

# **Single Subject Research – ein Evaluierungsverfahren auch für die Soziale Arbeit ?**

## **Bruni Vestner**

Kreisjugendring Nürnberger Land  
Waldluststraße 1  
91207 Lauf

## **Werner Wüstendörfer**

Fachbereich Sozialwesen  
Georg-Simon-Ohm Fachhochschule Nürnberg  
Bahnhofstr. 87  
90402 Nürnberg

## **Abstract**

Die Single Subject Research ist ein pragmatisches empirisches Verfahren um die Wirkungen professionellen Handelns auf Einzelsysteme im Zeitverlauf zu analysieren. Sie hat eine bestechende Attraktivität, da sie relativ einfach und ohne größeren Zeitaufwand durchführbar ist. Sie bezieht sich unmittelbar auf die interessierenden Interventionen bei den einzelnen Klienten und versucht die Praxis-Wirksamkeit des professionellen Handelns zu bestimmen. Sie kann somit dazu beitragen, das Handeln der Akteure in der Sozialen Arbeit zu professionalisieren.

Ernstzunehmenden Kritikpunkte sind jedoch zu beachten: Wie bei allen quantitativ orientierten Verfahren ist zu fragen, ob die Operationalisierungen valide und reliabel sind, in welcher Art und Weise dadurch auf die zugrundeliegenden (komplexen) Zielvariablen geschlossen werden kann und ob und inwieweit die Ergebnisse verallgemeinerbar sind.

Insgesamt betrachtet erscheint die SSR als eine gute Möglichkeit, sich mit der Frage auseinanderzusetzen, ob die (pädagogischen) Interventionen bei Klienten etwas bewirken und zielt insbesondere auf den prozessualen Ablauf. Sie könnte dadurch eine gute Ergänzung zu anderen (Selbst-)Evaluierungen bilden, jedoch auch als eigenständiges Verfahren eingesetzt werden.



## Single Subject Research – ein Evaluierungsverfahren auch für die Soziale Arbeit ?

In der amerikanischen Ausbildung der Sozialarbeiter wird seit über 20 Jahren die kontrollierte Einzelfallanalyse (Single Subject Research im Folgenden mit SSR abgekürzt) unterrichtet und in der sozialarbeiterischen Praxis angewandt. In Deutschland wurde das Verfahren bisher weder in der Ausbildung noch in der Praxis Sozialer Arbeit berücksichtigt<sup>1</sup>.

Die quantitativ orientierte Analyse von Einzelfällen kann allerdings auch in Deutschland - in der experimentellen und klinischen Psychologie und den ihr nahe stehenden Disziplinen, z.B. Psychiatrie, auf eine lange Tradition zurückblicken (vgl. Flügel 1925; Huber 1977; Jacobs 1998; Hornung und Hinkel 1998, Petermann 1989; 1996; Bortz, Döring 2002, S. 579ff). In manchen Bereichen, wie z.B. innerhalb des operanten Ansatzes (Sidman 1960) und der Verhaltenstherapie (Lazarus & Davison, 1971; Yates 1970, 1975) nahm sie zeitweise sogar eine dominierende Stellung ein.

Zwar setzte in den 1970er Jahren auch in Deutschland eine verstärkte Diskussion um den Stellenwert der Einzelfallanalyse in den Sozialwissenschaften ein, sie dürfte jedoch gegenwärtig außerhalb von Psychologie, Psychiatrie und Neurologie wenig bekannt sein. Neuere Anwendungen finden sich in verhaltenstherapeutischen Ansätzen in der Heil- und Sonderpädagogik, in der Biermann (2004, S. 22-31) die Hoffnung ausdrückt, mit diesen Verfahren den „naiven Ansatz des Förderns und Hoffens“ zu überwinden und auf wissenschaftlicher Basis wirksame Förderansätze zu identifizieren und deren Effektivität überprüfen“ zu können (Vgl. auch Kern 1997; Julius, Schlosser, Goetze, 2000; Mühl 2002; Lauth & Freese, 2003; Biermann, 2004).

Im Rahmen der Evaluierung von Kompetenzagenturen (vgl. Braun & Schaub 2002; Braun 2003) wurde die SSR beispielhaft in einem Feld der Sozialen Arbeit eingesetzt und auf ihre Brauchbarkeit überprüft (vgl. Vestner, 2004; Wüstendörfer, 2005)

### 1. Was ist Single Subject Research?

Die Bezeichnung *Single Subject Research* bezieht sich auf eine Sammlung empirischer Verfahren, mit denen Veränderungen im Verhalten eines Klienten oder Klienten-Systems beobachtet und innerhalb eines bestimmten Zeitraums wiederholt gemessen werden.

