

Birgit Brouër

Selbstbeurteilungsfähigkeit von Schülerinnen und Schülern als Aspekt des Selbstorganisierten Lernens¹

1. Problemstellung

Die Fähigkeit, die eigenen Leistungen selbst zu beurteilen, ist im Prozess des lebenslangen Lernens von fundamentaler Bedeutung. Während der Schulzeit versetzt diese Fähigkeit die Lernenden in die Lage, ein realistisches Bild von dem eigenen Leistungsvermögen zu bekommen, das für das Erkennen eigener Stärken und Schwächen unerlässlich ist. Die realistische Einschätzung der eigenen Fähigkeiten ist nicht minder wichtig, wenn man die Schule verlässt und das Berufsleben oder ein Studium beginnt, etwa wenn es darum geht, Entscheidungen bezüglich des eigenen beruflichen Werdegangs zu treffen oder die eigene Leistung in Beurteilungsgesprächen mit dem Vorgesetzten darzustellen. Das lernende Individuum ist stets darauf angewiesen, selbst einzuschätzen, wie weit ein bestimmter Prozess fortgeschritten ist und wie gut es seine Aufgaben bewältigen kann.

Ob der schulische Unterricht jedoch auf diese Herausforderung genügend vorbereitet, muss bezweifelt werden (vgl. z.B. *Jürgens & Sacher* 2000; *Sacher* 1994; *Stefani* 1994). In herkömmlichen schulischen Lernarrangements ist es üblich, dass die Lehrenden die Beurteilung der Schülerleistungen übernehmen und auch verantworten müssen (z.B. *Heymann* 2001; *Tent* 1998). Schüler fragen nicht danach, was sie können oder nicht können, sondern sie fragen danach, welche Note sie für ihre Leistung bekommen. Bei einer schlechten Note zweifeln sie nicht unbedingt ihre eigenen Leistungen an, sondern eher an der Fähigkeit des Lehrenden, eine objektiv richtige Zensur zu geben. Hinter jeder Note, als der gängigen Form der schulischen Leistungsbeurteilung, steht also eine Lehrperson, die diese gegenüber Schülern, Eltern, Kollegen und Vorgesetzten zu verantworten hat. Diese traditionell so gehandhabte Vorgehensweise wird auch in aktuellen erziehungswissenschaftlichen Publikationen kaum angezweifelt. Während eine rege Diskussion über Schulleistungsvergleiche, Bezugsnormen und Lernchancen im Gange ist, erstaunt es, dass eine Beurteilung der Leistungen durch den Schüler selbst kaum diskutiert wird. Ein Beispiel dafür ist das im vergangenen Jahr erschienene Buch mit dem Titel „Leistungsmessungen in Schulen“ (*Weinert* 2001). Auffällig ist das Fehlen der Stichworte „Selbstbeurteilung“ bzw. „Selbstbewertung“. Die Leistungsmessung durch den Lernenden selbst wird in keiner Weise thematisiert. Man kann dies mit der Notwendigkeit der Notengebung erklären oder mit dem Wissens- und Erfahrungsvorsprung der Lehrenden gegenüber

¹ Der Artikel erscheint in: van Buer, Jürgen (Hrsg.) (2003). Qualitätssicherung in der beruflichen Bildung. Frankfurt/Main et al.: Peter Lang

den Lernenden. Mit Blick auf die Notwendigkeit des permanenten Lernens auch weit über die Schulzeit hinaus hat diese Tradition jedoch fatale Konsequenzen. Sie erzieht die Schüler/innen dazu, sich auf das Urteil anderer zu verlassen und sich häufig keine eigene Meinung zu bilden. Wünscht man sich aber Lernende, die im Sinne einer Selbst- bzw. Persönlichkeitskompetenz die Verantwortung für das eigene Lernen übernehmen können (*Jürgens & Sacher* 2000, S. 10), so muss man die entsprechenden Kompetenzen gezielt fördern.

In diesem Artikel sollen Möglichkeiten dieser Förderung diskutiert werden. Dabei wird das Augenmerk auf die Entwicklung der Selbstbeurteilungsfähigkeit im Rahmen eines selbstorganisationsoffenen Unterrichts (sensu *Sembill* 1992 u. 1994) gerichtet. Es werden Daten aus einem Forschungsprojekt herangezogen, die zeigen, inwieweit selbstorganisiertes Lernen zu einer Verbesserung der Selbstbeurteilungsfähigkeit führen kann.

2. Theoretischer Hintergrund

2.1 Selbstbeurteilungsfähigkeit

Nach *Zimmerman* (1990) zeichnen sich selbstregulierte Lerner u.a. dadurch aus, dass sie ihr Lernen planen, sich Ziele setzen, Prozesse organisieren, sich dabei beobachten und den Prozess in verschiedenen Phasen evaluieren. Sie verfügen also in hohem Maße über die Fähigkeit, metakognitive Prozesse zu aktivieren. Außerdem gelingt es ihnen gut, sich selbst zu instruieren und sich zu bestärken.

Dem Thema dieses Beitrages entsprechend wird der Fokus auf die Beobachtung der eigenen Lernprozesse und die Evaluation der einzelnen Lernphasen gerichtet. Der Begriff der Selbstbeurteilungsfähigkeit wird dabei umfassender verstanden als z.B. Selbstbewertung, denn er beinhaltet neben dem Aspekt des Bewertens auch den Aspekt der Erfassung einer Leistung als unabdingbarer Voraussetzung für die Abgabe des Urteils über die Leistung.

So unterscheidet beispielsweise auch *Rheinberg* (2001) zwischen Leistungsmessung und Leistungsbeurteilung. Bei der Messung geht es um die Ermittlung eines Ergebnisses in Bezug auf die Menge und/oder Güte einer Leistung. Man erhält dadurch ein Resultat, das für sich alleine genommen noch nichts aussagt, z.B. 28 Punkte in einem Vokabeltest. Erst wenn dieses Ergebnis mit einem Standard verglichen wird, erhält es eine Aussagekraft. Der Vergleich mit dem Standard ist die Beurteilung.

Was nun als Standard definiert wird, hängt von verschiedenen Faktoren ab. Zum einen muss man unterscheiden, ob die sachliche, soziale oder individuelle Bezugsnorm gewählt wird. Für die Selbsteinschätzung ist es sicher wichtig, seine eigenen Entwicklungen und Leistungsveränderungen zu analysieren. Es ist aber ebenfalls von

Bedeutung zu erkennen, wo man im Vergleich zu anderen steht. Zum anderen sind Standards in manchen Bereichen relativ leicht zu definieren (10 gelöste Aufgaben bei 10 Aufgaben entspricht sehr gut usw.), in anderen deutlich schwieriger. Ist bei einem Aufsatz die Quantität ein Gütemaß oder zählt nur die Qualität? Woran kann man die Qualität erkennen? In vielen Unterrichtsfächern sind keine eindeutigen Kriterien für die Bestimmung der Güte einer Leistung festlegbar. Im Rahmen der schulischen Leistungsbewertung rettet man sich über dieses Faktum hinweg, indem man die soziale Bezugsnorm als Maßstab nimmt, auch wenn dies nicht immer die beste Bezugsnorm für einen einzelnen Schüler ist. Problematisch ist die soziale Bezugsnorm u.a. deshalb, weil sie sich zumeist auf eine bestimmte Klasse bezieht (ebd., S. 64). Wie gut diese Klasse jedoch im Vergleich mit anderen Klassen bzw. gleichaltrigen Schülern einzuschätzen ist, lässt sich daraus nicht ermitteln (*Schusser 1996, S. 28f*). Es kann also sein, dass ein Schüler zwar in seiner Klasse im Mittelfeld liegt, im Vergleich zu Schülern anderer Klassen jedoch sehr gut abschneidet. In diesem Fall ist die soziale Bezugsnorm sehr ungünstig für ihn. Die soziale Bezugsnorm erweist sich auch dann als ungünstig, wenn ein Schüler sich zwar im Laufe des Schuljahres in seinen Leistungen steigert, seine Vergleichsgruppe jedoch auch. Er wird dann trotz des Leistungszuwachses keine bessere Bewertung erfahren.

Was bedeuten diese Überlegungen im Umkehrschluss wiederum für die Selbstbewertung und die Selbstbeurteilungsfähigkeit der Schüler/innen über einen – sicherlich zu vermutenden – Motivationsverlust hinaus? Sie unterstreichen nachhaltig die Bedeutsamkeit und Notwendigkeit der Entwicklung und Förderung einer Selbstbeurteilungsfähigkeit von Schüler/innen: Ein Lernender, der seine eigenen Leistungen richtig beurteilen kann, vergleicht sich zwar i.d.R. durchaus mit seinen Mitschülern, kann aber auch seine eigenen Lernfortschritte erfassen und daraus die Erkenntnis seiner persönlichen Stärken und Schwächen ableiten. Er wird damit unabhängiger vom Urteil seiner Lehrer und lernt, seine Fähigkeiten und Möglichkeiten selbst einzuschätzen.

Wodurch jedoch zeichnet sich die Fähigkeit zur Selbstbeurteilung aus? In Anlehnung an Black und Wiliam (1998) können (1) die Wahrnehmung der eigenen bzw. der im Unterricht gesetzten Lehr- und Lernziele, (2) die Wahrnehmung des eigenen aktuellen Kenntnis- oder Fähigkeitsstandes und (3) die Bewertung des Lernergebnisses als wesentliche Komponenten dieser Fähigkeit angesehen werden. Diese drei Komponenten sollen im Folgenden näher betrachtet werden.

ad 1) Die Wahrnehmung der Lern- und Lehrziele ist als entscheidende Voraussetzung für die Beurteilung des Lernergebnisses zu betrachten, denn an ihnen gemessen definiert sich letztendlich der Erfolg des Unterrichts. Dabei muss unterschieden werden zwischen den vom Lehrer gesetzten Lehrzielen und den Lernzielen, die ein Lernender selbst anstrebt. Diese Ziele müssen nicht unbedingt übereinstimmen. Damit

aber die Lernenden entscheiden können, ob sie die Ziele der Lehrperson für sich selbst als relevant betrachten, muss es der Lehrperson gelingen, ihre Ziele für die Lernenden transparent und einsichtig zu machen.

