

## Reflexionsfortschritte im Umgang mit einem bioethischen Problem

### Versuch einer empirischen Überprüfung

Der Autor stellt Ergebnisse einer Arbeit vor, die den Reflexionsfortschritt empirisch zu überprüfen versuchte, den Schülerinnen und Schüler in der unterrichtlichen Auseinandersetzung mit bioethischen Problemen des therapeutischen Klonens vollzogen haben. Dabei zeigt er exemplarisch Möglichkeiten der empirischen Evaluation des Philosophie- und Ethikunterrichtes.

#### 1. Einleitung

In den letzten Jahren gab es einige Versuche, didaktische Kompetenzmodelle für den Philosophie- und Ethikunterricht zu entwickeln. Sie orientieren sich in der Regel an philosophischen *Methoden* und suchen diese in *Fähigkeiten* zu übersetzen, die die Schülerinnen und Schülern entwickeln sollen. So zeigt ein Schüler etwa nach einem Modell Christian Geferts in Anschluss an Ekkehard Martens philosophische Reflexionsfähigkeit, wenn er 1. von brüchig gewordenen Alltagsdeutungen ausgeht und empirische Daten berücksichtigt oder umfassend beschreibt, was sie oder er in Bezug auf diese Deutungen beobachtet, erfahren oder wahrgenommen hat, und ob sie oder er dabei die Ergebnisse der Einzelwissenschaften einbezogen hat (phänomenologische Methode); 2. Lehrmeinungen und Interpretationen oder ideengeschichtliches Wissen als Deutungsangebote und die eigenen, alltäglich Deutungsmuster heranzieht (hermeneutische Methode); 3. zentrale Begriffe und Argumente innerhalb des Prozesses der Deutungsfindung hervorhebt und prüft (analytische Methode); 4. unterschiedliche Deutungsangebote zuspitzt und gegeneinander abwägt (dialektische Methode); 5. ungeschützte Einfälle und Phantasien bei der Entwicklung von Deutungen zulässt und sie spielerisch erprobt (spekulative Methode).<sup>1</sup>

Weitgehend unerforscht ist jedoch bislang, ob solche Modelle, die ihre Kategorien deduktiv aus dem etablierten philosophischen Methodenkanon ableiten, sich

auch empirisch erhärten lassen. Darin spiegelt sich die grundsätzliche Schwierigkeit wider, eine so komplexe Größe wie philosophische oder ethische „Reflexionsfähigkeit“ sinnvoll zu operationalisieren und für Messungen brauchbar zu machen.

In einem Philosophie-Grundkurs des elften Jahrgangs einer Gesamtschule habe ich unter Rückgriff auf Geferts Kompetenzmodell eine solche Messung versucht.<sup>1</sup> Das Unternehmen hatte eher experimentellen Charakter: Erfasst wurden insgesamt siebzehn Schülerinnen und Schüler; die Stichprobe ist also sehr klein, die statistische Aussagekraft der Ergebnisse gering. Dennoch mag der Versuch dazu anregen, die Möglichkeiten empirischer Messung auch für die Evaluation des Philosophie- und Ethikunterrichtes zu nutzen.

#### 2. Zur untersuchten Unterrichtsreihe

Meine Untersuchung bezog sich auf eine problemorientiert strukturierte Unterrichtsreihe zu den ethischen Fragen, welche die (noch hypothetischen) Möglichkeiten therapeutischen Klonens aufwerfen.<sup>2</sup> Dabei hatte ich Geferts Modell auch in der Planung der Reihe berücksichtigt. Zunächst erarbeiteten die Schüler mit Hilfe einer Internet-Recherche Aspekte zur Embryonalentwicklung, zum technischen Verfahren des therapeutischen Klonens und zur nationalen und internationalen Rechtslage (vgl. Dimension 1). Es folgte eine intensive Auseinandersetzung mit den Positionen Robert Spaemanns, Norbert Hoersters, Peter Singers und