Die SSR wird im deutschen Sprachraum u. a. auch als kontrollierte Einzelfallforschung, kontrollierte Einzelfallanalyse oder N=1 Forschung bezeichnet. In der SSR wird versucht, experimentelle Versuchsanordnungen (Designs, Versuchspläne) bei Einzelfällen nachzubilden, um dadurch die Wirkung von Interventionen zu erfassen. Dementsprechend können die SSR Designs als Experimentierpläne zur kontrollierten individuellen Interventionsplanung bezeichnet werden.

Yegidis und Weinbach (1996, S 233 – 234) beschreiben SSR als ein Forschungsverfahren, bei dem - wie bei anderen Forschungsverfahren - Daten gesammelt, aufgezeichnet, analysiert und dann für die praktische Entscheidungsfindung interpretiert werden. Wie bei der Evaluierung von Programmen geht es darum, die Praxis-Wirksamkeit zu beurteilen. Wie in der Längsschnittforschung wird eine wiederholte Messung derselben Variablen über die Zeit vorgenommen. Wie in experimentellen Designs wird eine unabhängige Variable als 'Kontrollgruppe' eingeführt. Anstelle des prinzipiellen Vergleichs zwischen verschiedenen Gruppen tritt der Vergleich zwischen verschiedenen Bedingungen bei einem Klienten oder Klienten-System über die Zeit.

### 2. Vorgehensweise in der Single Subject Research

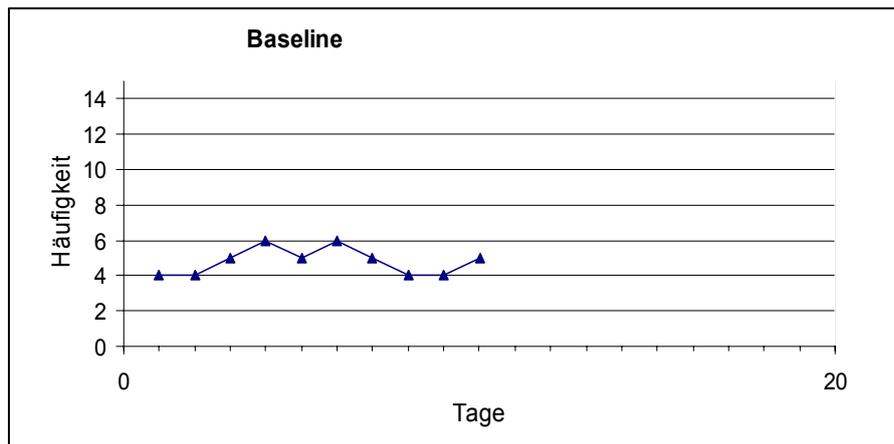
Um die Verhaltens- bzw. Einstellungsunterschiede dokumentieren und nachweisen zu können, muss zunächst das erwünschte Verhalten, die abhängige Variable (AV), genau spezifiziert und gemessen werden. Anschließend erfolgt durch wiederholte Beobachtung (monitoring) eine fortlaufende Messung (repeated measurement) der AV zu verschiedenen Zeitpunkten und unter verschiedenen Bedingungen sowohl in der Nicht-Interventions- als auch während der Interventionsphase. Die Auswertung der für SSR typischen Graphen erfolgt vorwiegend anhand einer visuellen Analyse.

---

<sup>1</sup> Trotz umfangreicher Literaturrecherche konnte in den Datenbanken SOLIS, FORIS, BLISS, IHSLIT, DZI und WAO sowie in der Aufsatzdatenbank JADE - kein Hinweis auf die Single Subject Research bzw. die Durchführung von kontrollierten Einzelfallstudien im Bereich der Sozialen Arbeit in Deutschland bzw. im deutschsprachigen Raum gefunden werden (vgl. Vestner, 2004).

Von wenigen Ausnahmen abgesehen beginnen alle SSR Designs mit der Bestimmung einer Grundrate (engl.: Baseline), die auch als A-Phase bezeichnet wird: sie bildet das „übliche“ Verhalten des Klienten ohne die Einführung oder Präsenz der zuvor spezifizierten Intervention ab.

Während der A-Phase wird zunächst ein bestimmtes Verhalten, das beeinflusst werden soll, über einen festgelegten Zeitraum beobachtet, quantifiziert (mehrmals gemessen) und grafisch dargestellt (vgl. Abbildung 1). Während dieser Zeit erfolgt die zu evaluierende Intervention nicht. Andere Interventionen (the usual treatment) können jedoch initiiert oder fortgesetzt werden (vgl. Yegidis & Weinbach, 1996, S. 235).



**Abbildung 1:** Beispiel einer hypothetischen Baseline mit 10 Messpunkten und den Auftretenshäufigkeiten im beobachteten Verhalten (zwischen 4 und 6).

