University of Groningen

SEN and the art of teaching
van der Kamp, Antoinette Jacqueline

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 11-01-2020
Chapter 3

Systematic academic instruction for students with EBD: the construction and use of a tool for teachers

This chapter is a slightly adapted version of:
3.1 Introduction

As indicated in Chapter 1, students with emotional and behavioural difficulties (EBD) are a serious challenge to education systems (Cooper & Jacobs, 2011). Such students not only show a wide range of externalising and internalising behavioural problems, numerous of these students also show very little progress in their academic learning over the course of a full academic year (Siperstein et al., 2011; Yell, Meadows, Drasgow, & Shriner, 2009). Moreover, the developmental delays that many students with EBD experience compared with the development of typical students increase rapidly over the years (Ledoux et al., 2012), including negative emotions about learning (Al-Hendawi, 2012). Although not all students with EBD show this gap in academic progress, it is a matter of continuing concern. Due to governmental pursuits of higher academic outcomes for all students (e.g., US: 'No Child Left Behind'; UK: 'Every Child Matters'; the Netherlands: 'Tailored Education'), the focus on academic instruction for these students in particular is growing. Given the fact that a large degree of problem behaviour is precisely related to problems in academic learning (Umbreit et al., 2007), academic instruction for EBD students is clearly worth studying to develop and implement efficacious interventions to tackle these difficulties.

Much problem behaviour observed in classrooms appears to originate from a discrepancy between the demands of the tasks offered and the skills of students with EBD (Lewis, Hudson, Richter, & Johnson, 2004; Umbreit et al., 2007). Kern and Clemens (2007) note that: “Frequently, problem behaviours result from a mismatch between the environment and a student’s skills, strengths, or preferences. For instance, work assignments that are too difficult for a student are a common cause of problem behaviour in the classroom. Appropriately matching instruction to a student’s skill and performance corrects this environmental problem” (2007, p. 66).

In Chapter 2 we concluded that, given the diversity of types and needs of students with EBD, knowing and anticipating on these needs is the first step to success. We referred to this as the key to teaching academics and stated that the curriculum continuously has to be adjusted to the needs of the student in an ongoing process. This calls for systematically planned academic instruction that is carefully orchestrated and appropriately adapted to fit each unique student’s needs (Simpson, Peterson, & Smith, 2011). Deming (1986) proposed that processes of planning should be placed in a feedback loop, making it
possible to change the parts of a plan that was not working (did not match) and needed improving. He created a diagram to illustrate this iterative, on-going process, commonly known as the Plan-Do-Check-Act (PDCA) cycle. This widely used popular concept, frequently used for improving the quality of education (Kartikowati, 2013), is one of the core tactics in Dutch education (The Dutch Inspectorate of Education, 2011).

Another important aspect of academic instruction and directly related to school performance outcomes of students with EBD is instructional time (Kurz, Talapata, & Roach, 2012; Matheson & Shriver, 2005; Vannest & Hagan-Burke, 2010). Winn, Menlove and Zsiray (1997) stated that the link between instructional time and learning is one of the most consistent findings in educational research. The time a teacher spends on academic instruction is in inverse proportion to the time a teacher spends correcting misbehaviour in the classroom (Berliner, 1988; Brophy & Good, 1984). Thus, much precious instruction time for students with EBD is lost because teachers pay a great deal of attention to controlling behaviour (Pianta & Hamre, 2009; Wehby, Tally, & Falk, 2004). Confronted with too challenging tasks, students often develop problem behaviour that ‘helps’ them to avoid academic settings (Gunter & Coutinho, 1997; Scott et al., 2001). Teachers who tend to shift their attention from academic instruction to handling problem behaviour often just reinforce that behaviour (Sutherland & Oswald, 2005). Chapter 2, however, reveals that increasing students’ exposure to academic instruction could have demonstrably positive impact on classroom behaviour as well as the academic achievement of students with EBD (Brigham, Gustashaw, Wiley, & Brigham, 2004; van der Worp-van der Kamp et al., 2014). Surely, as stated by Kern, Hilt-Panahon and Sokol (2009), academic instruction is closely linked to behavioural instruction. We agree with Hagaman (2012) who warned against addressing academic learning and behaviour as separate issues. However, as long as teachers give too little attention to academic instruction (Vannest & Hagan-Burke, 2010), it still makes sense to consider this distinction and thereby to emphasise both.