<sup>1</sup> Gefert, Christian: Bildungsziele, Kompetenzen und Anforderungen – Perspektiven für die Entwicklung von Bildungsstandards in philosophischen Bildungsprozessen. In: Martens, Ekkehard/Gefert, Christian/Steenblock, Volker (Hrsg.): Philosophie und Bildung. Beiträge zur Philosophiedidaktik. Münster 2005. S. 141. Ein weit differenzierteres, aber auch schwerer handhabbares Modell hat Johannes Rohbeck entwickelt, das sechsunddreißig verschiedene Kompetenzen benennt (Philosophische Kompetenzen. In: Ders.: Didaktik der Philosophie und Ethik. Dresden 2008. S. 91–104).

<sup>2</sup> Die Messung hat im Rahmen einer von Matthias Schulze betreuten Examensarbeit stattgefunden. Meine Kollegin Laura Nowack hat die evaluierte Unterrichtsreihe gemeinsam mit mir geplant und im *team-teaching* durchgeführt. Beiden sei für ihre Unterstützung gedankt.

Dietmar Mieths (vgl. Dimension 2). Dabei standen in der Diskussion dieser Positionen das Konzept der „Menschenwürde“ und Kriterien der Zuschreibung des Mensch- bzw. Personbegriffs im Mittelpunkt; in diesem Zusammenhang experimentierten die Schülerinnen und Schüler unter anderem mit verschiedenen Visualisierungsmöglichkeiten für Begriffsumfänge (vgl. Dimension 3). Die so gewonnenen Erkenntnisse haben sie immer wieder an konkreten Beispielen erprobt (vgl. Dimension 5). Zum Abschluss der Reihe teilten sich die Schülerinnen und Schüler in Gruppen auf, die jeweils eine der behandelten Positionen vertraten. Diese sollten sie bei einem Besuch des „Deutschen Referenzzentrums für Ethik in den Biowissenschaften“ in Bonn durch eigene Recherche in der dortigen Bibliothek argumentativ unterfüttern und auf Thesepapieren darstellen. Schließlich haben sie die Argumente der jeweiligen Gruppen in einer sehr engagierten Podiumsdiskussion miteinander konfrontiert (vgl. Dimension 4).

Die Reflexionsfähigkeit der Schülerinnen und Schüler versuchte ich durch Konfrontation mit einer konkreten Entscheidungssituation zu diagnostizieren. Sie erhielten als Einstieg in die Unterrichtsreihe ein Arbeitsblatt mit folgender Dilemmageschichte:

#### Rettung durch therapeutisches Klonen?

Paul hatte im vergangenen Sommer einen schlimmen Sportunfall, bei dem er sich schwere Rückenverletzungen zuzog. Seitdem ist er vom Lendenwirbelbereich ab bis zu den Füßen gelähmt. Paul hat zwar bei der Rehabilitation gelernt, weitgehend selbständig zurechtzukommen; dennoch ist sein Leben mit vielen Einschränkungen verbunden.

Stellen wir uns vor, Paul würde in der näheren Zukunft leben, und das therapeutische Klonen wäre schon anwendbar. Dann wäre es für ihn ein Angebot, das ihm Heilung in Aussicht stellen würde. Bei diesem Verfahren werden menschliche Embryonen künstlich hergestellt. Sind sie einige Tage alt, werden ihnen so genannte „Stammzellen“ entnommen. Solche Zellen sind in der Lage, sich beliebig oft durch Teilung zu vermehren; außerdem kann man dafür sorgen, dass sie sich zu Zellen mit ganz unterschiedlicher Spezialisierung entwickeln. Man könnte daraus zum Beispiel auch Nervenzellen des Rückenmarks

herstellen, um sie anschließend auf Paul zu übertragen. Wenn er Glück hätte, könnten sich die Lähmungen dadurch zurückbilden. Die Entnahme der Stammzellen verläuft allerdings für die Embryonen auf jeden Fall tödlich.

