tr>
<tr>
<td>Instructor</td>
<td>9</td>
<td>person teaching target student</td>
<td>regular class teacher, special ed., aide, student teacher, volunteer, related server, substitute, peer, none</td>
</tr>
<tr>
<td>Focus</td>
<td>5</td>
<td>to whom teacher behaviour is directed</td>
<td>whole class/group, other/s in whole class, other/s in small group, target individual, no one</td>
</tr>
<tr>
<td>Instruction/management responses</td>
<td>15</td>
<td>teacher’s responses relative to target student</td>
<td>question/academic, question/management, question/discipline, command academic, command/management, command/discipline, talk/academic, talk/management, talk/discipline, talk/non-academic, non-verbal prompt/cue, attention, reading aloud, singing, no response</td>
</tr>
<tr>
<td>Position</td>
<td>6</td>
<td>teacher position relative to target student</td>
<td>in front of room, at side of room, beside target student, at own desk, at back of room, out of room</td>
</tr>
<tr>
<td>CLASSROOM ECOLOGY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organised activity</td>
<td>13</td>
<td>topic/content of instruction</td>
<td>business/management, silent reading, reading to/with student/s, encoding/decoding, sight words, comprehension, writing, language, reading activity, spelling, handwriting, other, no organised activity</td>
</tr>
<tr>
<td>Implemented process</td>
<td>11</td>
<td>process of instruction</td>
<td>reviewing, outlining lesson content, presenting content, modelling/demonstrating, guiding practice, independent practice/learning, monitoring/feedback, reinforcing/praising, transition, other, no implemented activity</td>
</tr>
<tr>
<td>Physical arrangement</td>
<td>3</td>
<td>seating arrangement</td>
<td>entire group, divided group, individual</td>
</tr>
<tr>
<td>Instructional grouping</td>
<td>5</td>
<td>instructional patterns</td>
<td>whole class, small group, one-on-one, independent, no instruction</td>
</tr>
<tr>
<td>Specific Task</td>
<td>10</td>
<td>curriculum materials</td>
<td>reader/s, workbooks, worksheet, exercise book/pen/pencil/paper, listen/lecture, discussion, other media, computer, fetch/put, no task</td>
</tr>
</tbody>
</table>
Student variables included 19 codes within 3 sub-categories - academic responding, task management and competing behaviours. Teacher variables comprised 38 codes within 5 sub-categories - approval, instructor, focus, instruction/management and position. The classroom ecological variables included 42 codes within 5 sub-categories - organised activity, implemented process, physical arrangement, instructional grouping and specific task. The implemented process sub-category was added to provide greater depth to the analysis of the classroom ecology category, to match the instrument more closely with the effective instruction and differentiated instruction literature and for later comparisons to the results of the survey.

Data collection procedure

The overall data collection procedure for the pilot study involved the researcher and a trained assistant observing three classes, across three stages of learning, in the one primary school. One primary school was randomly selected from a stratified sample, that is, from one region of NSW primary schools. When the Principal of the first selected school was contacted he agreed to participate. Teachers were invited to participate until three teachers agreed, which resulted in three classes across three stages of learning. The participating teachers were then asked to divide their class into three student groups according to the students’ level of reading ability and randomly select six students – two from the low-reading ability group, two from the average-reading ability group and two from the high-reading ability group. All participants were provided with correspondence informing them about the study and to obtain their permission for participation. The final participants were three teachers, one male and two females, with varying degrees of teaching experience and qualifications and 18 students from three different grades – Year 2, Year 4 and Year 6.

The 18 target students’ reading abilities were assessed by the researcher using formal assessment tools such as the Neale Analysis of Reading Ability (Neale, 2000). The results
confirmed the students correct categorisation into the three student ability groups. Teachers were aware of the dates for all training and observation sessions, but students weren’t.