The degree of systematic teaching and academic instruction seems to link closely to the problem behaviour of students with EBD. We define teachers’ activities as systematic if they are (1) planned with concern for students’ special needs (plan), (2) realised according to prior planning (do), (3) monitoring students’ progress regarding the defined goals (check) and (4) acting on the outcomes of the check phase (act). This general cycle of continuous improvement and adaptation can be filled with substantive educational content by using core aspects from ‘effective instruction’ as identified by Glaser (1962), Van Gelder, Oudkerk Pool, Peters, and Sixma (1973), Yell and Rozalski
(2013) (in Yell et al., 2009) and The Dutch Inspectorate of Education (2011). Reactive teaching styles, instant solutions, frequent changes of plan, following the curriculum blindly, activities without a clear goal and instruction that does not match the students' needs are all examples of a non-systematic approach. Furthermore, we define teachers' instruction as academic when they predominantly focus on academic content through explaining, motivating, asking and answering academic questions, whereas instructions concerning behaviour focus on explaining behavioural rules, reacting on behaviour, providing non-academic tasks as well as punishing and rewarding behaviour.

These two important aspects of instruction, 'degree of systematic teaching' and 'degree of academic instruction', may be visually displayed in a coordinate system (Figure 3.1), depicting instruction along the two dimensions. The y-axis displays the PDCA cycle. The x-axis involves a bipolar dimension: the more teachers focus on academic instruction, the higher they score on the x-axis. The more teachers focus on behaviour, the lower they score on the x-axis.

Although a simplified representation of reality, the model can function as a framework to categorise different types of teachers. Research shows that teachers confronted with problem behaviour of students in their classrooms react in different ways. For instance, some teachers with a high level of systematic teaching focus mainly on behaviour. Such teachers see reducing problem behaviour as a prerequisite for academic instruction (Sutherland, Lewis-Palmer, Stichter, & Morgan, 2008). Confronted with problem behaviour, they tend to focus on systematic behavioural or emotional interventions using
allocated instruction time for redirecting behaviour (these teachers are in the upper left quadrant). Conversely, other teachers with a high level of systematic teaching focus on academic rather than behavioural instruction during instructional time. These teachers (in the upper right quadrant) reinforce their instruction techniques or adapt the task to the skills of the students in order to increase on-task behaviour (Raggi & Chronis, 2006; van der Worp-van der Kamp et al., 2014).

Teachers with a low level of systematic teaching, on the other hand, often work unprepared and more informal, regularly implementing interventions on an ad hoc basis (Kern et al., 2009; Mooij & Smeets, 2009). Confronted with problem behaviour, some of these teachers may persist on focusing on academic tasks, without considering students' skills. The offered tasks are too easy or too difficult and the curriculum is not always carefully constructed, for example because the teachers feel pressure to complete the curriculum regardless of student mastery (Brigham et al., 2004). These teachers are placed in the lower right quadrant. Other teachers (in the lower left quadrant) spend most of their time redirecting student behaviour at the expense of academic instruction, or they may even remove these students from the lessons (Pianta & Hamre, 2009).

Literature shows that many teachers work somewhat ad hoc (Banks & Zionts, 2009; Kern et al., 2009; Mooij & Smeets, 2009) with a focus on behaviour (Levy & Vaughn, 2002; Pianta & Hamre, 2009). However, based on the aforementioned theory, systematic academic instruction seems to bode best for positive academic outcomes and could be a key strategy for decreasing problem behaviour. The empirical basis for this is, however, still rather small. To date, scarce research has been done on the lasting effect of teaching academic learning through a systematic, cyclic, on-going approach to the academic and behavioural outcomes of students with EBD (van der Worp-van der Kamp et al., 2014). Further research is necessary to ascertain if teachers in the upper right quadrant actually show better results (behaviourally as well as academically) than teachers in the other quadrants. Therefore, tools are needed to measure systematic teaching and academic instruction. This chapter describes the development of such a tool.