Paul bekommt eine solche Therapie angeboten. Er denkt lange darüber nach. Schließlich entscheidet er sich dagegen.<sup>3</sup>

Die Schülerinnen und Schüler sollten Pauls Entscheidung zunächst beurteilen („Ich meine, Pauls Entscheidung war ...“), indem sie sechs Kästchen zwischen „völlig falsch“ und „völlig richtig“ ankreuzten – durch die gerade Anzahl von Antwortmöglichkeiten (d. h. die fehlende Mittelposition) mussten sie Stellung beziehen („forced choice“). Danach sollten sie angeben, wie groß ihr Entscheidungskonflikt war, d. h. wie leicht oder schwer ihnen die Beurteilung gefallen ist (sechs Ankreuzmöglichkeiten zwischen „sehr leicht“ und „sehr schwer“). Schließlich wurden sie gebeten, stichwortartig die Gründe für ihre Beurteilung darzustellen.<sup>4</sup>

Das gleiche Arbeitsblatt erhielten die Schülerinnen und Schüler ein zweites Mal zum Abschluss der Reihe. Damit ergab sich die Möglichkeit, den Stand der Reflexionsfähigkeit der Schülerinnen und Schüler – im Sinne einer sogenannten *Prä-post-Messung* – zu zwei Erhebungszeitpunkten, nämlich vor und nach der Unterrichtsreihe, zu vergleichen. Dazu habe ich die Begründungen der Schüler nach Geferts Kategorien eingeschätzt. Um hierbei eine gewisse Objektivität zu gewährleisten, habe ich die Bewertung von einer Kollegin durchführen lassen. Sie erhielt die Begründungen anonymisiert und in transkribierter Form. Die Kollegin sollte auf einer Skala zwischen 0 („gar nicht“) und 5 („in großem Maße“) einschätzen, ob und inwieweit die Schüler

- empirisches Wissen zur Embryonalentwicklung, zur Technik des therapeutischen Klonens und zur Rechtslage, aber auch alltägliches Erfahrungswissen einfließen lassen,
- Kenntnisse über relevante philosophisch-ethischen Positionen heranziehen,
- wichtige Begriffe (wie „Lebensrecht“, „Menschenwürde“ etc.) klar und reflektiert verwenden,
- sorgfältig begründen und dabei unterschiedliche Argumente abwägen

<sup>3</sup> Für einen Überblick siehe: *Oduncu, Fuat/Schroth, Ulrich/Vossenkuhl, Wilhelm* (Hrsg.): *Stammzellenforschung und therapeutisches Klonen*. Göttingen 2002. Diese Dilemmageschichte geht auf eine Idee von Dr. Michael Fuchs, Institut für Wissenschaft und Ethik Bonn, zurück

<sup>4</sup> Der Aufbau dieses Arbeitsblattes orientiert sich an Vorschlägen von Georg Lind (*Moral ist lehrbar*. Stuttgart 2003. S. 96f.). Er propagiert die Verwendung solcher Dilemmageschichten in Anlehnung an die Theorie Kohlbergs. Diese soll hier jedoch – auch in der Diagnose der „Reflexionsfähigkeit“ der Schülerinnen und Schüler – keine Rolle spielen.

- sowie Phantasie und Kreativität im Umgang mit dem Problem zeigen.

Die so gewonnenen Daten sollten zum einen Aufschluss darüber geben, ob und wie die Schülerinnen und Schüler im Laufe der Unterrichtsreihe ihre ursprüngliche Haltung verändert hatten und ob sie in ihrem Urteil dabei sicherer oder unsicherer wurden. Zum anderen sollte den Daten zu entnehmen sein, in welchen Bereichen sich die Reflexionsfähigkeit der Schüler verändert hatte.