**Training of the data collectors, trialling the observations and reliability checks**

Two observers were trained for the observations using clinical trials, videos of classroom practice, the MS-CISSAR instructional manuals and observing an infants class. Using the observation checksheet, the researcher and research assistant subsequently trained for several short sessions and for three full sessions in the pilot classrooms to be observed. Estimates of inter-reliability were undertaken at randomly selected intervals (Alberto & Troutman; 1999; Gay, 1992).

**Observation and analyses of the three classrooms**

Observations occurred across five weeks of the first term of the school year. The modified version of the MS-CISSAR was used to observe and assess accumulated response events for targeted students, teacher events and ecology events within three 1¾ hour literacy blocks in each of the three classes. The method of data collection included momentary time sampling of multiple events across the three categories: i) student variables, ii) teacher variables and iii) classroom ecology (Greenwood & Delquadri, 1988). Momentary time sampling involved adhering to three cues within a 50 second interval to observe one student and code the student’s responses in relation to teacher and ecology events. Consecutively, the next student was observed within the next 50 second interval and so forth. All six students and associated teacher and ecological variables were observed and recorded cyclically within a 300 second or 5 minute timeframe. At the conclusion of each 1¾ hour observation period student work samples were collected and photocopied².

---

² These were then used to identify if the process of collection would be effective and to see if analysis could assess differentiated content.
Recording and analyses of data were undertaken using the *Statistical Package for the Social Sciences* (SPSS version 11.0). The data were analysed by combining the intervals of each student, teacher and ecology category within each of the three 1¾ hour observation periods. Percentage scores were tabulated for each of the student, teacher and ecology events for all students combined and cross-tabulated with student categories of low-, average- and high-reading ability. Data analyses included descriptive data analyses of frequencies of events for all students combined and for students with LRA, ARA and HRA. Crosstabulations of relationships between most frequently coded events and academic-engagement for all students combined and students with LRA, ARA and HRA were also undertaken. Only a summary of the most frequently occurring events analysed are presented for the pilot study. Most frequent refers to those events greater than 10% in frequency or the three to five most frequently coded events in each sub-category as recommended by Greenwood, Delquadri, Stanley, Terry and Hall (1985).

Each of the 18 target students was observed for 60 intervals. Each interval was considered a data point, therefore there were 1080 data points altogether or 360 data points for each student category observed. Inter-observer reliability estimates were obtained in each class following several short reliability checks. Two observers recorded simultaneously and continuously in the pilot phase and 25% of their observations were randomly selected and calculated for estimates of inter-observer reliability. The random selection process occurred for each of the nine observations and involved selecting 25% of the numbers from one to 120 out of a container, as 120 intervals were observed during each 1¾ hours. The reliability estimates were checked on all observations by dividing the number of agreements by interval, by the number of agreements, plus disagreements multiplied by 100 as expressed in the following formula:

\[
\text{Reliability} = \frac{\text{Number of agreements}}{\text{Number of agreements} + \text{Number of disagreements}} \times 100
\]

Results are expressed as percentages here and will be subject to more complex statistical analysis in the main study. The purpose here is to identify key differences and patterns.
The results of the nine inter-observer reliability checks ranged from 85% to 96% reliability. Some results will now be provided.

Results of the pilot observations

The relationships between the some of the most frequent student, teacher and ecology variables for all students combined and students with low-, average- and high reading ability are presented. The relationships between the most frequent teacher and ecology variables and the level of student academic-engagement are also provided.

Student academic and competing behaviours

The most frequently observed student academic and competing behaviour events will now be reported. The preliminary descriptive analyses suggested the most frequently coded student academic response was writing (32.4%), with a large portion of non-academic responding (43.4%) (refer to Figure 1 for the results). The LRA students were recorded responding less academically than their peers (no academic response: LRA: 47.2%; ARA: 41.7%; HRA: 41.4%), while ARA and HRA students were mainly involved in writing activities (LRA: 28.3%; ARA: 35.3%; HRA: 33.6%).