In Bezug auf die Selbstbeurteilungsfähigkeit der Lernenden sind die Lernziele die Grundlage der Beurteilung, denn der Lernende muss prüfen und entscheiden, ob und mit welchem Erfolg er seine Lernziele erreicht hat. Wenn sich ein Lernender z.B. als Ziel des Lernens das Wissen über einen bestimmten Sachverhalt gesetzt hat, kann er am Ende des Lernprozesses überprüfen, ob dieses Wissen erlangt wurde. Dafür müssen mit der Formulierung der Lernziele jedoch auch Kriterien für die Beurteilung des Lernerfolges einher gehen.

Im traditionellen Unterricht wird nur selten eine Transparenz der Lehr- und Lernziele geschaffen. Normalerweise werden von der Lehrperson Lehrziele festgesetzt, die den Inhalt des Unterrichts bestimmen und an denen sich sein Aufbau orientiert. Die Lehrperson versucht dann, die Steuerung der vielfältigen Lernprozesse der Schüler/innen zu übernehmen. Unterricht kann jedoch nur gelingen, wenn die Lernenden selbst aktiv die Steuerung ihrer Lernprozesse übernehmen. Die Schüler wissen allerdings häufig nicht, welche Ziele die Lehrperson mit ihrem Unterricht verfolgt und welchen Weg sie einschlagen sollten, um diese Ziele zu erreichen (vgl. *Wagner 1999; Wagner 2000*). Zudem kommen in herkömmlichen Lehr-Lern-Arrangements eher selten Phasen vor, in denen die Lernenden darüber reflektieren können oder gar sollen, welche Ziele sie selbst in diesen Unterrichtsprozessen erreichen möchten.

Der Mangel an Transparenz der Lehr- und Lernziele führt letztlich dazu, dass die Selbstbeurteilungsfähigkeit der Lernenden nicht kriteriumsorientiert sein kann. Wenn transparente Ziele gesetzt sind, kann der Lernende davon auch Kriterien zur Beurteilung des Lernerfolges ableiten. Ist dies nicht der Fall, hängt die Selbstbeurteilung der Schüler eher von ihren generellen Überzeugungen, Emotionen und ihrer Selbstwirksamkeit ab (*Cameron & Pierce 1994*).

ad 2) Die Erreichung eines Lernzieles läuft häufig in verschiedenen Lernschritten ab (vgl. *Brouër 2001; Oser & Patry 1990; Wagner 1999*). Damit der Lernende bewusst und aktiv auf diesen Lernprozess Einfluss nehmen kann, muss er zu jedem Zeitpunkt des Unterrichts in der Lage sein, seinen aktuellen Kenntnis- oder Fähigkeitsstand wahrzunehmen und einzuschätzen. Das bedeutet nicht nur, dass er erkennen muss, was er bisher bereits geleistet hat, sondern er muss hier auch miteinbeziehen, wie viel Anstrengung ihn das gekostet hat, welche Mittel er mit welchem Erfolg eingesetzt hat und wie hoch seine Motivation zur Erreichung des Lernzieles war und ist.

Auch hier kommt der Transparenz der Lehr- und Lernprozesse eine Schlüsselrolle zu, denn nur wenn die anzustrebenden Ziele und die jeweils geltenden Bewertungs-

kriterien bekannt sind, kann der Lernende zu jedem Zeitpunkt seines Lernprozesses feststellen, ob seine Leistungen den eigenen bzw. den Anforderungen der Lehrperson entsprechen.

ad 3) Die Überprüfung des aktuellen Kenntnis- bzw. Fähigkeitsstandes beinhaltet auch eine Bewertung derselben, die schließlich zu Entscheidungen über den weiteren Verlauf des Lernprozesses führen sollte. Ist der Lernende dann am Ende des Lernprozesses angekommen, findet eine abschließende Beurteilung des Lernerfolges statt. Er überprüft und bewertet, ob er das sich selbst gesetzte Lernziel erreicht hat oder nicht, ob die eingesetzten Mittel adäquat waren, ob die Motivation ausreichend hoch war und das Ergebnis gemessen am Aufwand zufriedenstellend ist. Unabdingbar für diesen Reflexionsprozess ist die Kenntnis der Standards und Kriterien zur Bewertung des Lernergebnisses. Idealtypischerweise werden diese Standards und Kriterien ebenso wie die möglichen Lernziele und Lernprozesse gemeinsam mit den Lernenden diskutiert (z.B. *Barton & Collins* 1997; *Brunner & Schmidinger* 2000; *Sembill, Wolf, Wuttke, Santjer & Schumacher* 1998). Dass dies im herkömmlichen Unterricht häufig nicht geschieht, hat verschiedene Gründe, die im einleitenden Kapitel bereits angesprochen wurden. Dass es prinzipiell jedoch nicht unmöglich ist, zeigen einige Schulprojekte (z.B. *Stallings & Tascione* 1996; *Stefani* 1994), die sich in erster Linie durch eine hohe Schülerorientierung i.S. einer aktiven Beteiligung der Lernenden an den verschiedensten Prozessen des Unterrichts auszeichnen.

Es kann auf der Grundlage des bislang Dargelegten davon ausgegangen werden, dass zur Förderung der Selbstbeurteilungsfähigkeit der Lernenden ein geeigneter unterrichtlicher Rahmen von Nöten ist, der den Lernenden genügend Freiraum zur Entwicklung ihrer Fähigkeiten bietet und die Möglichkeit zur Verantwortungsübernahme gibt. Einen solchen Rahmen bieten selbstorganisationsoffene Lehr-Lern-Arrangements, in denen die Lernenden sich die Inhalte des Unterrichts an konkreten Problemstellungen selbstständig erarbeiten. Im Folgenden soll die Theorie des Selbstorganisierten Lernens von Sembill vorgestellt werden, die zur Durchführung solcher Lehr-Lern-Arrangements entwickelt wurde. Es wird anschließend geprüft, ob dieses Konzept zur Förderung der Selbstbeurteilungsfähigkeit hinreichend geeignet ist.

2.2 Selbstorganisationsoffene Lehr-Lern-Arrangements

Nicht erst seit der Veröffentlichung der Ergebnisse aus TIMSS und PISA wird konstatiert, dass die Fähigkeit zu lebenslangem Lernen, Flexibilität und Eigenverantwortlichkeit Schlüsselqualifikationen für eine erfolgreiche berufliche Tätigkeit sind

(Beck, Guldimann & Zutavern 1997; Boekaerts 1997; Schunk & Zimmermann 1994; Sembill 2000). Mit modernen Lehr-Lern-Arrangements sollen diese Qualifikationen stärker als bisher gefördert werden. Das von Sembill aus den Ideen des Forschenden Lernens (1992) entwickelte und mit seinen Mitarbeitern in diversen empirischen Projekten realisierte und überprüfte Konzept des Selbstorganisierten Lernens (im folgenden auch SoLe genannt) entspricht den Ansprüchen an diese modernen Lehr-Lern-Arrangements. In selbstorganisierten Unterrichtseinheiten werden die Lernenden an der Gestaltung des Unterrichts aktiv beteiligt, indem sie eigenständig und in Gruppen komplexe Problemstellungen bearbeiten (vgl. z.B. Sembill 1994; Sembill 1997; Sembill 2000; Sembill, Wolf, Wuttke, Santjer & Schumacher 1998).

Durch die Möglichkeit, in Gruppen über ihre Arbeit und Arbeitsprodukte zu sprechen, werden die Lernenden dazu angeregt, ihre Lernziele und Lernprozesse bewusst wahrzunehmen, weil sie sich darüber verständigen müssen, wie sie gemeinsam vorgehen wollen. In selbstorganisationsoffenen Lehr-Lern-Arrangements werden außerdem die Kriterien zur Bewertung der Arbeitsergebnisse gemeinsam mit den Lernenden festgelegt. Dadurch sind gute Voraussetzungen für eine angemessene Selbstbeurteilung der Lernenden gegeben, denn der Unterricht enthält Reflexionsmöglichkeiten und transparente Beurteilungskriterien. Ob eine selbstorganisationsoffene Lernumgebung jedoch schon ausreicht, um die Selbstbeurteilungsfähigkeit zu fördern, haben wir im Rahmen eines Forschungsprojektes geprüft, das an einer kaufmännischen Berufsschule im Fach Rechnungswesen durchgeführt wurde. Dieses Projekt soll im folgenden dargestellt werden. Für eine ausführlichere Beschreibung des Projektes und seiner Ergebnisse verweisen wir auf bereits publizierte Berichte (Brouër, Seifried & Sembill 2001; Seifried, Brouër & Sembill 2001; Seifried 2002; Seifried & Sembill 2002; Sembill, Seifried & Brouër 2002)². An dieser Stelle soll jedoch bereits erwähnt werden, dass in diesem Projekt nicht überprüft werden konnte, ob die Schüler ihre Lernziele bewusst wahrgenommen und ihre Lernergebnisse daran gemessen haben, sondern es wurde erfasst, wie sie eine eigene, klar umrissene Leistung selbst einschätzten.

3. Forschungsprojekt „Selbstorganisiertes Lernen im Fach Rechnungswesen“

Im Schuljahr 2000/2001 wurden an einer Bamberger Berufsschule zwei Klassen (erstes Ausbildungsjahr in einer kaufmännischen Berufsausbildung) über 72 Unterrichtsstunden hinweg im Fach Rechnungswesen nach den Kriterien des Selbstorganisierten Lernens unterrichtet.³ Der Unterricht wurde jeweils in vier Unterrichtsstunden pro

² Das Projekt wird auch in dem Beitrag von J. Seifried in diesem Band dargestellt.

³ Das Projekt wurde vom Bayerischen Kultusministerium sowie der internen Forschungsförderung der Otto-Friedrich-Universität Bamberg (Projekt-Nr. 060801) unterstützt.

Woche geblockt, so dass genügend Zeit für die vielfältigen Aktivitäten der Schüler zur Verfügung stand. Am Ende eines jeden Unterrichtsblockes waren die Schüler aufgefordert, ein Protokoll anzufertigen und damit ihre Arbeits- und Lernprozesse schriftlich festzuhalten. Dieses Arbeits- und Lernprotokoll kann als reflexionsanregende Fördermaßnahme betrachtet werden, denn die Schüler/innen wurden dazu ermuntert, ihre Lernziele, Lernprozesse und Lernprodukte damit zu protokollieren. Es soll geprüft werden, ob diese Maßnahme dazu beitragen kann, die Selbstbeurteilungsfähigkeit der Lernenden zu fördern. Es sei jedoch an dieser Stelle auch darauf hingewiesen, dass der Fokus der hier vorzustellenden Untersuchung auf der Fähigkeit der Schüler/innen, ihre eigene Leistung richtig einzuschätzen, gerichtet wird. Damit wird nur ein Teilaspekt der komplexeren Selbstbeurteilungsfähigkeit geprüft, nämlich der Aspekt der Selbstbewertung⁴.