### 3. Ergebnisse

Wie Tabelle 1 zu entnehmen ist, stimmten vor der Unterrichtsreihe fast doppelt so viele Schüler *pro* reproduktives Klonen (lehnten also das Verhalten des Protagonisten in der Dilemmageschichte ab) wie *contra*; nach der Unterrichtsreihe positionierten sich etwa gleich viele Schüler *pro* und *contra* therapeutisches Klonen. Der Beurteilungswechsel fand lediglich von *pro* nach *contra* statt; er fand sich ausschließlich bei Mädchen.

Tabelle 1: Beurteilung des therapeutischen Klonens (erste Zeile: gesamt, zweite Zeile: nach Geschlechtern getrennt)

	Contra	Pro	Gesamt
Erster Erhebungszeitpunkt (vor der Unterrichtsreihe)	6 (4♂ + 2♀)	11 (3♂ + 8♀)	17 (7♂ + 10♀)
Zweiter Erhebungszeitpunkt (nach der Unterrichtsreihe)	9 (4♂ + 5♀)	8 (3♂ + 5♀)	17 (7♂ + 10♀)

Hat die Unterrichtsreihe eher zu einer Verunsicherung oder zu einer Festigung im Urteil der Schülerinnen und Schüler geführt? Um das herauszufinden, habe ich eine Variable „Eindeutigkeit“ gebildet, die darüber Aufschluss gibt, wie dezidiert sie *pro* oder *contra* Stellung bezogen haben. Diese Variable nimmt den höchsten Wert 2 an, wenn bei der ersten Frage – es ging um die Beurteilung von Pauls Verhalten auf einer sechsstufigen Skala – eine der beiden äußeren Ankreuzmöglichkeiten („völlig falsch“ oder „völlig richtig“) markiert wurde; die Werte 0 und 1 bilden einen niedrigen bzw. mittleren Grad der Eindeutigkeit ab.

Weiterhin habe ich aufgrund der zweiten Frage die Variable „Konflikt“ gebildet; sie nimmt den höchsten Wert 5 an, wenn die Schüler angaben, sich bei ihrem Urteil „sehr unsicher“ gefühlt zu haben, und entsprechend, auf der anderen Seite der Skala, den Wert 0 bei der Antwort „sehr sicher“.

Interessant ist zunächst, dass Eindeutigkeit und Konflikt kaum korrelieren ( $r = -.20$  für den ersten Erhebungszeitpunkt und  $-.14$  für den zweiten;  $p$  jeweils weit oberhalb des kritischen -Fehlerniveaus von 5%<sup>5</sup>); die beiden Dimensionen scheinen also tatsächlich trennscharf zu sein. Die Abbildungen 1 und 2 zeigen Eindeutigkeit und Konflikthaftigkeit der Beurtei-

lung – getrennt nach den beiden Erhebungszeitpunkten und nach der Beurteilung *pro* und *contra*.

Abbildung 1: Eindeutigkeit in Abhängigkeit von Erhebungszeitpunkt (EZP) und Beurteilung

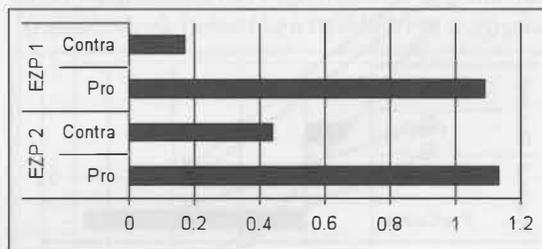
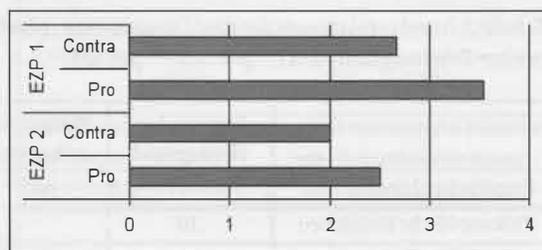