Damit erfüllt die Baseline zentrale Aufgaben. In erster Linie verfolgt sie eine deskriptive Funktion, indem sie über das Ausmaß des erwünschten bzw. unerwünschten Verhaltens informiert. Sofern die Baseline eine gewisse Kontinuität aufzeigt, können Voraussagen über das zukünftige Verhalten – ohne intervenierenden Eingriff – gemacht werden (voraussagende Funktion). Sie gibt Auskunft darüber, ob mit Einführung der Intervention eine Verhaltensänderung eintritt oder nicht (feststellende Funktion) und anhand ihres Musters können Charakterisierungen hinsichtlich der Vergleichbarkeit von Klienten vorgenommen werden (charakterisierende Funktion). Dies ist ein entscheidendes Kriterium für eine spätere Wiederholbarkeit (Replikation) der Untersuchung unter möglichst ähnlichen Ausgangsbedingungen (vgl. Julius, Schlosser & Goetze, 2000, S. 37 – 44).

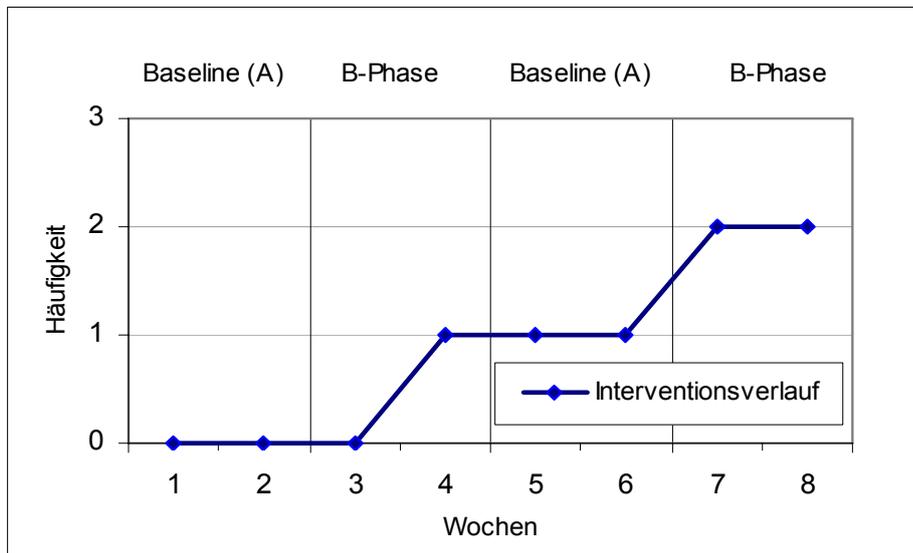
A-Phasen erfüllen im SSR die Funktion von Kontrollgruppen in experimentellen Forschungsdesigns: sie bieten eine Vergleichsquelle indem sie anzeigen, ob und in welchem Ausmaß das angestrebte Verhalten auftritt, wenn die Intervention entweder nicht eingeführt - oder nach einer B-Phase zurückgezogen wurde (vgl. Yegidis & Weinbach, 1996, S. 239). Dem Mangel an einer realen Kontrollgruppe wird durch Designalternativen begegnet, um die Wirkungen von konkurrierenden Hypothesen und Designausrichtungen - in der Interpretation der Beziehung zwischen der Intervention und Veränderungen im angestrebten Verhalten - zu kontrollieren.

Eine an die A-Phase anschließende gleichlange Zeitspanne, in welcher die Intervention präsent ist, wird als Interventions- oder B-Phase bezeichnet. In den B-Phasen wird über eine festgelegte Zeit eine bestimmte Intervention als unabhängige Variable (UV) eingeführt und beobachtet, wie sich das Zielverhalten (engl.: target behavior) entwickelt. B-Phasen können deshalb mit Experimentiergruppen in experimentellen Forschungsverfahren verglichen werden.

## 2. 1 Spezifika der Messungen

Die mit SSR Designs erhobenen Messwerte werden in der Regel graphisch dargestellt, weil dadurch eine einfache, effektive und meist auch überzeugende Auswertung möglich ist. Im Prinzip ist die Aufzeichnung der Messungen unkompliziert: sie können entweder auf Standardmillimeterpapier oder mit Hilfe eines Kalkulationsprogramms, z.B. Excel erfasst werden. Darüber hinaus ist inzwischen auch spezielle Software (z.B. SingWin) zur Aufzeichnung und Auswertung verfügbar (vgl. Bloom, Fischer, Orme, 2002).

Die An- oder Abwesenheit von Interventionen bei der Messung wird auf dem Graphen deutlich gekennzeichnet: die Zeitabstände werden entweder als A- oder als B-Phase markiert. Das Zielverhalten (AV) wird entlang der vertikalen Achse (Y) und die Zeiteinheiten entlang der horizontalen Achse (X) angezeigt. Jeder Punkt auf dem Graphen spiegelt eine Messung des gezeigten Verhaltens und den Zeitpunkt der Messung wider. Die Punkte werden zu einem Liniendiagramm verbunden (vgl. Abb. 2).