Teachers' behaviour can be measured through direct observation, interviews or questionnaires. Given the fact that the on-going process of systematic teaching cannot be captured in one or two observations and because certain parts of the PDCA cycle cannot be observed immediately and take place in the teachers' mind before and after the overt lesson, observation is not a suitable instrument to measure the complete PDCA cycle. The
The advantage of questionnaires over individual interviews is that the former requires less time than the latter, making it possible to assess a larger number of teachers. Research shows that teacher self-reports are reasonably accurate (Clunies-Ross, Little, & Kienhuis, 2008; Porter, 2002). A literature search, however, did reveal neither questionnaires concerning teacher use of the PDCA cycle, nor questionnaires on academic versus behavioural instruction. Therefore, we decided to develop two questionnaires, one concerning the PDCA cycle and the other concerning academic instruction.

### 3.2 Method

#### 3.2.1 Design

The development of the questionnaires began with an orientation visit to a real educational setting for students with EBD. Based on a study of lesson plans, observation of lessons, expert meetings with school counsellors and cognitive interviews with teachers (Figure 3.2, step 1), items for both questionnaires were drawn up. Next, data were collected, with teachers as respondents and used as the basis for calculating the reliability of both questionnaires (step 2). In order to validate both questionnaires (step 3), the agreement between teachers’ self-reported and observed use of systematic teaching and academic instruction was studied. For this purpose, the items of the questionnaires were transferred into observational items. The resulting observation scales comprise the PDCA cycle and academic instruction (step 2.1). The procedure for developing the questionnaires and observation scales is given in Figure 3.2. To differentiate between the questionnaires and the observation scales, these were named PDCA_{Q} versus PDCA_{Obs} and AI_{Q} versus AI_{Obs}.

![Figure 3.2. Construction of the questionnaires and observation scales.](image-url)
3.2.2 Participants

All participating schools were situated in Northern Netherlands.

Step 1 took place in a particular school for special education, whose school policy is to offer students adequate academic instruction so much time is scheduled for teaching academic skills. Eight teachers participated in this phase. During the period of item construction, these teachers used extended lesson plans (based on the Deming cycle of PDCA).

In step 2, the questionnaires were sent to the administrators of five special schools for students with EBD (aged 7–12), who distributed the questionnaires among all their 80 teachers. Fifty-six (72%) teachers completed and returned the questionnaire. Years of teaching experience ranged from <5 (28%), 5–10 (19%), 10–15 (21%), 15–20 (12%) to >20 (20%). Of these teachers, 19 percent were teaching students in grade 3/4, 26 percent in grade 5/6 and 40 percent in grade 7/8 while the remaining 15 percent were teaching in another combination of grades.

Thirty teachers (54%) agreed to be observed (step 3), but due to illness, too many demands in the classroom and switching jobs, eventually 20 teachers (36%) were observed for step 3. Five of these teachers were selected, based on their availability, to be observed by two observers in order to determine inter-observer agreement (step 2.1). All observations were conducted by an experienced, primary school teacher. The first author of this paper participated in the joint observations necessary for the determination of the IOA.

3.2.3 Development

3.2.3.1 Questionnaire

The item construction were grounded upon open lesson observations, accompanied by semi-structured interviews. Teachers were asked to hand over their lesson plans before every observation. Plans and observations were discussed with the teachers afterwards. This empirical input and the extensive feedback from teachers resulted in a number of key topics. These were discussed during two meetings with school counsellors and translated into items for the questionnaires. Next, these items were optimised through cognitive interview techniques (Willis, 2005) in which the researcher submitted the items to each of the eight teachers in turn, inviting them to think out loud about what they felt each item was about and what they felt certain words and phrases in the item meant. Ambiguous
items and terms were replaced and read again to the next teacher. This resulted in two
finite questionnaires comprising 36 items about the PDCA cycle (PDCAQ) and 24 items on
behavioural and academic instruction (AIQ). A 4-point Likert response scale varying from
'no or rarely' (1) to 'very often' (4) was used. The mean score determined respondents'
degree of systematic and academic instruction. Table 3.1 shows a couple of examples of
the final items used.

The items of the AIQ comprised both behavioural and academic instruction. High
scores on academic instruction were expected to go along with low scores on behavioural
instruction. Since these scores have the opposite result, the items on behavioural
instruction were re-coded. The higher the teachers score on items concerning behaviour,
the lower they score on the AIQ.