Abbildung 2: Konflikt in Abhängigkeit von Erhebungszeitpunkt (EZP) und Beurteilung



<sup>5</sup>  $r$  ist eine Abkürzung für den Korrelationskoeffizienten, der sich zwischen -1 und +1 (vollkommen negativer bzw. positiver Zusammenhang) bewegen kann; je höher der Koeffizient (im Betrag), desto stärker der Zusammenhang.  $p$  bezieht sich auf die statistische Signifikanztestung und gibt den -Fehler wieder, hier: die Wahrscheinlichkeit, die Annahme eines Nullzusammenhangs *fälschlicherweise* zu verwerfen und einen substantiellen Zusammenhang anzunehmen; im Allgemeinen wird in den Sozialwissenschaften ein -Fehler  $< .05$  als statistische Signifikanzgrenze akzeptiert. Vgl. hierzu: Bortz, Jürgen/Döring, Nicola: Forschungsmethoden und Evaluation. Berlin 1999. S. 492ff.

Insgesamt ergibt sich, dass die Konflikthaftigkeit der Beurteilung bei allen Schülerinnen und Schülern durchgehend *abgenommen*, die intensive Beschäftigung mit dem Thema also zu einer *Festigung* des Urteils geführt hat – und zwar, was für die Ausgewogenheit des Unterrichts sprechen mag, sowohl bei Befürwortern als auch bei Gegnern des therapeutischen Klonens. Die Eindeutigkeit der Positionierung nimmt bei den Gegnern zu, bei den Befürwortern bleibt sie in etwa gleich.

Nun hatte sich oben gezeigt, dass eine kleine Gruppe von Schülerinnen ihre Position verändert hat. Kann man hier Besonderheiten hinsichtlich der Eindeutigkeit und Konflikthaftigkeit ihres Urteils erkennen? Um dies zu klären, habe ich Differenzwerte gebildet. Pro Schüler wurde jeweils vom Wert des zweiten Erhebungszeitpunkts derjenige des ersten Erhebungszeitpunkts abgezogen, so dass ein positiver Wert eine Zunahme der Eindeutigkeit bzw. Konflikthaftigkeit anzeigt. Die folgenden Abbildungen 3 und 4 zeigen *Veränderungen der Variablen* „Eindeutigkeit“ (Spannweite zwischen -2 und +2) und „Konflikt“ (Spannweite zwischen -5 und +5) je nach Beurteilungsstabilität und -wechsel.

Abbildung 3: Veränderungen von Eindeutigkeit in Abhängigkeit von Stabilität und Wechsel der Beurteilung

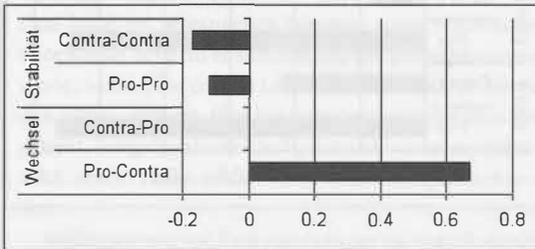
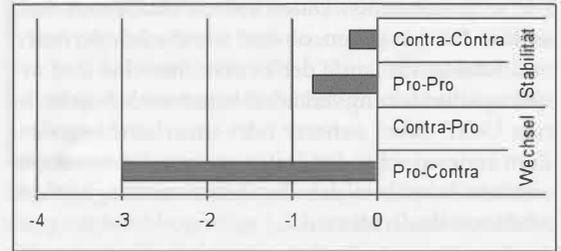


Abbildung 4: Veränderungen von Konflikt in Abhängigkeit von Stabilität und Wechsel der Beurteilung