Overall, competing behaviours (10.9%) were minimal. However, there was a small proportion of looking around and students with LRA displayed this behaviour the most (ALL: 9.9%; LRA: 12.8%, ARA: 8.1%, HRA: 8.9%) (refer to Figure 2 for the results).
Figure 1 The most frequently observed student academic-behaviours

Figure 2 The most frequently observed student competing behaviours
Teacher behaviours and classroom ecology and student academic-engagement

The most frequently observed teacher and ecology events will now be presented in relation to student academic-engagement.

Teacher behaviours and student academic-engagement

The most frequently observed teacher instructor, instruction and academic-engagement will now be reported.

Teacher as instructor and academic-engagement

Overall, the teacher was coded as the main instructor for 97.4% of the time the instructor was observed. When engagement was investigated, HRA students (57.6%) were coded as more engaged than their peers, while the teacher was the instructor and LRA students (51.6%) were coded the least engaged (refer to Figure 3).

![Figure 3 Most frequently observed instructor events and academic-engagement](image-url)
Teacher instruction and academic-engagement

In regards to teacher instruction, teachers were observed using academic talk (19.4%) or questioning (15.5%) predominantly during instruction, with a large amount of non-response (18.4%) (refer to Figure 4). Teachers were coded directing academic questions consistently to LRA students (LRA:12.8%; ARA: 11.1%; HRA: 12.2%), attention (ALL: 12%; LRA: 12.2%; ARA: 17.8%; HRA: 16.4%) and academic talk (LRA: 18.3%; ARA: 21.9%; HRA: 18.1%) to ARA students and disciplinary talk to HRA students (ALL: 9.9%; LRA: 10%; ARA: 7.5%; HRA: 12.2%).

Figure 4 Most frequently observed teacher instruction events

Students with LRA were observed as less engaged than their peers during all instructional contexts, except when the teacher was talking to the class academically (refer to Figure 5).
Classroom ecology and student academic-engagement

The most frequently observed grouping events, implementation processes and task material events will be presented in relation to student academic-engagement.

Grouping and academic-engagement

In the instructional ecology context, the classes seemed to be arranged and taught more often in whole class groupings (whole class: 56%, independent learning: 24%), with the classroom teacher (97.4%) as the instructor positioned at the front of the class (ALL: 58.8%, LRA: 58.3%, ARA: 57.8%, HRA: 60.3%) (refer to Figure 6).
Students with LRA were coded as more engaged than their peers during small group (ALL: 76.6%; LRA: 82.8%; ARA: 62.5%; HRA: 70%) situations, ARA students were observed as more engaged than their peers during independent activities (ALL: 81%; LRA: 82.4%; ARA: 83.2%; HRA: 78.7%), while HRA students seemed more engaged than their peers during whole class instruction (ALL: 40.2%; LRA: 35.9%; ARA: 40.9%; HRA: 43.9%) (refer to Figure 7).
Implemented processes and academic-engagement

Independent practice/learning (38.4%) was the most frequently observed implemented process. Other frequently coded processes included guided practice (12.3%), presenting content (11.9%), transition (10.6%) and other activities (18.4%), such as assessment or using contracts (refer to Figure 8).

Figure 7 Most frequently observed instructional grouping events and academic-engagement

Figure 8 Most frequently observed implemented process events
The LRA students were observed being guided (LRA: 16.1%; ARA: 10.8%; HRA: 10%) through activities or in transition (LRA: 11.9%; ARA: 8.6%; HRA: 11.4%) more than their peers, ARA students were coded working independently (LRA: 33.3%; ARA: 43.3%; HRA: 38.6%) more often than their peers and HRA students were coded more often when involved in other activities (LRA: 18.6%; ARA: 16.7%; HRA: 20%).

All students were observed as quite engaged during independent learning opportunities (82.9%) (refer to Figure 9). However, students with LRA presented as more engaged during guided practice (ALL: 55.7%; LRA: 67.2%; ARA: 51.3%; HRA: 50%), but less engaged than their ARA and HRA peers during independent (LRA: 81.7%; ARA: 83.3%; HRA: 83.5%) or other activities (ALL: 50.8%; LRA: 41.9%; ARA: 48.6%; HRA: 62.5%) or when the teacher was presenting content (ALL: 48.4%; LRA: 38.6%; ARA: 55%; HRA: 52.3%).