**Abbildung 2:** Beispiel eines hypothetischen Interventionsverlaufs mit 8 Messpunkten über einen Zeitraum von 8 Wochen mit einer Auftretenshäufigkeit des gezeigten Verhaltens von 0 – 2.

In der SSR werden die Messungen des Zielverhaltens normalerweise nach Häufigkeit, Dauer oder Zeitabstand variiert (vgl. Yegidis & Weinbach, 1996, S. 236). Die Interventionsabsicht erscheint dabei als wichtigstes Kriterium bei der Entscheidung, wie die AV zu messen ist.

Will man das Auftreten eines bestimmten Verhaltens erhöhen, wird man z.B. bei einem Schulschwänzer die Häufigkeit der Anwesenheit in der Schule messen. Soll dagegen die Zeitspanne, in der ein erwünschtes Verhalten gezeigt wird, erhöht werden, wird man gewöhnlich die Zeitdauer messen, die z.B. ein Auszubildender für die Vorbereitung auf eine Prüfung aufwendet. Soll eine gewisse Kontinuität im erwünschten Verhalten erreicht werden, wird man die Zeitabstände messen, in denen das angestrebte Verhalten nicht gezeigt wird, z.B. die Länge der Zeit zwischen einzelnen Bewerbungsaktivitäten.

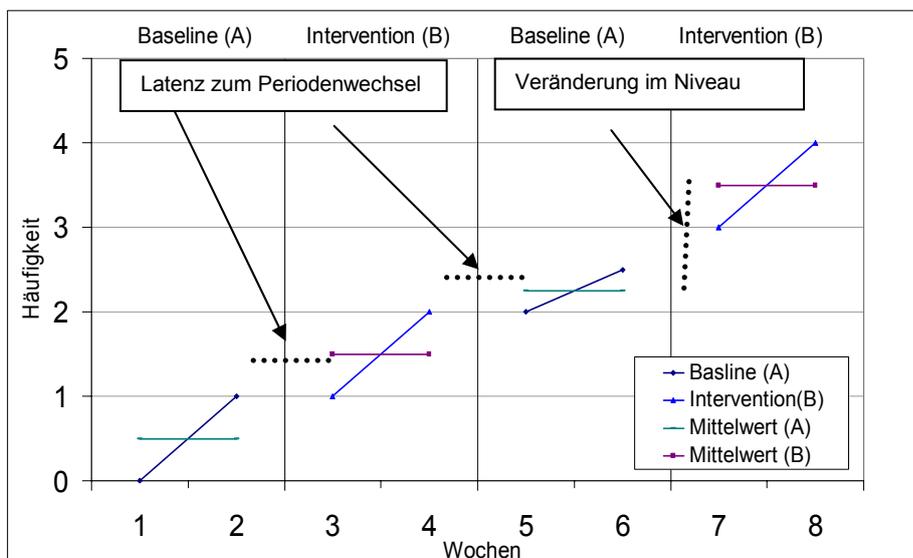
## 2.2 Analyse der Aufzeichnungen

Die Analyse erfolgt anhand von sichtbaren Veränderungen im Niveau des Datenmusters. Besondere Aufmerksamkeit ist dem Übergang von einer Phase zur anderen zu widmen, vor allem in Bezug auf das Vorhandensein oder Fehlen eines Rückgangs oder Anstiegs innerhalb einer Baseline- oder Interventionsphase sowie von Änderungen in der Stabilität des Datenmusters beim Übergang von einer Phase zur nächsten.

Die visuelle Analyse der Graphen erfolgt im Wesentlichen nach den Kriterien Mittelwert, Level, Diskontinuität bzw. Stabilität sowie dem Trend. Der Mittelwert gibt Auskunft über Veränderungen der Durchschnittsrate eines Verhaltens. Der Unterschied im *Niveau (Level)* zwischen der Baseline- und der Interventions-Phase gibt Aufschluss über die Diskontinuität eines Verhaltens. Stabilität im Datenmuster einer vorausgehenden Phase bestimmt die Voraussagbarkeit eines Verhaltens zu einer nachfolgenden Phase. Bei Stabilität des Verhaltens zwischen Ende und Beginn einer Phase wird von einer *Latenz zum* Periodenwechsel gesprochen. Der Trend zeigt die Richtung der Daten, also die systematische Zu- oder Abnahme des Verhaltens über die Zeit. Von einem Slope spricht man, wenn sich innerhalb einer Phase Zu- oder Abnahme im Datenmuster zeigt. Um von einem Trend über zwei oder mehr Phasen - einem Drift (Antrieb) - sprechen zu können, muss der Slope jeder

einzelnen Phase allerdings in die gleiche Richtung, entweder auf- oder absteigend verlaufen (vgl. Bloom, Fischer, Orme, 2002, S. 555 ff.)