**Table 3.1.** Examples of items of the Questionnaires.

<table>
<thead>
<tr>
<th>PDCAQ</th>
<th>No or rarely</th>
<th>occasionally</th>
<th>regularly</th>
<th>very often</th>
</tr>
</thead>
<tbody>
<tr>
<td>I adjust the learning goals to the capabilities of the students in advance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A great deal of allocated learning time is lost to unexpected happenings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the complete lesson, I check the attainability of the learning goals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I use evaluation data for the preparation of my lessons</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I discuss the realization of the learning goals with my students</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AIQ</th>
<th>no or rarely</th>
<th>occasionally</th>
<th>regularly</th>
<th>very often</th>
</tr>
</thead>
<tbody>
<tr>
<td>I spend allocated instruction time actually on teaching academic skills</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In case of disruptive behaviour, I check the appropriateness of the learning task</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In case of disruptive behaviour I discuss behavioural rules</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I mainly reward good behaviour</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.2.3.2 Observation scale

All items of the PDCAQ were used for constructing the PDCA_{obs} observation scale. The items included specific descriptions of teacher behaviour concerning the PDCA cycle and was rated from 1 to 4 points. Teachers’ preparation and evaluation of the lessons (some items of the plan and most items of the act phase) took place before and after the observed lessons. These items were hard to observe and therefore were looked at by checking written lesson plans and by interviewing teachers after these observed lessons. The mean score determined teachers’ degree of systematic instruction. Table 3.2 shows some examples of the items, with corresponding points.

**Table 3.2. Examples of observation items of the PDCA_{obs} (Observation scale).**

<table>
<thead>
<tr>
<th>1 point</th>
<th>2 points</th>
<th>3 points</th>
<th>4 points</th>
</tr>
</thead>
<tbody>
<tr>
<td>The teacher appointed no/a superficial goal for the whole class</td>
<td>The teacher appointed an adaptive/measurable goal for the whole class</td>
<td>The teacher appointed an adaptive and measurable goal for the whole class</td>
<td>The teacher appointed an adaptive, measurable goal for groups/individual students</td>
</tr>
<tr>
<td>In case of incidents, the teacher reacts ad hoc</td>
<td>In case of incidents, the teacher uses a planned general approach</td>
<td>In case of incidents, the teacher uses a planned approach adapted to the group</td>
<td>In case of incidents, the teacher uses a planned approach adapted to specific students</td>
</tr>
<tr>
<td>The teacher offers the whole group a general instruction, strictly from the method/book</td>
<td>The teacher offers the whole group a specific/adaptive instruction.</td>
<td>The teacher offers specific groups a specific/adaptive instruction.</td>
<td>The teacher offers specific students a specific/adaptive instruction</td>
</tr>
<tr>
<td>The teacher does not address the learning goal/process slightly</td>
<td>The teacher addresses the learning goal/process slightly</td>
<td>The teacher addresses the learning goal/process extensive</td>
<td>The teacher verifies the learning goal/process and discusses it with the students</td>
</tr>
</tbody>
</table>

When developing the A_{obs}, a distinction was made between academic and behavioural oriented remarks. Everything a teacher said regarding the content of the academic task was obviously noted as academic instruction, i.e., ‘Look at assignment three on the whiteboard’, ‘Which part of this task do you not understand?’ or ‘Continue with page three’. Likewise, every remark about students’ behaviour was noted as behavioural instruction, i.e., ‘Sit on your chair’, ‘Keep working’, ‘Be quiet’ or ‘Pay attention please’. Although these remarks may be intended to get pupils to work, they do not assign concrete academic tasks and are thus behavioural focused. Time spent on matters unrelated to the lesson, i.e., small talk, handing out medication, talking to persons outside the classroom and periods of silence were noted as ‘other’. For the precise scoring of the
instruction, teachers were asked to record the observed lesson using a small portable audio device. Afterwards, the time spent on academic, behavioural and 'other' instruction was noted. The AlObs score was determined through dividing the percentage of time a teacher spends on academic instruction by the time spend on both academic and behavioural instruction. Thus, the more teachers spent on behavioural instruction, the lower the score on the AlObs. To facilitate the comparison with the outcomes from the 4-point scale of the AIq, this score was multiplied by 4.