Parallel zu den beiden Experimentalklassen erhielt eine dritte Klasse traditionellen Rechnungswesenunterricht und fungierte damit als Kontrollklasse. In allen drei Klassen wurden vor, während und nach der Untersuchung zu insgesamt vier Messzeitpunkten Daten erhoben wie z.B. biografische Daten, Vorwissen, Intelligenz und Problemlösefähigkeit. Die im folgenden dargestellten Analysen beziehen sich auf die Daten aus den Vorwissens- bzw. Wissenstests.

Weil Vergleiche zwischen dem Pre- und dem Post-Test angestellt werden, gehen in die Datenauswertung nur die Schüler/innen ein, die an beiden Messungen teilgenommen haben.

3.1 Pre-Test-Ergebnisse der Experimentalgruppe: Selbsteinschätzung im Vergleich zur Fremdeinschätzung

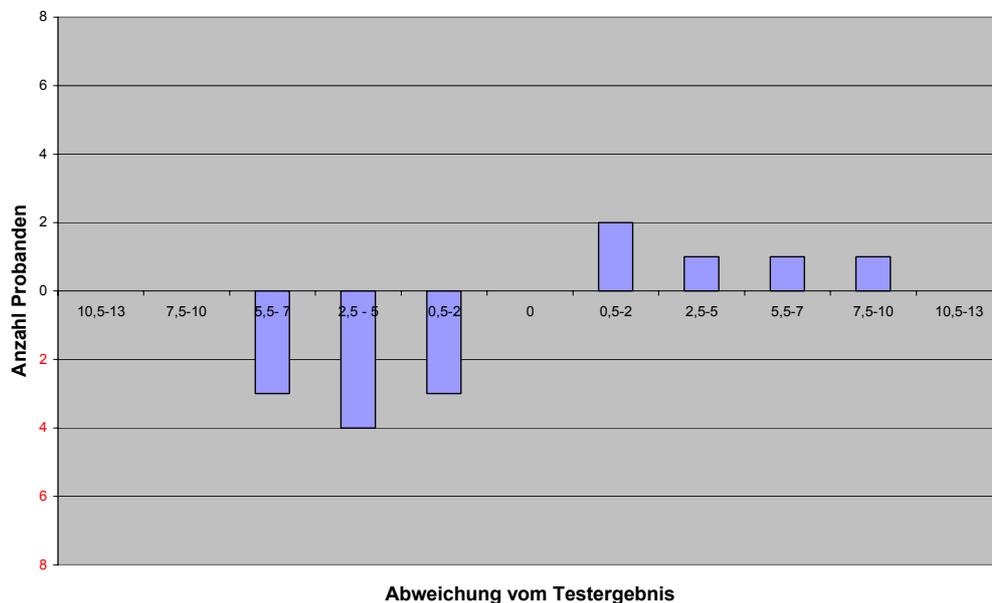
Zu Beginn der selbstorganisationsoffenen Unterrichtsphase wurden die Vorkenntnisse der Lernenden mit einem Test geprüft, der 25 fachliche Fragen zum Rechnungswesen und zur Allgemeinen Wirtschaftslehre enthielt. Am Ende dieses Tests wurden die Schüler/innen dazu aufgefordert selbst einzuschätzen, wie viele der 25 Fragen sie richtig beantwortet haben. Diese Selbsteinschätzung wurde dann verglichen mit der tatsächlichen Zahl der richtig beantworteten Fragen, also mit der Fremdeinschätzung des Testresultates.

Zu erwähnen ist hier, dass die Schüler/innen zu diesem Zeitpunkt des Projektes noch keinen Rechnungswesenunterricht erhalten hatten und sich die Testfragen auf Wissen bezogen, das sie möglicherweise in vorausgehenden Schulen oder in anderen Lebensbereichen erworben haben könnten. Der Test diente im Rahmen des Projektes u.a. dazu festzustellen, wie sich die drei Klassen hinsichtlich ihres fachlichen Vorwissens unterschieden.

⁴ Im Folgenden werden die Begriffe Selbsteinschätzung und Selbstbewertung synonym verwendet.

Die Abbildungen 1 und 2 zeigen die Differenzen zwischen der Selbst- und der Fremdeinschätzung der Richtigkeit der Wissensaufgaben für die beiden Experimentalklassen getrennt. Dabei werden die Ergebnisse nach der Höhe der Abweichung gruppiert und nach Unter- und Überschätzung der Probanden sortiert. Die Werte links des Nullpunktes zeigen die Häufigkeiten und Ausmaße der geringeren Selbsteinschätzung (auch: Unterschätzung) der Probanden im Vergleich zum Testergebnis an, die Werte rechts des Nullpunktes die höhere Selbsteinschätzung (auch: Überschätzung).

Abbildung 1: Häufigkeit der **geringeren** und **höheren** Selbsteinschätzung der Probanden der Experimentalklasse 1 im Vergleich zur



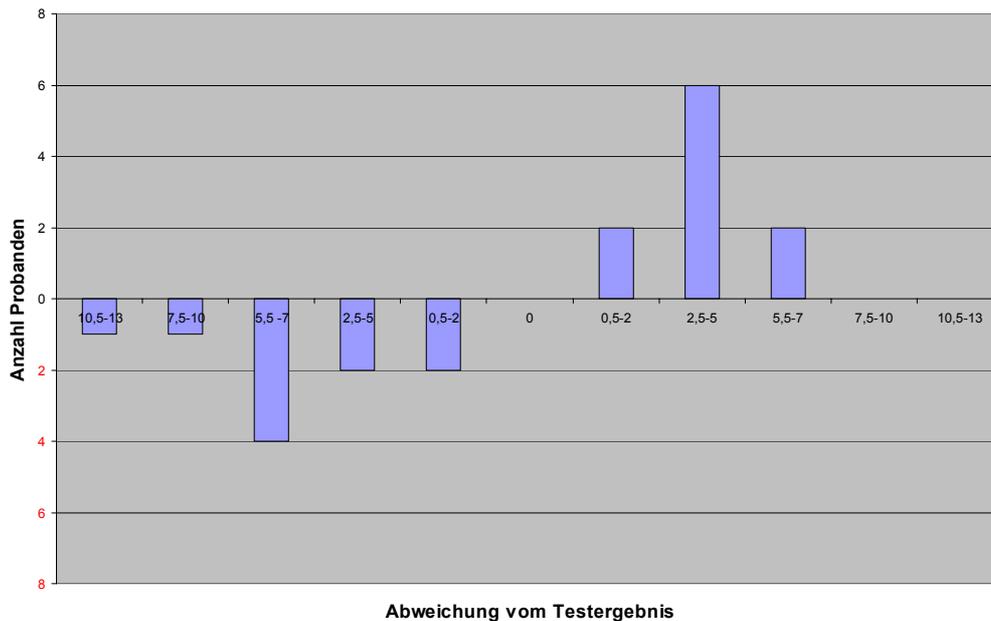
Fremdeinschätzung durch Experten im Pre-Test, N = 15

Wie Abbildung 1 zeigt, hat keiner der 15 Lernenden seine Leistung im Wissenstest richtig eingeschätzt. Die Klasse zeigt insgesamt eine deutliche Tendenz dazu, die eigene Leistung eher zu unterschätzen: 10 Schülern, die ihre Leistung geringer einschätzten, als sie tatsächlich war, stehen 5 Überschätzende gegenüber. Dabei ist der Grad der Über- bzw. Unterschätzung in beiden Gruppen ähnlich breit gestreut, es kommen also geringfügige Fehleinschätzungen (mit Punktdifferenzen von 0,5 bis 2) ebenso vor wie drastische (mit Punktdifferenzen von 5 bis 10). Insgesamt ist ein Drittel der Schüler/innen der Experimentalklasse 1 dazu in der Lage, die eigene Leistung relativ gut, also mit einer Abweichung von bis zu 2 Punkten, einzuschätzen.

Berechnet man aus der Summe der Abweichungen den arithmetischen Mittelwert, so erhält man die mittlere Abweichung⁵ von 4,21 Punkten für diese Klasse ($s = 2,57$).

Die Einschätzungen der Schüler/innen der zweiten Experimentalklasse sind in der folgenden Abbildung 2 dargestellt.

Abbildung 2: Häufigkeit der **geringeren** und **höheren** Selbsteinschätzung der Probanden der Experimentalklasse 2 im Vergleich zur Fremdeinschätzung durch Experten im Pre-Test, N = 20



In der Experimentalklasse 2 haben 20 Schüler/innen an der Erhebung teilgenommen und es zeigt sich hier in Abbildung 2 ein ähnlich heterogenes Bild wie bei der Experimentalklasse 1, allerdings mit graduellen Unterschieden. Keiner der Probanden konnte seine Leistung korrekt einschätzen und 4 Probanden (also ein Fünftel der Klasse) wichen in ihrer Einschätzung leicht (0,5 bis 2 Punkte) von der Fremdeinschätzung ab. Bei den Abweichungen ist das Verhältnis zwischen Unter- und Überschätzungen ausgeglichen, in beiden Richtungen sind je 10 Schüler/innen zu verzeichnen. Bei den Überschätzern fällt auf, dass der größte Teil der Schüler/innen sich im Bereich von 2,5 bis 5 Punkten verschätzte, also deutlich vom korrekten Ergebnis abwich. Bei den sich unterschätzenden Schülern finden wir sogar Abweichungen von großem Ausmaß (einmal 7, einmal 10 und einmal 12,5 Punkte), d.h. diese Schüler/innen waren von einer korrekten Selbsteinschätzung weit entfernt. Die mittlere

⁵ In die Mittelwertberechnung wurde die Differenz zwischen der Selbsteinschätzung eines jeden Schülers und seinem eigentlichen Testergebnis einbezogen, ohne dabei die Vorzeichen, also die Über- bzw. Unterschätzung, zu berücksichtigen. Auf die Vorzeichen konnte verzichtet werden, weil diese Analyse ausschließlich auf den Grad der Abweichung und nicht auf ihre Richtung fokussiert.

Abweichung beträgt hier 4,92 ($s = 2,87$) und ist damit etwas schlechter als der Mittelwert der Experimentalklasse 1.