Zur Verdeutlichung zeigt der folgende Graph einen exemplarischen Interventionsverlauf über einen Zeitraum von 8 Wochen, mit jeweils einer Messung pro Woche. In allen vier Phasen zeichnet sich ein Slope ab, der jeweils in die gleiche Richtung (Aufwärtstrend) zeigt und somit einen *Drift* darstellt. Dabei spielt es keine Rolle, dass an den Übergängen von der ersten A- zur B-Phase und von der B- zur zweiten A-Phase nur eine latente Veränderung im Verhalten zu beobachten ist. Der Unterschied im Niveau von der zweiten A-Phase zur zweiten B-Phase basiert hier auf dem Vergleich der Mittelwerte, könnte aber genauso auf dem Vergleich einzelner Datenpunkte verschiedener Phasen bestehen.



**Abbildung 3:** Beispiel eines hypothetischen Interventionsverlaufs, der die Veränderungen im Niveau und Trend des Zielverhaltens visualisiert

Darüber hinaus kann der Unterschied zwischen Nicht-Interventions-Phasen und Interventionsphasen, der für den Nachweis der Praxis-Wirksamkeit von Interventionen entscheidend ist, auf einfache Weise auch rechnerisch ausgedrückt werden. Der Prozentsatz sich nicht überlappender Daten (PND) liegt im Idealfall bei 100, wenn sich *alle* Messpunkte der Interventionsphase – je nach angestrebter Richtung der Verhaltensänderung – ober- oder unterhalb der Messpunkte der vorausgehenden Baseline-Phase befinden. Je nach der Problemlage des Einzelfalles können auch niedrigere PND-Werte bedeutende Fortschritte bedeuten. In dem in Abbildung 3 dargestellten hypothetischen Interventionsverlauf liegen die zwei Werte der ersten Baseline bei „0“ und „1“, die zwei Werte der ersten Interventionsphase bei „1“ und „2“. Dies ergibt einen Prozentsatz nicht überlappender Daten (PND) von 50 (1 von 2 Datenpunkten). Die zweite Baseline-Phase hat die Werte „2“ und „2,5“, die zweite Interventions-Phase die Werte „3“ und „4“. Der PND liegt bei 100 (2 von 2 Datenpunkten) (vgl. Julius, Schlosser & Goetze, 2000, S. 137 ff).

Darüber hinaus können auch deskriptive Statistiken und Signifikanztests, z.B. t-Tests und Varianzanalysen gerechnet werden.

## 2. 3Design-Varianten von Single Subject Research

Ein SSR-Design besteht aus einer oder mehreren - zeitlich vorbestimmten - Baselines (A-Phasen), die mit einer oder mehreren Interventionsphasen (B-Phasen) systematisch variieren. Unterschiedliche Zusammenstellungen und Abfolgen von A- und B-Phasen (unterschiedliche SSR-Designs) haben verschiedene Merkmale und Verwendungsmöglichkeiten.

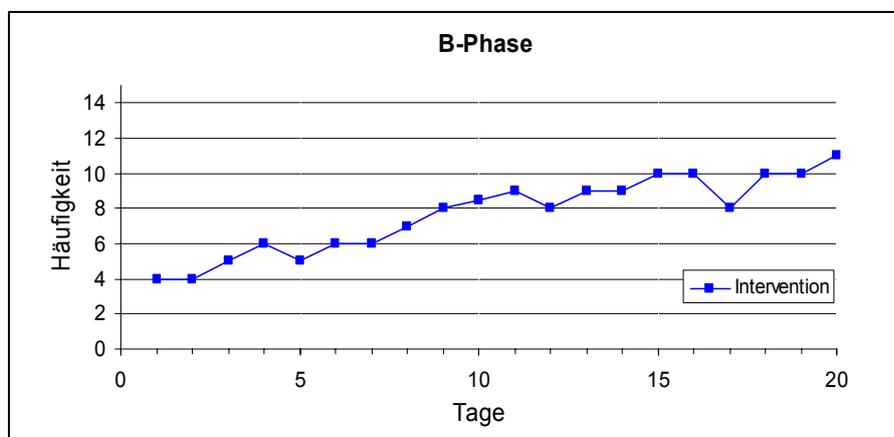
### 2.3.1 Basale Single Subject Research Designs

#### B-Design

Die grundlegendste und meist angewandte Methode der Evaluation Sozialer Arbeit könnte als B-Design bezeichnet werden. Es besteht aus der Einführung der Intervention mit anschließender Beobachtung des Verhaltens. Wird die Beobachtung jedoch als ein Teil von SSR durchgeführt, finden – im Gegensatz zur üblichen Praxis - kontrollierte Messungen im Sinne einer Quantifizierung des Zielverhaltens statt, die aufgezeichnet werden und die das Verhalten des Klienten während der Intervention widerspiegeln.