3.2.3. Data analyses

3.2.3.1 Questionnaires
The first step of the construction of the questionnaire involved an on-going and iterative process of collecting and analysing data. The analyses of all next steps as well as the final outcomes of the questionnaires in the coordinate system were performed in SPSS (IBM, Armonk, New York, United States). The reliability of both questionnaires and the PDCAObs was assessed by calculating Cronbach's alpha. Non-essential items that correlated poorly with the total score (item total correlation <0.2) were removed.

3.2.3.2 Observation scale
The inter-observer agreement of the PDCAObs and AlObs was established by using Bland–Altman plots (Bland & Altman, 1986). The validity of the questionnaires was determined by the correlation between the outcomes of the questionnaires and observation scales. Because a correlation does not automatically imply an agreement between two measurements, Bland–Altman plots were also used for measuring the agreement between the questionnaires and observation scales.

Bland–Altman plots are based on graphical techniques and provide information about the agreement and nature of the differences between two measurements, i.e., in this study those of two different observers concerning the PDCAObs and the AlObs (inter-observer agreement) on the one hand and systematic teaching and academic instruction by two different instruments, namely the questionnaires and the observation scales on the other. Provided the differences have an approximate normal distribution, about 95 percent of the calculated differences will fall between the mean difference, plus or minus two standard deviations (the latter being entitled as the limits of agreement). The smaller the range between these two limits the better the agreement. The mean differences between the two measurements as well as the limits of agreement are also presented in the plots. The acceptability of the differences between the two measurements
was determined by the limits of agreement. A difference in a score exceeding these limits was regarded as a lack of agreement.

Finally, the outcomes of the questionnaires of the sample of 56 teachers were processed in a scatterplot representing the above-mentioned coordinate system of systematic academic instruction.

### 3.3 Results

**3.3.1 Questionnaires**

The first step resulted in the following key topics on systematic academic instruction:

1) Preparation of concrete academic goals as well as their communication and evaluation.
2) Consideration of differences between students during planning, throughout instruction and as a result of the evaluation.
3) Understanding of problem behaviour and knowledge how to handle it.
4) Anticipation on problem behaviour and incidents.
5) Teachers' focus on behavioural versus academic approach.

All these topics formed the basis for 60 items.

Before establishing the reliability of both instruments (step 2), six items of the PDCAQ and four of the AIQ were deleted because they correlated extremely weakly with the total score. One item with a low item – total score (0.15) was considered to be essential for the PDCA process (preparing additional tasks for students who completed their task) and was therefore not deleted. After removal of the 10 items, the alphas for the remaining items of the PDCAQ and AIQ were 0.89 and 0.76 respectively, suggesting that the scale scores with these items were reliable.

**3.3.2 Observation scale**

Concerning the inter-observer agreement of the observation scales, the Bland–Alman plot for the PDCA_{obs} (Figure 3.3) showed that observer 1 scored an average 0.03 point higher than observer 2 [standard deviation (SD) 0.16]. All differences fell between the limits of agreement (−0.30 and 0.36) and were nicely scattered around the mean difference.
Regarding the $A_{\text{obs}}$, observer 1 scored an average 0.31 points (SD 0.20) lower than observer 2 (Figure 3.4). All differences fell within the lines of agreement (−0.71 and 0.09) but the disagreement about teacher 3 (a difference of 0.63 points) was remarkable. To understand this disagreement, the audiotape of that particular teacher was played once again by both observers. It turned out that observer 1 and observer 2 had a different interpretation of one aspect of academic instruction, namely those concerning learning conditions, like ‘get your book’, ‘open your book’, ‘pick up your pencil’ etc. Observer 1 scored these actions as being behavioural focused, while observer 2 as academically based. Particularly teacher number 3 spent much time on these particular actions. But because these remarks do not include specific academic instruction, the $A_{\text{obs}}$ was improved by describing these actions as behavioural focused. Apart from this difference, both observers were in far-reaching agreement.
Figure 3.4. Bland–Altman plot concerning inter-observer agreement observation scale concerning academic instruction (Alobs).