Zusammenfassung der Pre-Test-Ergebnisse

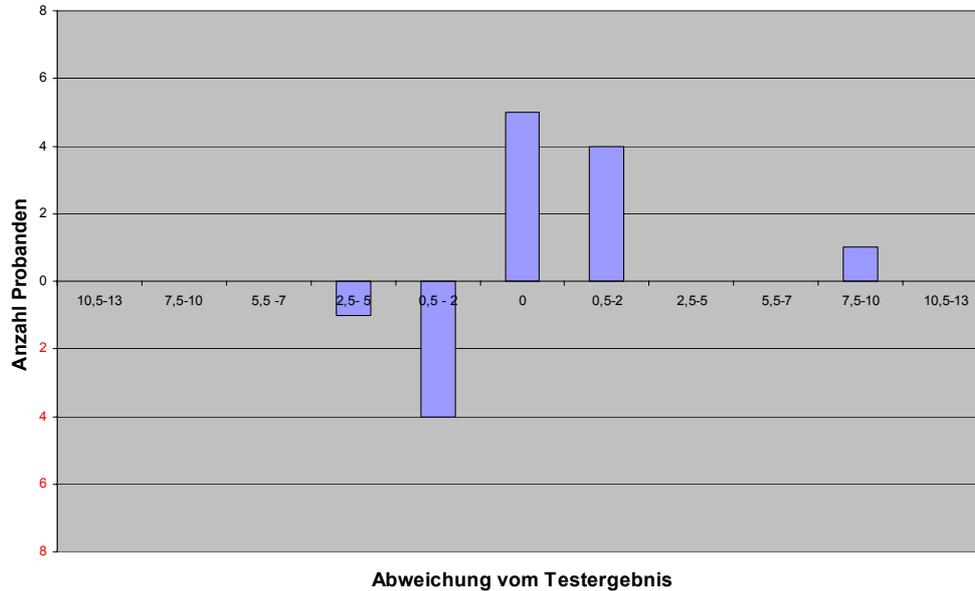
Für beide Experimentalklassen zusammen lässt sich feststellen, dass 57 Prozent der Abweichungen (entspricht 20 Personen) zwischen der Selbst- und der Fremdeinschätzung eine geringere Selbsteinschätzung anzeigen, während knapp 43 Prozent (entspricht 15 Personen) auf eine höhere Selbsteinschätzung hinweisen. Dabei zeigen sich jedoch graduelle Unterschiede. 5 Probanden schätzten ihre Leistung leicht schlechter ein (Abweichung von 0,5 bis 2 Punkten), als sie tatsächlich war, und 4 Probanden neigten zu einer geringfügigen Überschätzung (ebenfalls mit einer Abweichung von 0,5 bis 2 Punkten) ihrer Leistung. Die übrigen Schüler/innen (26 Personen) wichen sehr deutlich von der richtigen Einschätzung ihrer Leistung ab.

Eine Betrachtung der Unterschiede zwischen den Geschlechtern muss vorsichtig bleiben, da in den beiden Gruppen zusammen 29 Mädchen nur 6 Jungen gegenüber stehen. Es zeigt sich ein in der erziehungswissenschaftlichen Forschung nicht unbekannter Trend: Während von 29 Mädchen 18 ihre eigene Leistung unterschätzten und nur 11 sich überschätzten, ist das Verhältnis bei den Jungen umgekehrt, denn hier schätzte sich einer zu gering und 5 schätzten sich zu gut ein.

3.2 Post-Test: Selbsteinschätzung im Vergleich zur Fremdeinschätzung

Das gleiche Testverfahren wurde während der selbstorganisationsoffenen Unterrichtseinheit zweimal und ein letztes Mal nach Abschluss der selbstorganisierten Unterrichtseinheit angewendet. Die folgende Darstellung beschränkt sich auf die Ergebnisse aus dem Post-Test, da dies im Rahmen dieser Arbeit eine ausreichende Aussagekraft hat. In den Abbildungen 3 und 4 werden die Ergebnisse für beide Experimentalklassen getrennt präsentiert, wobei wiederum unterschieden wird zwischen den möglichen Richtungen der Abweichung der Selbsteinschätzung der Probanden von der Fremdeinschätzung durch die Experten, gruppiert nach der Höhe der Abweichungen.

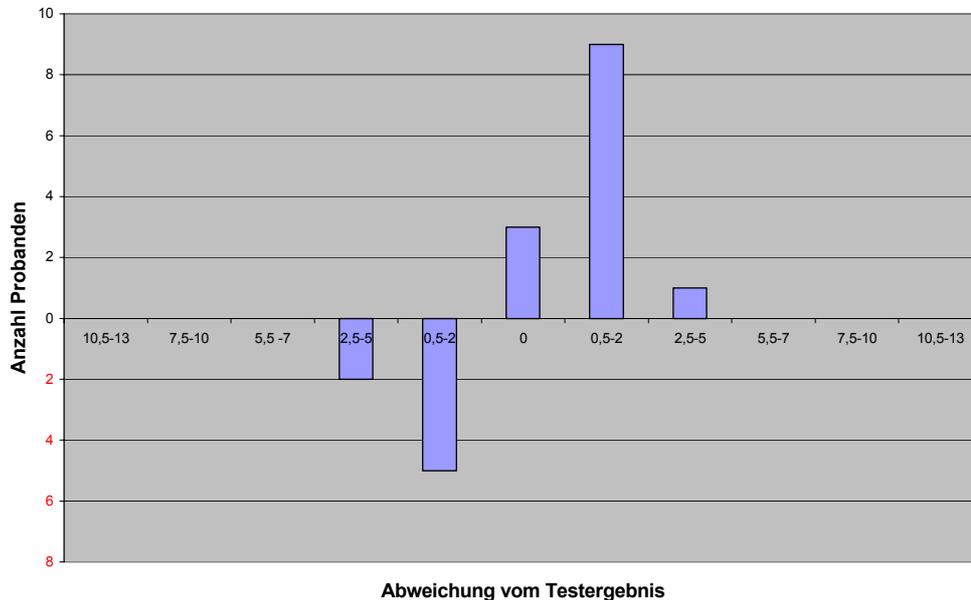
Abbildung 3: Häufigkeit der **geringeren** und **höheren** Selbsteinschätzung der Probanden der Experimentalklasse 1 im Vergleich zur Fremdeinschätzung durch Experten im Post-Test, N = 15



Die Abbildung 3 lässt eine deutliche Veränderung der Selbsteinschätzung der Experimentalklasse 1 im Vergleich zum Pre-Test erkennen: Nach der selbstorganisationsoffenen Unterrichtseinheit waren die Schüler/innen auffallend gut dazu in der Lage, ihre eigenen Leistungen richtig einzuschätzen. Fünf von 15 Probanden stimmten in ihrer Einschätzung völlig mit den Experten überein. Zu erkennen ist auch, dass geringfügige Abweichungen von 0,5 bis 2 Punkten den größten Teil der Abweichungen ausmachen, während stärkere und starke Abweichungen nur noch in zwei Fällen vorkamen. Die Schüler/innen dieser Klasse haben also im Verlaufe der selbstorganisierten Unterrichtsphasen gelernt, ihre eigene Leistung realistischer einzuschätzen, wobei nun das Zahlenverhältnis zwischen sich selbst unterschätzenden und sich selbst überschätzenden Schülern in dieser Klasse ausgeglichen ist. Es lässt sich hier ein mittlerer Abweichungswert von 1,58 berechnen, der ebenfalls eine deutliche Verbesserung der Klasse im Vergleich zum Pre-Test anzeigt ($s = 2,35$).

Die Ergebnisse für die zweite Experimentalklasse sind in der folgenden Abbildung 4 grafisch dargestellt.

Abbildung 4: Häufigkeit der **geringeren** und **höheren** Selbsteinschätzung der Probanden der Experimentalklasse 2 im Vergleich zur Fremdeinschätzung durch Experten im Post-Test, N = 20



Auch in der zweiten Experimentalklasse ist eine deutliche Verbesserung der Selbsteinschätzungsfähigkeit im Vergleich zum Pre-Test zu verzeichnen. 3 von 20 Schülern schätzten ihre Leistung nun korrekt ein und weitere 14 wichen nur schwach vom korrekten Wert ab. Die Abweichungen der übrigen 2 Schüler sind bei näherer Betrachtung ebenfalls nicht besonders hoch und bewegen sich im Rahmen von 2,5 bis 3,5 Punkten.

Bei dieser Klasse ist außerdem auffällig, dass das Verhältnis von sich selbst überschätzenden zu sich selbst unterschätzenden Schüler/innen im Vergleich zum Pre-Test nun zugunsten der Überschätzer (10 zu 7) gewichtet ist, während es beim Pre-Test ausgeglichen war. Die mittlere Abweichung beträgt 1,24 ($s = 1,04$) und ist damit sehr viel günstiger als beim Pre-Test.

Zusammenfassung der Post-Test-Ergebnisse

Erfreulicherweise ist festzustellen, dass in den beiden Experimentalklassen die Zahl der Schüler/innen, die ihre eigene Leistung korrekt eingeschätzt hat, von 0 auf 8 gestiegen ist. Das entspricht einem Anteil von 23 Prozent aller Probanden. Mit einem Anteil von 43 Prozent überschätzte der weitaus größte Teil der Probanden beider Klassen die eigenen Leistungen, während 34 Prozent der Probanden ihre Leistung unterschätzten. Dieses Verhältnis hat sich also im Vergleich zum Pre-Test umge-

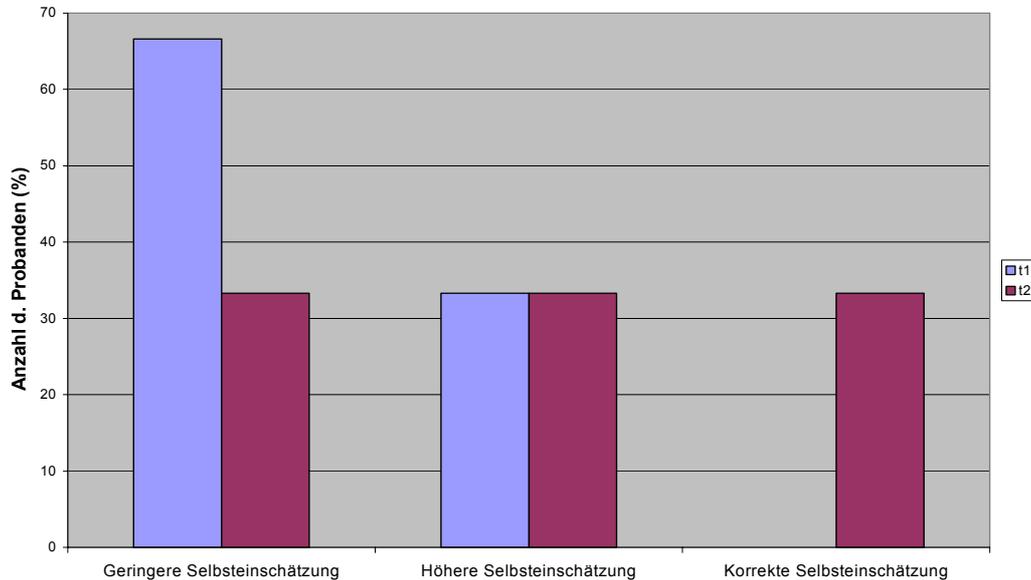
kehrt, denn dort neigten die Probanden noch eher dazu, ihre Leistungen zu unterschätzen. Auffallend bei den Ergebnissen des Post-Tests ist auch, dass die Abweichungen von der korrekten Leistungseinschätzungen nicht mehr so breit gestreut sind, wie sie es vor dem Treatment waren. Nur fünf von 35 Probanden schätzten ihre Leistungen eindeutig falsch ein (wobei sich einer stark überschätzte), während es beim Pre-Test 16 von 35 Probanden waren, die um mehr als 2 Punkte vom richtigen Ergebnis abwichen (und 8 davon wichen um mehr als 5 Punkte ab).