Der Graph in Abb. 4 demonstriert die Grundzüge einer B-Phase und zeigt, dass das gemessene Verhalten systematisch über die erhobenen 20 Messpunkte in der gewünschten Richtung zunimmt. Dennoch lässt sich nicht sagen, dass die *Intervention* zu der gewünschten Verhaltensänderung geführt hat, denn es fehlt die Erhebung der Baseline und damit das für SSR zentrale Kriterium der experimentellen Kontrolle, mit dem ein Kausalzusammenhang zwischen Intervention und Verhaltensänderung hergestellt werden kann.

Die Durchführung einer isolierten B-Phase macht allerdings dann Sinn, wenn – wie in der Prozessforschung - die Veränderungsprozesse im Interventionsverlauf untersucht und dokumentiert werden sollen. Mit einem B-Design kann z.B. untersucht werden, ob spezifische Interventionen in Zusammenhang mit systematischen Änderungstrends stehen (vgl. Julius, Schlosser & Goetze, 2000, S. 46, 47).

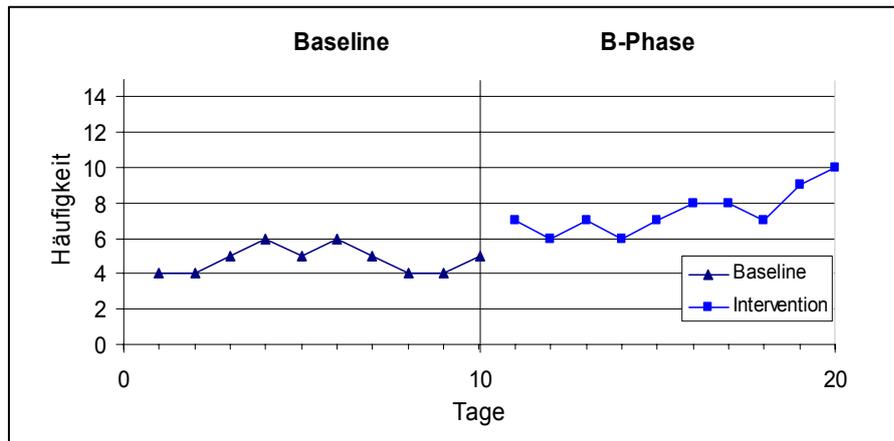


**Abbildung 4:** Hypothetisches Beispiel für einen B-Design mit 20 Messpunkten und den Auftretenshäufigkeiten im beobachteten Verhalten zwischen 4 und 11.

#### AB-Design

Ein AB-Design besteht aus zwei Phasen, einer Nicht-Interventions-Phase (A-Phase) in der die Baseline bestimmt wird, gefolgt von einer Interventionsphase (B-Phase).

Obwohl es sich auch beim AB-Design nicht um einen sog. experimentellen Design handelt, erlaubt er doch – wie der Graph in Abb. 5 verdeutlicht - den logischen Vergleich einer Interventions- mit einer Nicht-Interventionsphase, als minimales Arrangement von Bedingungen, für eine mehr oder weniger objektive Beurteilung der Situation des Klienten vor, während oder nach einer Intervention (vgl. Bloom, Fischer, Orme, 2002, S. 381).



**Abbildung 5:** Beispiel für einen hypothetischen Interventionsverlauf an einem A-B Design mit 20 Messpunkten und den Auftretenshäufigkeiten im beobachteten Verhalten während der Baseline zwischen 4 und 6 und während der Interventionsphase zwischen 6 und 10.

Mit AB-Designs können nur vorläufige Schlussfolgerungen hinsichtlich des Ursache-Wirkungszusammenhangs von Interventionen getroffen werden.

Obwohl die interne Validität eines AB-Designs als sehr gering einzuschätzen ist, könnte es als „hypothesengenerierendes“ Verfahren eingesetzt werden, das in praktischen Kontexten Sozialer Arbeit immer dann angezeigt wäre, wenn zusätzliche Kontrollbedingungen aufgrund mangelnder Praktikabilität bzw. ethischer Bedenken nicht eingeführt werden können (vgl. Julius, Schlosser & Goetze, 2000, S. 51).

### 2.3.2 Experimentelle Single Subject Research Designs

Will man tatsächlich die Wirkung von Interventionen demonstrieren, muss der AB-Design mindestens um eine weitere experimentelle Phase erweitert werden, wobei zu den Bedingungen bei der Bestimmung der Baseline zurückgekehrt werden muss.