Figure 9  Most frequently observed instructional process events and academic-engagement
Specific tasks or materials and academic-engagement

Overall, the students were observed mostly using worksheets (29.8%), with frequent use of exercise books (19.4%), readers (12.3%), other media (15.6%), such as games, or having no organised materials (11.4%) (refer to Figure 10). The LRA students were coded using readers (LRA: 13.6%; ARA: 11.1%; HRA: 12.2%) or having no specific task materials (LRA: 12.8%; ARA: 10.3%; HRA: 11.1%) more often than their peers, ARA students were observed using worksheets (LRA: 27.2%; ARA: 32.8%; HRA: 29.4%) and exercise books (LRA: 19.4%; ARA: 20%; HRA: 18.6%) more frequently than other students, while HRA students used other media more regularly than LRA or ARA students (LRA: 15.8%; ARA: 14.4%; HRA: 16.4%).

![Figure 10: Most frequently observed specific task or materials events](image)

Figure 10 Most frequently observed specific task or materials events

There seemed to be more engagement for all students together when using exercise books or worksheets (79.9% and 76.4% respectively) (refer to Figure 11). Students with LRA were coded as more engaged using readers (ALL: 62.4%; LRA: 65.3%; ARA: 60%; HRA: 61.4%) or other media (ALL: 50.6%; LRA: 52.6%; ARA: 50%; HRA: 49.2%) than students in other
ability groups. Students with ARA were observed as more engaged than the other target students when using *exercise books* (ALL: 79.9%; LRA: 70%; ARA: 86.1%; HRA: 83.6%), while students with HRA were more engaged than their peers when using *worksheets* (ALL: 76.4%; LRA: 74.7%; ARA: 74.6%; HRA: 80.2%).

![Bar chart showing specific task or materials and engagement]

**Figure 11 Most frequently observed specific task or materials events and academic-engagement**

**Collection of the student work samples**

Student work samples were collected at the conclusion of each classroom observation. Content analysis of collected student work samples suggested there was some differentiation in the form of *different levels of reading materials, contracts or activity booklets, tutoring* and *guided reading in small groups*. However, most activities were observed taking place in whole class instructional contexts, using one level text or worksheet. Some implications for the main study will now be presented.

**Implications**

The results provided the basis for further refinement of the observation tool and research
process, but did the results address the research aims and questions? The implications for the main study related to enlarging the cohort, streamlining the research process, amending the observation tool, obtaining additional outcomes data and refining terminology. These recommendations will now be described.

**Did the results address the aims and answer the research questions?**

The results suggested some differences in frequencies of events within sub-categories and between the student ability groupings, providing some indication that instruction was differentiated for different student ability levels. However, results also suggested instructional differentiation was minimal, even though some different relationship patterns relating to students’ engagement responses, teaching behaviours and classroom ecology, were observed across the student ability levels.

**Enlarging the cohort**

While the aim of the pilot study was to explore the feasibility of the research design with a smaller cohort, the study only involved three classes in one school, but these classes were across three different stages of learning. While stages reflect common practice in NSW schools, instructional practice in each stage can be very different. Therefore, the findings may not be an accurate reflection of teachers’ practice or students’ responses within one particular stage. Hence, the main study involved a broader sample across two schools, but within one stage of learning, that is, Stage 2.

**Streamlining the research process**

Amendments to the research process were needed to ensure more effective procedures ensued. Specific amendments to the research process included streamlining coding for inter-reliability checks through continued co-coding trialling and on-going formative debriefing sessions. More complex statistical analyses of the main study’s results to identify significant
differences between variables investigated were undertaken.