Ein Vergleich von Schülerinnen und Schülern zeigt hier, dass von sechs männlichen Probanden einer seine Leistung korrekt eingeschätzt hat, fünf haben sich überschätzt (davon einer mit 9 Punkten stark). Bei den Schülerinnen finden wir nun ein ausgeglicheneres Verhältnis von Unter- und Überschätzungen (12 zu 10) sowie 7 korrekte Selbsteinschätzungen. Vor allem die Mädchen konnten also das selbstorganisationsoffene Lehr-Lern-Arrangement positiv für sich nutzen und ihre Selbsteinschätzung deutlich verbessern.

3.3 Entwicklung der Selbstbeurteilungsfähigkeit der Experimentalklassen

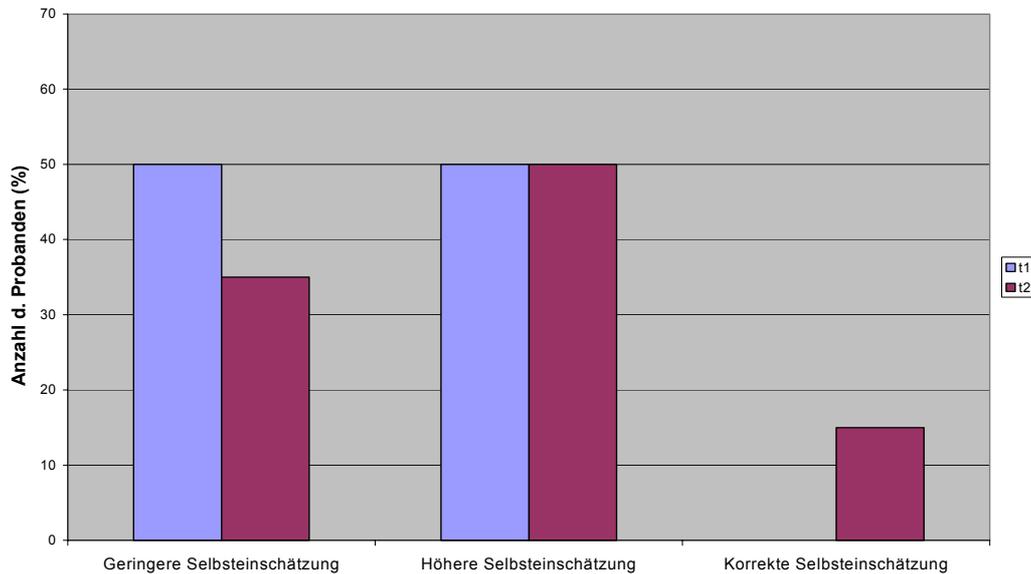
Um die Unterschiede zwischen den Ergebnissen des Pre-Tests und des Post-Tests grafisch zu veranschaulichen, ist in Abbildung 5 die relative Häufigkeit der Über- und der Unterschätzungen zu den jeweiligen Meßzeitpunkten für die Experimentalklasse 1 und in Abbildung 6 für die Experimentalklasse 2 dargestellt.

Abbildung 5: Entwicklung der Selbsteinschätzung der Probanden der Experimentalklasse 1 im Vergleich der beiden Messzeitpunkte t1 und t2 (t1 = Pre-Test, t2 = Post-Test). N = 15



Die in Abbildung 5 dargestellten Ergebnisse sind in zweierlei Hinsicht interessant: Zum einen zeigt der Vergleich zwischen dem Pre-Test und dem Post-Test, dass der Prozentanteil der korrekten Selbsteinschätzung im Verlaufe des Treatments stark zugenommen hat, und zum anderen hat sich der Anteil der Schüler/innen, die ihre eigene Leistung höher einschätzten, überhaupt nicht verändert. Während zum Zeitpunkt des Pre-Tests ca. Zweidrittel aller Probanden ihre Leistungen geringer einschätzten, als sie tatsächlich waren, ist dieser Prozentanteil zum Zeitpunkt des Post-Tests deutlich gesunken. Dieser drastische Abfall der geringeren Selbsteinschätzung und der deutliche Anstieg der korrekten Selbsteinschätzung könnten Indikatoren dafür sein, dass besonders die sich unterschätzenden Schüler/innen vom selbstorganisationsoffenen Lehr-Lern-Arrangement profitiert haben. Diese Vermutung wird in Kapitel 3.5 geprüft. Zuvor soll in Abbildung 6 die Entwicklung der zweiten Experimentalklassen vom ersten zum letzten Messzeitpunkt gezeigt werden.

Abbildung 6: Entwicklung der Selbsteinschätzung der Probanden der Experimentalklasse 2 im Vergleich der beiden Messzeitpunkte t1 und t2 (t1 = Pre-Test, t2 = Post-Test). N = 20



In dieser Abbildung ist zunächst auffällig, dass auch bei dieser Klasse der Anteil der höheren Selbsteinschätzung der Schüler/innen zu beiden Messzeitpunkten gleich hoch ist. Die Hälfte der Klasse neigte demnach dazu, die eigene Leistung zu gut zu bewerten. Die geringere Selbsteinschätzung fiel mit 50 Prozent im Pre-Test ebenfalls hoch aus, nahm im Verlaufe der Intervention jedoch zugunsten einer realistischen, korrekten Einschätzung der eigenen Leistung ab. Im Vergleich zur Experimentalklasse 1 ist jedoch der Anteil der Schüler/innen, die ihre Leistung völlig korrekt einschätzen konnten, zum Zeitpunkt t2 geringer, was dafür spricht, dass das selbstorganisierte Lehr-Lern-Arrangement für die Entwicklung der Selbsteinschätzungsfähigkeit der Experimentalklasse 2 nicht so förderlich war wie für die Experimentalklasse 1. Allerdings zeigt der Vergleich der Abweichungsindizes für die Experimentalklasse 2 eine noch günstigere Entwicklung an als für die Experimentalklasse 1, so dass der Vergleich der beiden Experimentalklassen keine eindeutige Entscheidung darüber zulässt, welche Klasse sich besser entwickelt hat.

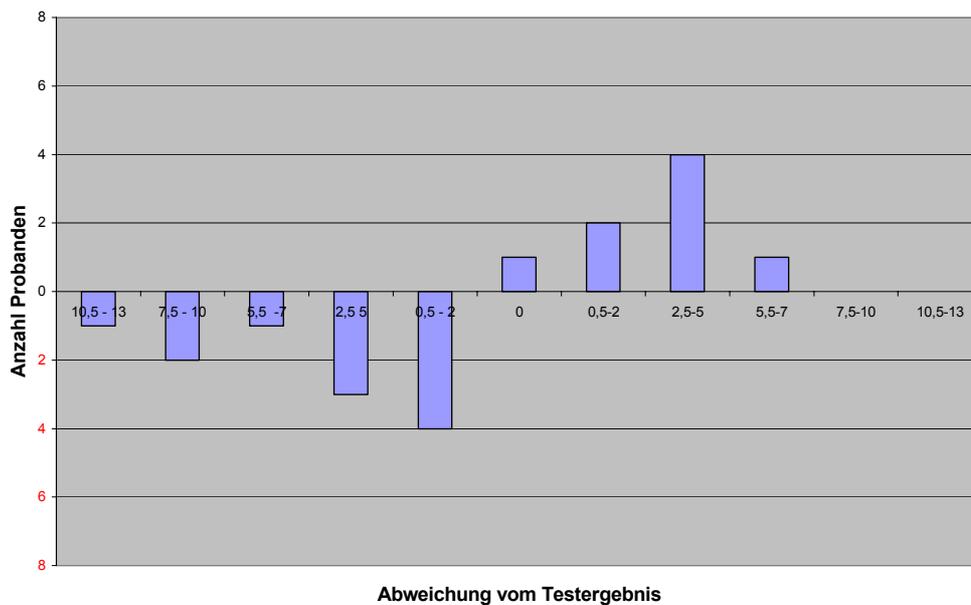
Interessant ist für unsere Fragestellung jedoch in erster Linie, wie sich die beiden Klassen im Vergleich zur Kontrollklasse entwickelt haben. Deshalb werden im folgenden Kapitel die Ergebnisse der Kontrollklasse dargestellt.

3.4 Entwicklung der Selbstbeurteilungsfähigkeit der Kontrollklasse

Zusätzlich zu den beiden Experimentalklassen erhielt eine parallele Kontrollklasse traditionellen Rechnungswesenunterricht ohne gezielte Selbstorganisation und reflexionsanregende Maßnahmen (kein Unterrichtsprotokoll). In dieser Klasse wurden die gleichen Produktdaten erhoben wie in den Experimentalklassen, so dass Vergleichsdaten zu den Schülerleistungen und der Selbsteinschätzung vorliegen.

Im folgenden wird dargestellt, wie sich die Selbstbeurteilungsfähigkeit dieser Kontrollklasse im Verlaufe des Unterrichts entwickelte. Dafür zeigt Abbildung 7 die Verteilung der Einschätzung der eigenen Leistungen der Schüler für den Pre-Test und Abbildung 8 zeigt die Verteilung für den Post-Test.