Den experimentellen SSR Designs liegen im Wesentlichen zwei grundlegende Schemata zugrunde (vgl. Horning & Hinkel, 1998): das sind zum einen der Rücknahme-Design (withdrawal), zu dem auch der Umkehrungs-Design (reversal) gerechnet wird und zum anderen Designs mit mehreren Baselines (multiple baseline).

#### Die Rücknahme- und Umkehr-Designs

Bei den Rücknahme- bzw. Umkehr-Designs wechseln sich Nicht-Interventionsphasen (A-Phasen) mit Interventionsphasen (B-Phasen) in unterschiedlichen systematischen Anordnungen ab. Dabei wird beim Übergang von einer Phase zur nächsten immer nur eine Variable verändert (changing *one* variable at a time).

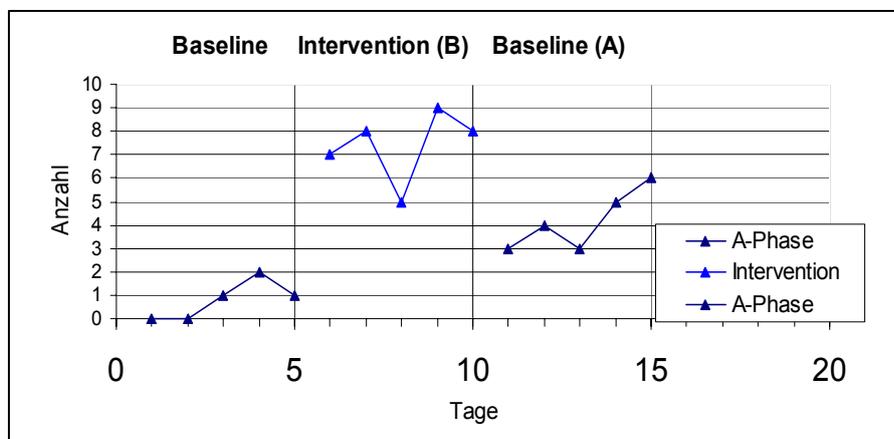
Rücknahme oder Abzug (withdrawal) bedeutet das Stoppen der Intervention, während weiterhin gemessen wird und die Messungen des Zielverhaltens aufgezeichnet werden. Diese Methode wird üblicherweise in zweiten und nachfolgenden A-Phasen im SSR verwendet.

Umkehrung (reversal) dagegen ist der bewusste Versuch, mit einer anderen Intervention, die Richtung der Veränderung ins Gegenteil umzukehren, die während der B-Phase durch die vorherige Intervention auftrat.

## ABA-Design

Der ABA-Design besteht aus drei Phasen. Einer Baseline-Phase folgt eine Interventionsphase mit einer anschließenden zweiten Baseline-Phase. Die zweite A-Phase des ABA-Designs sowie anderer Designs, die mehr als „eine“ A-Phase anwenden, kann entweder durch die Rücknahme (withdrawal) oder durch Umkehrung (reversal) der Intervention charakterisiert werden. Damit erfüllt ein ABA-Design die Mindestanforderungen eines experimentellen Designs.

Wenn sich nach Rücknahme der Intervention das Zielverhalten wieder der Ausgangssituation der Baseline nähert, lässt sich relativ sicher sagen, dass die Intervention für die Veränderung im Zielverhalten verantwortlich zeichnet. Im Umkehrschluss bedeutet dies: wenn sich vorhergesagte Veränderungen in der B- und zweiten A-Phase zeigen, ist es sehr unwahrscheinlich, dass sich diese Änderungen auch ohne die eingeführte Intervention ergeben hätten (vgl. Julius, Schlosser & Goetze, 2000, S. 57). Der Graph in Abbildung 6 zeigt eine deutliche Zunahme des beobachteten Verhaltens während der B-Phase, während sich das gezeigte Verhalten nach Abzug der Intervention in der zweiten A-Phase wieder der Ausgangssituation annähert. Entsprechend obiger Ausführungen kann man bei einem derartigen Kurvenverlauf annehmen, dass die Intervention wirksam war.