The 20 observed lessons had a mean duration of 46.2 (SD 8.5) minutes. The results indicated that most of the observed instructional time was spent on academic instruction (56%). It should be noted that the audio recordings revealed that much of the academic instruction was given in a one-to-one interaction during independent practice: learning tasks actually assigned to be done by students without supervision. All outcomes of the observed teachers are displayed in Table 3.3.

Table 3.3. Outcomes questionnaires and observation scales (N = 20).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDCAQ</td>
<td>2.90</td>
<td>0.27</td>
</tr>
<tr>
<td>AIQ</td>
<td>2.66</td>
<td>0.38</td>
</tr>
<tr>
<td>PDCAObs</td>
<td>2.54</td>
<td>0.53</td>
</tr>
<tr>
<td>AlObs</td>
<td>2.97</td>
<td>0.49</td>
</tr>
</tbody>
</table>

Percentage of time spent on academic versus behavioural instruction

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic instruction</td>
<td>56.0%</td>
<td>12.8</td>
</tr>
<tr>
<td>Behavioural instruction</td>
<td>19.5%</td>
<td>10.2</td>
</tr>
<tr>
<td>Other</td>
<td>24.5%</td>
<td>11.3</td>
</tr>
</tbody>
</table>
For the PDCA$_{Q}$ and the PDCA$_{Obs}$ as well as for the AI$_{Q}$ and AI$_{Obs}$, we found a moderate Pearson correlation of 0.32 and 0.31 respectively. Bland–Altman plots (Figures 3.5 and 3.6) provided us with more specific information on the agreement between the questionnaires and observation scales. As Figure 3.5 shows, the mean difference between the scores on the PDCA$_{Q}$ and PDCA$_{Obs}$ was 0.36 (SD 0.53). The 95 percent limits of agreement were −0.7 and 1.42 and all differences fell within these limits of agreement. Generally, teachers ranked higher on the PDCA$_{Q}$ than the PDCA$_{Obs}$, while the distribution of the dots in the plot reveals that teachers with a high mean score on the x-axis of the plot (>3) also ranked higher on the observation scale. Consequently, most teachers assessed their systematic teaching higher than the observer did, except for a couple of teachers with a high mean score on the PDCA$_{Q}$ and PDCA$_{Obs}$. Analyses of the differences between the scores revealed that the greatest differences lay in the plan and do part of the cycle. The degree of planning as indicated by the outcomes of the questionnaires was not observed during the lessons, mainly because an overt, written plan was seldom accessible (except for the teachers with a high mean score).

As Figure 3.6 shows, the mean difference between the scores on the AI$_{Q}$ and AI$_{Obs}$ was −0.31 (SD 0.52). Generally, teachers ranked lower on the AI$_{Q}$ than the AI$_{Obs}$. There was no pattern in the relation between the mean score and the differences. The dots were nicely scattered around the mean difference and all differences fell between the limits of agreement (−1.35 and 0.73). Even though Figures 3.5 and 3.6 indicate that the differences between questionnaires and observation scales were acceptable, the limits of agreement were considered wide.
Figure 3.5. Bland–Altman plot concerning questionnaire and observation scale concerning Plan-Do-Check-Act (PDCA\textsubscript{Q} and PDCA\textsubscript{Obs}).

Figure 3.6. Bland–Altman plot concerning questionnaire and observation scale concerning academic instruction (AI\textsubscript{Q} and AI\textsubscript{Obs}).
The final outcomes of the questionnaires of all 56 teachers are shown in the coordinate system ‘systematic academic instruction’ (Figure 3.7). The mean score on the PDCAQ and AIQ were 2.89 (0.31 SD) and 2.68 (0.38 SD) respectively. In the upper quadrants, teachers scored 69% (upper right) and 22% (upper left). In the lower quadrant, teachers’ scores were 7% (lower right) and a meagre 2% (lower left).