Abbildung 7: Häufigkeit der **geringeren** und **höheren** Selbsteinschätzung der Probanden der Kontrollklasse im Vergleich zur Fremdeinschätzung durch



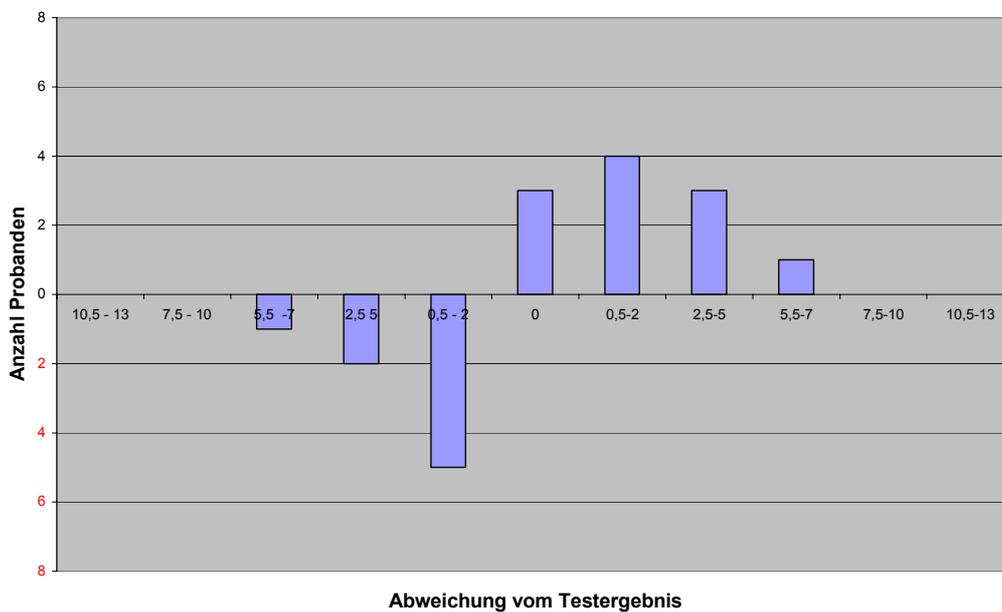
Experten im Pre-Test, N = 19

Die Abbildung 7 zeigt eine Verteilung, die mit denen der beiden Experimentalklassen vergleichbar ist. Es gibt sowohl Schüler/innen, die ihre eigenen Leistungen unterschätzten, als auch solche, die ihre Leistungen überschätzten. Dabei weist die Verteilung der Ergebnisse des Pre-Tests eine breite Streuung der Abweichungen auf. Eine richtige Einschätzung wurde nur von einem Probanden geleistet, und es finden sich deutlich mehr Unter- als Überschätzer (11 zu 7), wobei ein größerer Anteil der Schüler (6 von 19 Probanden = 32 %) nur leicht und ein großer Anteil (12 von 19

Probanden = 63 %) stärker bzw. stark von der korrekten Einschätzung abwich. Für diesen Messzeitpunkt wurde eine mittlere Abweichung von 4,03 berechnet ($s = 3,23$).

Im Vergleich dazu haben sich die Verhältnisse beim Post-Test etwas verändert. Die Ergebnisse sind in Abbildung 8 dargestellt.

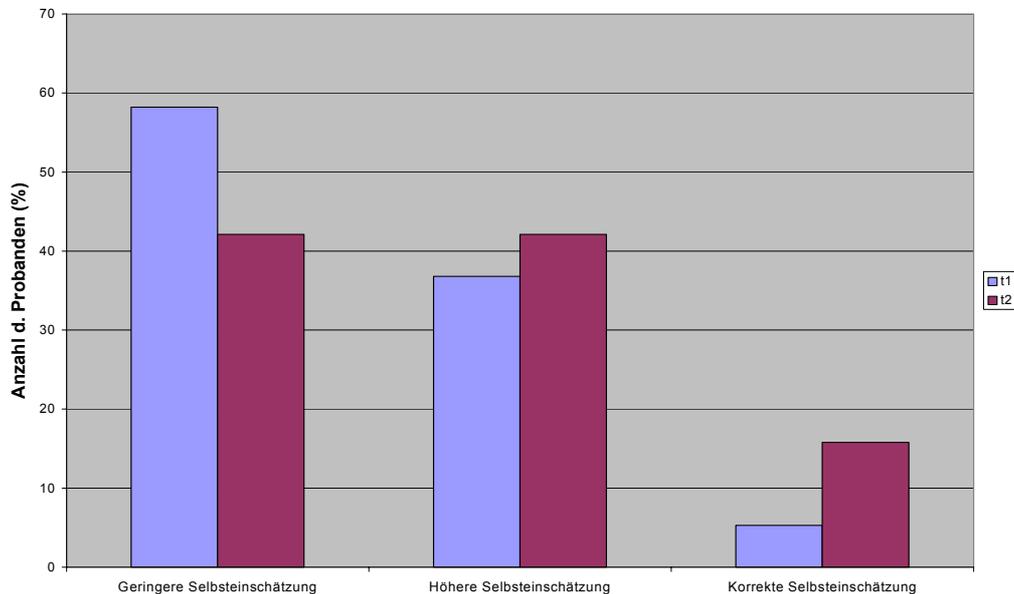
Abbildung 8: Häufigkeit der **geringeren** und **höheren** Selbsteinschätzung der Probanden der Kontrollklasse im Vergleich zur Fremdeinschätzung durch Experten im Post-Test, N = 19



Aus der Abbildung 8 ist zu erkennen, dass die Zahl der Probanden, die ihre Leistungen korrekt einschätzte, im Vergleich zum Pre-Test von 1 auf 3 angestiegen ist. 9 Schüler (= 47 %) wichen leicht von der korrekten Einschätzung ab und weitere 7 (= 37 %) wichen stärker ab, zwei davon um mehr als 5 Punkte. Während beim ersten Messzeitpunkt 11 Probanden (= 58 %) ihre eigene Leistung geringer eingeschätzt hatten, als sie tatsächlich war, waren es beim Post-Test nur noch 8 Probanden (42 %). Entsprechend hat die Zahl der Unterschätzer leicht zugenommen, sie ist von 7 (= 37 %) im Pre-Test auf 8 (= 42 %) im Post-Test angestiegen. Der Mittelwert der Abweichung für diesen Messzeitpunkt beträgt 2,06 ($s = 1,73$). Es hat also eine deutliche Verbesserung gegenüber dem ersten Messzeitpunkt gegeben.

Abbildung 9 soll die Veränderungen über den Untersuchungszeitraum hinweg verdeutlichen.

Abbildung 9: Entwicklung der Selbsteinschätzung der Probanden der Kontrollklasse im Vergleich der beiden Messzeitpunkte t1 und t2 (t1 = Pre-Test, t2 =



Post-Test). N = 19

Insgesamt hat sich die Selbstbeurteilungsfähigkeit der Schüler/innen der Kontrollklasse im Verlaufe des Schuljahres verbessert, obwohl keine besonderen Maßnahmen zur Förderung dieser Fähigkeit eingeführt wurden. Während beim ersten Messzeitpunkt mehr als die Hälfte der Probanden ihre Leistung geringer eingeschätzt hat, als sie tatsächlich war, hat dieser Anteil zum Post-Test hin abgenommen. Stattdessen haben die korrekte Selbsteinschätzung deutlich und die höhere Selbsteinschätzung leicht zugenommen. Beim Post-Test sind Über- und Unterschätzungen ausgeglichen und machen zusammen ca. 84 Prozent der Selbsteinschätzungen aus.

Auch aus diesen Ergebnissen lässt sich die Vermutung ableiten, dass sich eine eventuelle anfängliche Unterschätzung der eigenen Leistungen zugunsten einer realistischen bzw. höheren Selbsteinschätzung verändert hat. Um jedoch einen genaueren Einblick in die Veränderungen der drei Klassen nehmen zu können, soll im folgenden betrachtet werden, ob tatsächlich die Schüler/innen, die ihre eigenen Leistungen eher unterschätzten, im zweiten Test zu einer korrekten Einschätzung gelangt sind, oder ob die Selbsteinschätzungen stark variierten.

3.5 Konsistenz der Selbsteinschätzung über die Zeit

In der folgenden Tabelle wird für alle drei Klassen⁶ dargestellt, wie sich die Einschätzung der Schüler vom ersten zum letzten Messzeitpunkt verändert hat. Es ist aufgeführt, ob beispielsweise die Probanden, die ihre Leistung im Pre-Test unterschätzt haben, im Post-Test ebenfalls Unterschätzer waren, oder ob Veränderung über die Zeit zu beobachten sind. Die Betrachtung der Veränderung der Abweichungen der Schüler/innen zeigt, dass sie in der Art ihrer Selbsteinschätzung nicht gänzlich konsistent sind.

Tabelle 1: Veränderung in den Selbsteinschätzungen aller drei Klassen (N = 54)

Pre-Test	Post-Test	Anzahl Probanden			Summe
		EG 1	EG 2	KG	
Überschätzung	Überschätzung	3	7	4	14
Unterschätzung	Unterschätzung	4	4	7	15
Überschätzung	Unterschätzung	1	3	0	4
Unterschätzung	Überschätzung	2	3	4	9
Überschätzung	Korrekte Einschätzung	1	0	3	4
Unterschätzung	Korrekte Einschätzung	4	3	0	7
Korrekte Einschätzung	Überschätzung	0	0	0	0
Korrekte Einschätzung	Unterschätzung	0	0	1	1
Korrekte Einschätzung	Korrekte Einschätzung	0	0	0	0

Es ist in der Tabelle 1 sehr deutlich zu erkennen, dass mehr als die Hälfte der Schüler/innen (= 29 Probanden) in der Art ihrer Selbsteinschätzung (Über- bzw. Unterschätzung) über den Untersuchungszeitraum hinweg gleich geblieben sind. Weitere 9 Schüler/innen haben ihre eigene Leistung beim Pre-Test unterschätzt und beim Post-

⁶ Die Unterschiede zwischen den Klassen sind diesbezüglich vernachlässigbar gering.

Test dann überschätzt. Auffallend ist auch, dass diejenigen, die beim Post-Test zu einer korrekten Einschätzung gekommen sind, beim Pre-Test überwiegend zu der Gruppe der Unterschätzer zählten.