Es fällt auf, dass bei den Schülerinnen und Schülern, deren Position stabil geblieben ist, die Konflikthaftigkeit, aber auch die Eindeutigkeit der Beurteilung ein wenig abgenommen hat. Das Urteil dieser Schüler ist also sowohl *sicherer* als auch *vorsichtiger* (weil differenzierter?) geworden. Die Schülerinnen, die einen Beurteilungswechsel vorgenommen haben, fallen aus diesem Muster jedoch heraus: Der Konflikt nimmt stark ab, die Eindeutigkeit der Beurteilung zu. Es sieht aus, als hätten die betreffenden Schülerinnen zu Beginn der Unterrichtsreihe dem therapeutischen Klonen unter Skrupeln zugestimmt und dann durch die Auseinandersetzung mit dem Thema eine ablehnende, nunmehr gefestigte Haltung entwickelt.

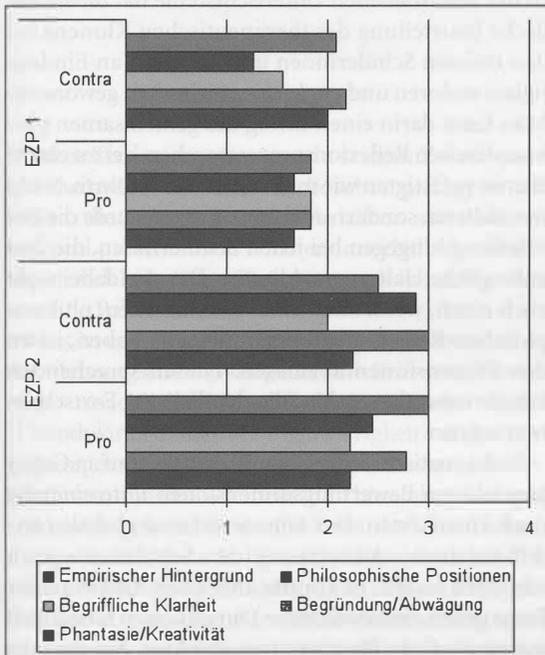
Inwiefern lässt sich nun aber an den Begründungen der Schülerinnen und Schüler eine Progression ihrer Reflexionsfähigkeit gemäß den fünf Gefertschens Dimensionen feststellen, die zugleich Planungsgrundlage der Unterrichtsreihe waren? In Tabelle 2 sind zunächst die Korrelationen zwischen den Dimensionen – inklusive der Signifikanztestungen – dargestellt.

Tabelle 2: Interkorrelationen der fünf Dimensionen (oberhalb der Diagonalen: erster Erhebungszeitpunkt, unterhalb: zweiter Erhebungszeitpunkt; \*  $p < .05$ , \*\*  $p < .01$ )

	Empirischer Hintergrund	Philosophische Positionen	Begriffliche Klarheit	Begründung/ Abwägung	Phantasie/ Kreativität
Empirischer Hintergrund	--	.61**	.75**	.81**	.73**
Philosophische Positionen	.10	--	.82**	.76**	.59*
Begriffliche Klarheit	.15	.68**	--	.82**	.61**
Begründung/ Abwägung	.71**	.61**	.56*	--	.83**
Phantasie/Kreativität	.70**	.56**	.65**	.80**	--

Die hohe Interkorrelation der Dimensionen – wenn man einmal von den Zusammenhängen zwischen „Empirischer Hintergrund“ und „Philosophischen Positionen“ sowie „Begrifflicher Klarheit“ zum zweiten Erhebungszeitpunkt absieht – wirft Fragen hinsichtlich ihrer Trennschärfe auf. Daher ist bei nachfolgenden Analysen einige Vorsicht bei der Interpretation der Ergebnisse geboten.