**Amending the observation tool and descriptors**

The MS-CISSAR has had extensive validation and has been used in a number of studies (e.g. Carta, Greenwood, Schulte, Arreaga-Mayer, & Terry, 1988; McDonnell, Mathot-Buckner, Thorson & Fister, 2001; Wallace, Reschly-Anderson & Hupp, 2002). Amended versions have also been used in other studies and it is generally considered to be a reliable tool (e.g. McDonnell, Thorson & McQuivey, 1998; Conway, Arthur-Kelly & Pascoe, 2004).

The initial amended version of the MS-CISSAR enabled the pilot research questions to be addressed. However, the results suggested further amendments to the instrumentation and research process were required. Difficulties in using the MS-CISSAR were noted in a small scale pilot study by Conway, Arthur-Kelly and Pascoe (2004), who found no discernible differentiated instruction for students with varying abilities. However, their study compared two classes, whereas this pilot study grouped three classes together and did find some possible differences. In order to accommodate the limitations in using the MS-CISSAR, the researcher enlarged the data set, increased the training sessions, examined the position of observers and the process of observations was trialled consistently, before, during and after the pilot, until these concerns were reduced. Additionally, anecdotal, survey, interview, work samples, outcomes and assessment data were included in the main study.

The observational assessment instrument needed to be modified as there were items within it that were not required to answer the research questions for the main study. Planned rest periods were needed and this impacted on the timing of the intervals, which, in turn, had to be reduced. The time to record the observations needed to be trialled again, to see if a reduction from 50 seconds to 40 seconds could be accomplished. The on-going training of the inter-observers was used as the period to trial the timing again. Reliability estimates of
inter-observer agreement protocols continued. While a second observer was used to simultaneously and consistently co-coded in the pilot phase, a second observer was needed to alternatively co-code for both agreement checks on at least 25% of collected data. This process aimed to relieve the primary observer to minimise potential for observer bias and drift in the main phase. These changes required a larger focus on the training and trialling portion of the main study to ensure a more refined preparation for the final observations.

**Obtaining additional outcomes data**

As the main study was also looking at the relationship between syllabus outcomes and practice, the data needed to link more closely with outcomes-based data. Therefore, an outcomes checklist for teachers to complete prior to each observation, student portfolio items and class assessments were included. This enabled a closer link with the English syllabus and reading outcomes, as students were grouped according to levels of reading ability.

**Redefining terminology**

The categorisation of students into low-, average- and high-reading ability was used for the pilot study. Teachers often divide their classes in such a way when implementing group activities and standardised testing and collected student work samples supported this classification. However, support needs is the current focus of the NSW Department of Education and Training (1997). As the terms high-, low- and average-support needs are not widely used, and teachers could identify with ability levels, a combination of terms ability/support needs was used for the main study.

**Conclusion**

This paper provided an outline of the aims, research questions, method, a summary of the results of a pilot observation study and the implications for the main study. The results suggested that observation, used within this methodology, may indicate some differentiation...
between different student ability groups observed and identify relationship patterns between variables. However, the results also suggested further amendments and more observer training were needed for the next phase of the study and additional methods were needed for a more comprehensive overview of perceived practice.

The findings suggested the instruments and process were reliable and could be used to investigate the frequency and perceptions of instructional differentiation and academic-engagement across different student categories in regular classrooms. However, several refinements to instrumentation and procedures were identified and implemented for the main study. Implications related to the need to enlarge the cohort, streamline the research process, further refine the observation tool, attain additional outcomes data and refine some terminology. Other parts of the study included a survey questionnaire and interviews to investigate teachers’ perceptions of instructional practice for student diversity. The amendments to the research process and instruments enabled the investigation of the relationships between differentiated instructional strategies, students with low-, average- and high-reading ability/support needs, academic-engagement and learning outcomes.

Thank you

Thank you to my University PhD supervisors Associate Professor Greg Robinson, Dr Michael Arthur-Kelly and Dr Phil Morgan, who guide my progress with diverse wisdom and invaluable expertise. Thankyou also to Mr Raymond Smith, for his exceptional research assistance, consistent support and encouragement.

For more information contact

Email: Susen.smith@newcastle.edu.au or Phone: +61 2 43 822936
References