Dieser Befund stimmt mit den Vermutungen überein, die aus der Analyse der Abbildungen 5, 6 und 9 gezogen werden konnten (s. Kap. 3.3 und 3.4). Dort wurde beobachtet, dass in beiden Experimentalklassen die Zahl der Schüler/innen, die ihre eigene Leistung überschätzten, im Vergleich von Pre- zu Post-Test gleich geblieben ist, während der Anteil der Unterschätzungen gesunken ist. Es kann also festgehalten werden, dass eher diejenigen Schüler/innen zu einer korrekten Selbsteinschätzung gekommen sind, die sich zu Beginn der Untersuchung noch unterschätzten. Daraus kann die Vermutung abgeleitet werden, dass insbesondere die eher vorsichtigen, evtl. wenig selbstbewussten Schüler mehr Sicherheit durch die selbstorganisationsoffene Lernumgebung gewonnen haben.

Die Betrachtung der Rohdaten zeigt außerdem, dass einige der Schüler/innen keine und einige Schüler/innen erhebliche Veränderungen zwischen den beiden Messzeitpunkten aufweisen. Bei fünf Probanden ist die Abweichung der Selbsteinschätzung von der Fremdeinschätzung zu beiden Messzeitpunkten exakt gleich, wohingegen drei Probanden drastische Veränderungen zeigen. Eine Schülerin hat ihre Leistung beim ersten Test um 12,5 Punkte schlechter eingeschätzt, als sie tatsächlich war, und beim Post-Test um 3,5 Punkte besser. Eine andere Schülerin hatte beim Pre-Test eine um 10 Punkte höhere Selbsteinschätzung und beim Post-Test eine um 1,5 Punkte höhere. Eine weitere Schülerin schließlich hat sich beim Pre-Test um 10 Punkte zu schlecht eingeschätzt, beim Post-Test jedoch völlig korrekt. Bei diesen drei Schülerinnen wie auch bei fast allen anderen, die Veränderungen in ihrer Einschätzung erlebt haben, war die Tendenz eine positive: sie sind in ihrer Einschätzung treffsicherer geworden. Nur in fünf Fällen hat eine, wenn auch wenig dramatische, Verschlechterung der Einschätzung stattgefunden.

Damit bestätigt sich der durch die Betrachtung der einzelnen Klassen bereits entstandene Gesamteindruck, dass die Schüler/innen der drei an der Untersuchung beteiligten Klassen im Laufe des Schuljahres in ihrer Selbsteinschätzung sicherer wurden, und zwar unabhängig davon, ob sie in einer selbstorganisationsoffenen Lernumgebung lernten oder traditionell unterrichtet wurden. Dennoch gibt es Unterschiede zwischen der Experimentalgruppe und der Kontrollgruppe, die im folgenden diskutiert werden.

3.6 Zusammenfassender Vergleich der Experimentalgruppe mit der Kontrollgruppe

Als erstes bemerkenswertes Resultat kann festgehalten werden, dass nach Ablauf des Versuchs die Zahl der Teilnehmer, die ihre Leistung korrekt einschätzen konnten,

stark anstieg. Dies gilt sowohl für die beiden Experimentalklassen (Anstieg von 0 auf 5 (EG1) und Anstieg von 0 auf 3 (EG2)) als auch für die Kontrollklasse (Anstieg von 1 auf 3). Wenngleich dieses Resultat keiner statistischen Überprüfung unterzogen werden kann, lässt es doch die Aussage zu, dass die Verbesserung der beiden Experimentalklassen insgesamt positiver zu bewerten ist als die der Kontrollklasse. Beachtlich ist insbesondere, dass vor allem in der Experimentalklasse 1 ein Drittel der Schüler/innen beim Post-Test die eigene Leistung vollkommen korrekt eingeschätzt hat. In dieser Klasse hat sich auch der Index für die durchschnittliche Abweichung sehr positiv verändert: Von anfänglich 4,2 Punkten sank die Maßzahl auf 1,6 Punkte. Ein Vergleich der Entwicklung der drei Klassen in Bezug auf die Abweichungsindizes ist in der folgenden Tabelle 2 dargestellt.

Tabelle 2: Vergleich der Entwicklung der Selbstbeurteilungsfähigkeit aller drei Klassen über die beiden Messzeitpunkte hinweg

	N	Pre-Test		Post-Test		Differenz zwischen Post- und Pre-Test	
		Treffer	Mittlere Abweichung	Treffer	Mittlere Abweichung	Treffer	Mittlere Abweichung
EG 1	15	0	4,2	5	1,6	+5	-2,6
EG 2	20	0	4,9	3	1,2	+3	-3,7
KG	19	1	4,0	3	2,1	+2	-1,9

Betrachtet man nicht nur die absolute Anzahl der richtig eingeschätzten Leistung (Treffer), sondern auch die Veränderung der mittleren Abweichung, so zeigt sich, dass sich die zweite Experimentalklasse sogar noch stärker verbessert hat als die erste. Bei ihr betrug die mittlere Abweichung des Pre-Tests 4,9 Punkte und beim Post-Test nur noch 1,2 Punkte. Das ist eine Verbesserung um 3,7 Punkte. Die Schüler/innen dieser Klasse sind also in ihrer Selbsteinschätzung sehr viel sicherer geworden. Die Veränderung der mittleren Abweichung der Kontrollklasse verdeutlicht, dass diese Klasse zum Zeitpunkt des Pre-Tests weniger große Abweichungen (4,0 Punkte) von der korrekten Einschätzung der eigenen Leistung vorwies als die beiden Experimentalklassen, sie sich im Vergleich dazu jedoch nicht so stark verbessert hat wie diese beiden Klassen (Abweichungswert Post-Test = 2,1 Punkte). Die Unterschiede zwischen dem ersten und dem letzten Messzeitpunkt sind für die beiden Experimentalklassen höchst signifikant (EG1: $p = .002$; EG 2: $p = .000$) und für die Kontrollklasse signifikant ($p = .015$). Die Unterschiede zwischen den Gruppen sind

allerdings nicht signifikant. Dennoch kann man begründet davon ausgehen, dass zwischen der ausgesprochen positiven Entwicklung der Experimentalklassen und dem Treatment ein positiver Zusammenhang besteht.

Diese Entwicklung der Selbstbeurteilungsfähigkeit ist insbesondere deshalb beachtlich, weil die Schüler in den selbstorganisationsoffenen Lehr-Lern-Arrangements in erster Linie selbst für die Einschätzung ihrer Leistungen verantwortlich sind. Während die Schüler im traditionellen Unterricht normalerweise zu jedem Sachverhalt von der Lehrperson mitgeteilt bekommen, was richtig und was falsch ist, korrigieren sich die Schüler im SoLe-Unterricht weitgehend selbst bzw. gegenseitig. Die Lehrperson greift nur ein, wenn sie merkt, dass die Schüler alleine nicht weiterkommen, und auch dann korrigiert sie nicht einfach, sondern leitet zur Fehlerfindung an. Diese in der hier präsentierten Studie praktizierte Maßnahme hat offensichtlich besser als die im traditionellen lehrerzentrierten Unterricht praktizierte Fehlerkorrektur dazu geführt, dass die Schüler/innen der beiden Experimentalklassen lernen konnten, ihre Fähigkeiten selbst realistischer einzuschätzen.

4. Diskussion und Ausblick

Für alle drei Klassen hat sich eine Verschiebung von der anfänglich deutlichen Unterschätzung der eigenen Testleistung hin zu einer realistischeren Einschätzung ergeben, die nicht mehr so stark vom tatsächlich erreichten Testergebnis abweicht. Diese realistischeren Einschätzung könnte daraus resultieren, dass die Schüler im Verlaufe eines Schuljahres den Unterrichtsinhalt zunehmend besser kennen lernen und deshalb zunehmend besser einschätzen können, wie gut sie diesen Stoff beherrschen. Für diese Vermutung spricht, dass das beim Post-Test abgefragte Wissen sich direkt an dem im Unterricht behandelten Stoff orientierte, während beim Pre-Test Vorwissen gefordert war, das von den Schülern nicht aktuell erworben wurde. Ein weiterer Grund für die bessere Selbsteinschätzung beim letzten Messzeitpunkt könnte sein, dass Lernende im Unterricht im allgemeinen eine Rückmeldung darüber erhalten, ob sie einen Stoff beherrschen oder nicht, entweder als direkte Rückmeldung von Lehrperson zu Schüler, etwa im Falle einer Schülermeldung, oder als Beobachtung im Unterrichtsverlauf, etwa wenn man feststellt, dass man den Stoff so verstanden hat, wie er bei der Präsentation auch dargestellt wird. Diese permanente Rückmeldung über die Richtigkeit des Stoffverständnisses könnte zu einer größeren Sicherheit in Bezug auf die Einschätzung der eigenen Kenntnisse des Stoffes führen. Während jedoch bei den Experimentalgruppenschülern die erworbene Sicherheit eine selbst erarbeitete Sicherheit ist, hängt die Sicherheit der traditionell unterrichteten Schüler in erster Linie davon ab, wie die Lehrperson mit den im Unterricht gemachten Schüleräußerungen und -fehlern umgeht.

Darüber, welche einzelnen Faktoren für die bessere Entwicklung der Selbstbeurteilungsfähigkeit in den beiden Experimentalgruppen ursächlich verantwortlich sind, können verschiedene Vermutungen angestellt werden. Ergebnisse früherer Untersuchungen zum Selbstorganisierten Lernen zeigen beispielsweise, dass die Experimentalklassenschüler ein ganz anderes Frageverhalten entwickeln als die traditionell unterrichteten Schüler. Zum einen stellen sie deutlich mehr Fragen als die Vergleichsgruppenschüler, zum anderen haben diese Fragen deutlich öfter einen „Deep-Reasoning-Charakter“ (vgl. *Sembill* im gleichen Band; *Sembill, Schumacher, Wolf, Wuttke & Santjer-Schnabel* 2001). Angesichts der Tatsache, dass die Lernenden der Experimentalklassen überwiegend in Gruppen an komplexen Problemen arbeiteten, sich dabei über ihre Fähigkeiten und Kenntnisse gegenseitig austauschten und sie ihre Arbeits- und Lernerfolge in ihre Lernprotokolle eintragen sollten, liegt außerdem die Vermutung nahe, dass eine vertiefte Reflexion der Lernvorgänge stattgefunden hat. Die Ergebnisse der vorliegenden Studie zeigen allerdings eine offensichtlich geringe Nutzung des Lernprotokolls, das die Schüler am Ende jeder Unterrichtseinheit zu schreiben hatten. Diese Protokolle dienten der Anregung von bewussten Reflexionen des eigenen Lernens, enthielten jedoch keine gezielten Aufforderungen zur Selbstbeobachtung und Selbstbewertung. Bei der Analyse der von den Schülern erstellten Protokolle stellte sich heraus, dass sie durchgängig keine selbst-reflexiven Elemente enthielten und folglich keine Anhaltspunkte über die Selbstbeurteilungsfähigkeit der Lernenden geben können. In der weiteren Forschungsarbeit stehen deshalb gezieltere Anregungen zur Reflexion und eine bessere Einbindung dieser Maßnahmen in das unterrichtliche Lernen bei der Suche nach geeigneten Fördermaßnahmen im Vordergrund.