Abbildung 5: Die fünf Bewertungsdimensionen in Abhängigkeit von Erhebungszeitpunkt (EZP) und Beurteilung



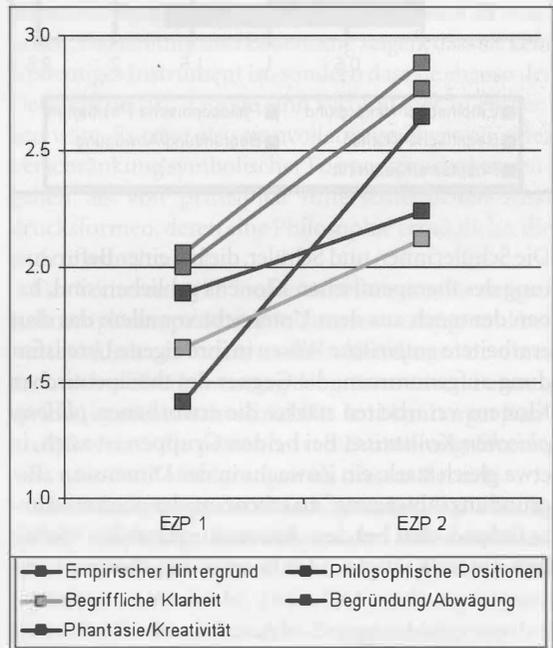
Wie Abbildung 5 zeigt (pro Schülerin/Schüler habe ich für jede Dimension vom Wert des zweiten Erhebungszeitpunkts denjenigen des ersten Erhebungszeitpunkts abgezogen, so dass ein positiver Wert einen Kompetenzzuwachs anzeigt), ist bei allen Schülerinnen und Schülern, gleich ob sie das therapeutische Klonen befürworteten oder ablehnten, ein deutlicher Fortschritt über alle Dimensionen hinweg erkennbar. Unterschiede hinsichtlich der fünf Dimensionen sind auf diesem Auflösungsgrad schwer auszumachen, daher sollten Veränderungen zwischen den Erhebungszeitpunkten zusätzlich unabhängig von den Beurteilungen der Schüler betrachtet werden.

Bei einer reduzierten Betrachtung der Veränderungen in den fünf Dimensionen, wie sie Abbildung 6 er-

möglicht, wird deutlich, dass sich die stärksten Verbesserungen in der Kategorie „Philosophische Positionen“ ergeben haben: Im Mittel fällt die Bewertung zum zweiten Erhebungszeitpunkt mehr als einen Skalenpunkt höher aus. Auch die Veränderung der Rangfolge ist bemerkenswert: Die Kompetenz, philosophische Positionen bei der Urteilsbildung zu berücksichtigen, nimmt zum zweiten Erhebungszeitpunkt in Relation zu den anderen Kompetenzen eine deutlich höhere Stellung ein.

Mittlere Verbesserungen betreffen die Kategorien „Begründung/Abwägung“ und „Empirischer Hintergrund“ (im Mittel annähernd ein Skalenpunkt); hier wurden anscheinend relativ starke Ausgangskompetenzen im Laufe der Unterrichtsreihe ausgebaut. Eher geringe Verbesserungen zeichnen sich in den Kategorien „Phantasie/Kreativität“ und „Begriffliche Klarheit“ ab (im Mittel ein halber Skalenpunkt).

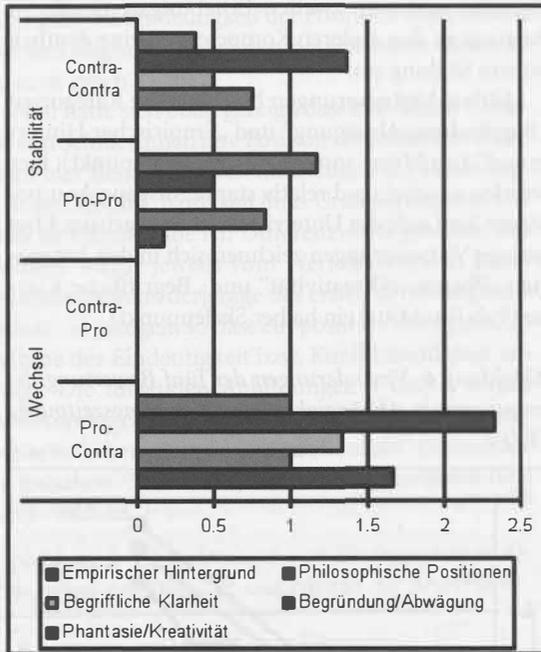
Abbildung 6: Veränderungen der fünf Bewertungsdimensionen in Abhängigkeit vom Erhebungszeitpunkt (EZP)