![Figure 3.7. Outcomes questionnaire concerning Plan-Do-Check-Act (PDCAQ) and questionnaire concerning academic instruction (AIQ) presented in coordinate system ‘systematic academic instruction’ (N = 56).](image)

### 3.4 Discussion

The aim of this study was to develop a tool measuring teachers’ systematic academic instruction to students with EBD, namely a questionnaire concerning the Plan-Do-Check-Act-cycle (PDCAQ) and a questionnaire concerning academic instruction (AIQ). Various steps were taken to assure that the tool, comprising two questionnaires, was reliable, valid and came close to the daily practice of teaching students with EBD. The empirical data showed that the reliability of both questionnaires was high (PDCAQ) and acceptable (AIQ) as demonstrated by Cronbach’s $\alpha$ of 0.89 and 0.76 respectively. Their validity was supported by fair correlations between the questionnaires and the observation scales (being 0.32 and 0.31). Because correlation coefficients can be misleading in method agreement studies (Bland and Altman, 1986), Bland–Altman plots were used to quantify
the differences between the two methods. These plots not only made it easy to interpret differences, they also showed their magnitude. The wide limits of agreement revealed some differences between questionnaires and observation scales. Regarding the different approaches of both methods (self-reported and observed behaviour), we consider the limits of agreement small enough to be confident that both questionnaires reflected teachers' systematic academic teaching in a satisfactory manner.

Next, the questionnaires made it possible to include all participants in the coordinate system 'systematic academic instruction'. As Figure 3.7 reveals, the majority of teachers scored in the upper right quadrant, indicating a focus on academic instruction in a systematic manner. These outcomes do not confirm statements from literature that numerous teachers tend to work in an ad hoc manner, focusing more on behaviour than on academic instruction. A possible explanation might be that systematic academic instruction is a point of particular interest in RENN4 (Doornenbal, Jonker, Bijstra, & Minnaert, 2006). The same applies to Dutch special education for students with EBD, especially since the Dutch education inspectorate has emphasised the importance of high academic outcomes for all students, including those with EBD. Further analysis of each single axis of the coordinate system may enhance our understanding of the outcomes. The aforementioned differences between questionnaires and observation scales can also provide us with certain indications.

Looking at the y-axis, we see that teachers scored relatively high on systematic teaching, with a mean score of 2.89. Observation, however, revealed that teachers hardly ever write out lesson plans. Although this finding is in agreement with literature (John, 2006; Morine-Dershimer, 1978; Sutherland & Oswald, 2005), the question arises to what extent this implicit, covert planning can still be regarded as systematic. Detailed and daily written lesson plans are considered to be important organisational blueprints of what will be taught (Yell, Busch, & Rogers, 2008). We agree with Shen, Poppink, Cui and Fan (2007) in describing lesson plans as “important yet often overlooked sources of professional growth”. The PDCA$_Q$ makes no explicit distinction between overt and covert planning. Future studies regarding the importance and feasibility of written lesson plans are recommended.

Considering the x-axis of the coordinate system, the mean score of 2.68 on behavioural versus academic instruction suggests that teachers tend to focus slightly more on the latter. The observations confirm this outcome, but also reveal that a striking
amount of academic instruction was offered to students individually. This corresponds with the outcomes of Vaughn, Levy, Coleman and Bos (2002) who concluded that students in special education services receive more individual instruction compared with their peers without disabilities. We can therefore assume that teachers who score high on both axes (upper right quadrant) provide sufficient systematic academic instructions to their classes.

The final aim when developing the questionnaires was to determine the position in the coordinate system of the most effective teachers. Comparing the positions of teachers in the coordinate system with academic and behavioural outcomes of students might reveal the most effective approach. For example, teaching students with EBD might require a very high score on both axes. If so, effective teachers have to be positioned in the extreme upper right corner of the coordinate system. On the other hand, it is also possible that academic instruction to students with EBD always requires a certain degree of behavioural instruction. In that case, effective teachers are positioned near the upper centre of the coordinate system. More research is required to define limits of effective teacher positions in the coordinate system. The questionnaires are very useful for this purpose and will therefore be used as an instrument in the next chapter. In that chapter we aim to assess the impact of systematic academic instruction on academic progress and behavioural problems of EBD-students in special education.

Bland–Altman plots proved to be very applicable to clarify the inter-observer agreement as well as the differences between observation scales and questionnaires. The interpretation of the width of the limits of agreement, however, is challenging. Only if the limits are narrow enough, two different methods of measurement could be evaluated as equivalent. The interpretation of ‘narrow enough’ is not statistically determined but is a question of clinical judgement. Notwithstanding this limitation, based on the agreement between questionnaires and observations, we have sufficient confidence in this tool to use it as an instrument in the following chapter.