Das in Kapitel 3.5 dargestellte Ergebnis, dass in erster Linie solche Experimentalklassenschüler beim Post-Test zu einer korrekten Einschätzung ihres Testergebnisses gelangt sind, die beim Pre-Test zur Gruppe der Unterschätzer gezählt wurden, spricht im Übrigen dafür, dass insbesondere die unsicheren Schüler durch das selbstorganisationsoffene Lehr-Lern-Arrangement mehr Selbstvertrauen gewonnen haben. Die Vergleiche von traditionellem und selbstorganisiertem Lernen in dieser Studie und in früheren Studien zeigen hierzu, dass auch die schwächeren Schüler bzw. Klassen stark von der selbstorganisierten Lernumgebung profitieren (z.B. *Seifried, Brouër & Sembill* in Vorb.). Um diesen Effekt besser interpretieren und erklären zu können, werden z.Zt. Audio- und Videoanalysen durchgeführt, über die zu einem späteren Zeitpunkt berichtet werden kann.

Im Hinblick auf die Förderung der Selbstbeurteilungsfähigkeit der Schüler/innen sind die Ergebnisse dieser Studie ermutigend, denn sie zeigen, dass das selbstorganisierte Lehr-Lern-Arrangement prinzipiell einen geeigneten Rahmen dafür bietet. Im Verlaufe der insgesamt ca. 72 Unterrichtsstunden, die nach den Prinzipien des Selbstorganisierten Lernens gestaltet wurden, gelang es den Schülern zunehmend besser, richtig einzuschätzen, wie viel Wissen sie in einem Test reproduzieren kön-

nen. Obwohl die Kontrollgruppe im Verlaufe des Schuljahres ebenfalls eine positive Entwicklung in Bezug auf die Selbstbeurteilungsfähigkeit erlebt hat, lassen die Daten die Schlussfolgerung zu, dass die positivere Entwicklung der Experimentalgruppe auf die selbstorganisationsoffene Organisation des Unterrichts zurückzuführen ist.

Abschließend ist noch darauf hinzuweisen, dass die Selbstbeurteilung sich in dieser Studie auf das Reproduzieren von Wissen bezog. Es ist jedoch zu vermuten, dass die Selbstbeurteilung bei komplexeren Aufgaben wie z.B. dem Lösen von Problemen schwerer fällt. Auch dafür kann die Studie Anhaltspunkte geben, denn es wurde ein Instrument zur Erfassung der Problemlösefähigkeit eingesetzt, das auch zum Vergleich der Selbst- und Fremdeinschätzung heran gezogen werden kann. Diese Daten sind jedoch noch nicht restlos ausgewertet und können erst zu einem späteren Zeitpunkt berichtet werden.

Literatur

- Barton, J. & Collins, A.* (Ed.) (1997). *Portfolio Assessment: A Handbook for Educators*. Menlo Park: Addison-Wesley
- Beck, E., Guldemann, T. & Zutavern, M.* (1997) (Hrsg.). *Lernkultur im Wandel*. Sankt Gallen: UVK
- Black, P. & William, D.* (1998). *Assessment and Classroom Learning*. In: *Assessment in Education*, 5 (1), 7-74
- Boekarts M.* (1997). *Self-Regulated Learning: A new Concept Embraced by Researchers, Policy Makers, Educators, Teachers, and Students*. In: *Learning and Instruction*, 7 (2), 161-186
- Brouër, B.* (2001). *Förderung der Wahrnehmung von Lernprozessen durch die Anwendung der Basismodelle des Lernens bei der Gestaltung von Unterricht*. In: *Unterrichtswissenschaft 2/2001*, S. 153-170
- Brouër, B., Seifried, J. & Sembill, D.* (2001). *Selbstorganisiertes Lernen im Rechnungswesen-Unterricht (1)*. In: *VLB Akzente 04-05/01 (10. Jahrgang)*, S. 17-20
- Brunner, I. & Schmidinger, E.* (2000). *Gerecht beurteilen. Portfolio: die Alternative für die Grundschulpraxis*. Linz: Veritas
- Cameron, J. & Pierce, D.P.* (1994). *Reinforcement, reward, and intrinsic motivation: a metaanalysis*. In: *Review of Educational Research*, 64, 363-423
- Heymann, H. W.* (2001). *Lernergebnisse sichern – aber wie?* In: *Pädagogik 10/2001*, S. 7-9
- Jürgens, E. & Sacher, W.* (2000). *Leistungserziehung und Leistungsbeurteilung: schulpädagogische Grundlegung und Anregung für die Praxis*. Neuwied & Kriftel: Luchterhand

- Oser, F. & Patry, J.-L.* (1990). Choreographien unterrichtlichen Lernens. Pädagogisches Institut der Universität Freiburg (Schweiz), Berichte zur Erziehungswissenschaft, Nr. 89
- Rheinberg, F.* (2001). Bezugsnormen und schulische Leistungsbeurteilung. In: *Weinert, F.E.* (Hrsg.). Leistungsmessungen in Schulen. Weinheim & Basel: Beltz Verlag, S. 59-71
- Sacher, W.* (1994). Prüfen – Beurteilen – Benoten. Bad Heilbrunn: Klinkhardt
- Schunk, D.H. & Zimmerman, B.J.* (1994) (Ed.). Self-regulation of learning and performance. Issues and educational applications. Hillsdale, NJ: Erlbaum
- Schusser, G.* (1996). Leistungsbeurteilung in der Schule. Osnabrück: Staperfeld
- Seifried, J.* (2002). Selbstorganisiertes Lernen im Rechnungswesen. In: Zeitschrift für Berufs- und Wirtschaftspädagogik, 98. Jg., H.1, S. 104-121
- Seifried, J., Brouër, B. & Sembill, D.* (2001). Selbstorganisiertes Lernen im Rechnungswesen-Unterricht (2). In: VLB Akzente 06/01 (10. Jahrgang), S. 21-22
- Seifried, J., Brouër, B. & Sembill, D.* (in Vorbereitung). Was lernen Schülerinnen und Schüler im selbstorganisationsoffenen Rechnungswesenunterricht?
- Seifried, J. & Sembill, D.* (2002). Emotionales, motivationales und kognitives Erleben im Fach Rechnungswesen (2). VLB-Akzente H. 6, 11. Jg., S. 19-21.
- Sembill, D.* (1992). Problemlösefähigkeit, Handlungskompetenz und Emotionale Befindlichkeit. Zielgrößen Forschenden Lernens. Göttingen, Toronto, Zürich: Hogrefe
- Sembill, D.* (1994). Feststellung und Bewertung selbstorganisierter Lernprozesse. In: *Olechowski, R. & Rollett, B.* (Hrsg.). Theorie und Praxis: Aspekte empirisch-pädagogischer Forschung. Frankfurt/M. et al.: Peter Lang, S. 258-264
- Sembill, D.* (1997). Prozessanalysen Selbstorganisierten Lernens. Zwischenbericht an die Deutsche Forschungsgemeinschaft im Rahmen des Schwerpunktprogramms „Lehr-Lern-Prozesse in der kaufmännischen Erstausbildung“, Juli 1997
- Sembill, D.* (2000). Selbstorganisiertes und Lebenslanges Lernen. In: *Achtenhagen, F. & Lempert, W.* (Hrsg.). Lebenslanges Lernen (IV): Formen und Inhalte von Lernprozessen. Opladen: Leske & Budrich, S. 60-90
- Sembill, D., Wolf, K.D., Wuttke, E., Santjer, I. & Schumacher, L.* (1998). Prozessanalysen Selbstorganisierten Lernens. In: *Beck, K. & Dubs, R.* (Hrsg.). Kompetenzerwerb in der Berufserziehung - Kognitive, motivationale und moralische Dimensionen kaufmännischer Qualifizierungsprozesse. Zeitschrift für Berufs- und Wirtschaftspädagogik, Beiheft Nr. 14, Stuttgart: Steiner Verlag

- Sembill, D., Schumacher, L., Wolf, K. D., Wuttke, E. & Santjer-Schnabel, I.* (2001). Förderung der Problemlösefähigkeit und der Motivation durch Selbstorganisiertes Lernen. In: *Beck, K. & Krumm, V.* (Hrsg.). *Lehren und Lernen in der beruflichen Erstausbildung. Grundlagen einer modernen kaufmännischen Berufsqualifizierung.* Opladen: Leske + Budrich, S. 257-281
- Sembill, D., Seifried, J. & Brouër, B.* (2002). SoLe – ein Schritt in die Zukunft? In: *VLB Akzente*, 11 (2), S. 13-15
- Stallings, V. & Tascione, C.* (1996). Student Self-Assessment and Self-Evaluation. In: *The Mathematics Teacher*, 89 (7), 548-554
- Stefani, L.A.J.* (1994). Peer, Self and Tutor Assessment: relative reliabilities. In: *Studies in Higher Education*, V. 19, No. 1, pp. 69-75
- Tent, L.* (1998). Zensuren. In: D.H. Rost. *Handwörterbuch Pädagogische Psychologie.* Weinheim & Basel: Beltz, S. 580-584
- Wagner, B.* (1999). *Lernen aus der Sicht der Lernenden.* Frankfurt et al.: Lang
- Wagner, B.* (2000). Wahrnehmung und Beurteilung von Lernprozessen durch Schülerinnen und Schüler. Unveröffentlichtes Vortragsmanuskript (59. Tagung der Arbeitsgruppe für Empirische Pädagogische Forschung in Bremen)
- Weinert, F.E.* (Hrsg.) (2001). *Leistungsmessungen in Schulen.* Weinheim & Basel: Beltz Verlag
- Zimmerman, B.J.* (1990). Self-Regulated Learning and Academic Achievement: An Overview. In: *Educational Psychology*, 25(1), 3-17