Ein interessanter Befund ergibt sich, wenn man auch hier die Veränderungen in den fünf Dimensionen wieder danach aufspaltet, ob die Schülerinnen und Schüler bei ihrer ursprünglichen Beurteilung geblieben sind oder nicht. Abbildung 7 zeigt entsprechend die

Differenzwerte der fünf Bewertungsdimensionen in Abhängigkeit von Stabilität und Wechsel der Beurteilung.

Abbildung 7: Veränderungen der fünf Bewertungsdimensionen in Abhängigkeit von Stabilität und Wechsel der Beurteilung



Die Schülerinnen und Schüler, die bei einer Befürwortung des therapeutischen Klonens geblieben sind, haben demnach aus dem Unterricht vor allem das dort erarbeitete *empirische Wissen* in ihre eigene Urteilsfindung aufgenommen; die Gegner des therapeutischen Klonens verarbeiten stärker die erworbenen *philosophischen Kenntnisse*. Bei beiden Gruppen ist auch, in etwa gleich stark, ein Zuwachs in der Dimension „Begründung/Abwägung“ zu erkennen; der oben erwähnte Befund, dass bei den „beurteilungsstabilen“ Schülern die Eindeutigkeit der Beurteilung abnimmt, ob-

gleich auch der Konflikt zurückgeht, könnte wohl tatsächlich auf eine zunehmende *Differenzierung des Urteils* zurückzuführen sein.

Auffällig ist, dass die Schülerinnen, die einen Beurteilungswechsel vollzogen, in fast allen Bewertungsdimensionen (außer „Empirischer Hintergrund“) die größten Fortschritte zeigen. Besonders bemerkenswert ist hier vor allem der Zuwachs in den Dimensionen „Phantasie/Kreativität“ und „Philosophische Positionen“.

#### 4. Schlussfolgerungen

In der beschriebenen Unterrichtsreihe hat die moralische Beurteilung des therapeutischen Klonens bei den meisten Schülerinnen und Schülern an Eindeutigkeit verloren und *zugleich* an Sicherheit gewonnen. Man kann darin einen Erfolg des gemeinsamen philosophischen Reflexionsprozesses sehen, der zu einem ebenso gefestigten wie maßvollen Urteil führte. Nicht nur sicherer, sondern auch eindeutiger wurde die Beurteilung hingegen bei jenen Schülerinnen, die ihre anfängliche Haltung revidierten. Dass sie dabei wohl auch einen, wenn auch anders verlaufenen, philosophischen Reflexionsprozess vollzogen haben, ist an den Dimensionen abzulesen: Die entsprechenden Schülerinnen lassen hier die deutlichsten Fortschritte erkennen.

Problematisch ist der Befund, dass die fünf an Gefert angelegten Bewertungsdimensionen *untereinander* stark korrelierten. Das könnte auf eine globale (undifferenzierte) Auswertung der Schülerantworten schließen lassen. Es könnte aber auch Anlass zu der Frage geben, inwieweit diese Dimensionen tatsächlich *unterschiedliche* Kompetenzen abbilden. Sie sind aus dem Methodenkanon der gegenwärtigen Philosophie abgeleitet und insofern sachlich wohlbegründet; ihre Trennschärfe müsste sich aber auch empirisch bestätigen lassen, wenn aussagekräftige Evaluationen auf Grundlage dieses Modells möglich sein sollen. Hier Klarheit zu schaffen, wäre Aufgabe weitaus größer angelegter empirischer Studien.