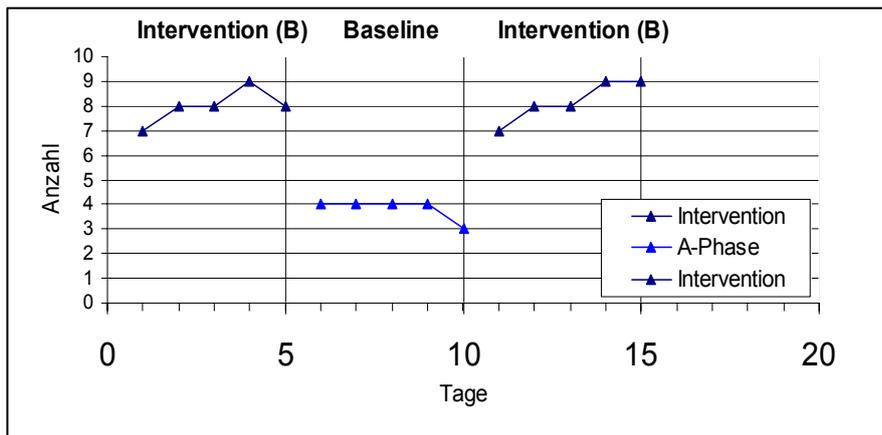


**Abbildung 6:** Beispiel eines hypothetischen Interventionsverlaufs an einem A-B-A Design mit 15 Messpunkten und den Auftretenshäufigkeiten im beobachteten Verhalten während der Baseline zwischen 0 und 6 und während der Interventionsphase zwischen 5 und 9.

## BAB-Design

Ein BAB-Design eignet sich dazu, die Wirkung einer bereits laufenden Intervention zu evaluieren. Wie ein ABA-Design besteht er aus drei aufeinander folgenden Phasen – nur in umgekehrter Abfolge.

Als nachteilig wird bei dieser Versuchsanordnung der Umstand betrachtet, dass der Sachverhalt, der zur Intervention geführt hat, in die Untersuchung nicht einbezogen ist. Will man allerdings wissen, inwieweit eine bereits etablierte Intervention wirkt, ist ein BAB-Design nach Julius, Schlosser & Goetze (2000, S. 65) durchaus sinnvoll. Der Graph in Abbildung 7 zeigt ein BAB-Design mit einem hypothetischen Kurvenverlauf, aufgrund dessen eindeutiger Ausprägung davon auszugehen wäre, dass die Intervention „gewirkt“ hat.



**Abbildung 7:** Beispiel für ein hypothetisches B-A-B-Design mit 15 Messpunkten und den Auftretenshäufigkeiten im beobachteten Verhalten zwischen 7 und 9 in der Interventions-Phase und zwischen 3 und 4 in der dazwischen liegenden A-Phase.

## ABAB-Design

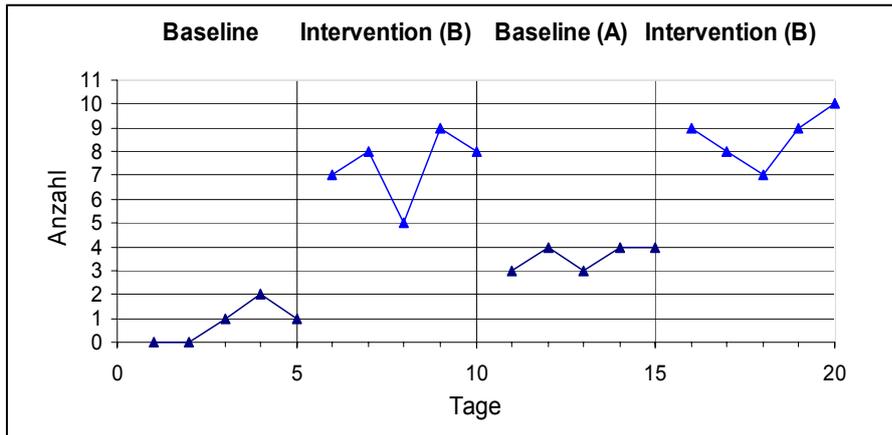
Ein ABAB-Design beginnt mit einer Baseline-Phase (A), wechselt in regelmäßigen Zeitabständen mit einer Interventionsphase (B), einer zweiten A-Phase und endet schließlich mit einer Interventionsphase. Im Gegensatz zu Julius, Schlosser & Goetze (a.a.O) stellt nach Yegidis & Weinbach (1996) erst diese Versuchsanordnung die grundlegendste Art der Durchführung eines Umkehrungs- oder Rücknahme-Designs dar.

Mit der ersten B-Phase wird das Ziel verfolgt, das Verhalten unter der Intervention genau zu beschreiben und eine Vorhersage über die weitere Entwicklung des Verhaltens zu treffen. Die zweite A-Phase hat die Funktion, eine Prognose über den künftigen Verlauf des Verhaltens ohne weitere Intervention zu treffen. Außerdem wird überprüft, ob sich der tatsächliche Verlauf von dem in der vorhergehenden B-Phase prognostizierten Verlauf deutlich unterscheidet und ob Unterschiede im Vergleich zur ersten A-Phase aufgetreten sind.

Erst durch die zweifache Wiederholung jeder Bedingung (ABAB) können erste kausale Aussagen über die Wirksamkeit der Intervention getroffen werden. Der Graph in Abbildung 8 zeigt den hypothetischen Interventionsverlauf bei einem ABAB-Design.

Falls die Intervention wirksam ist, wird angenommen, dass sich der Kurvenverlauf relativ schnell dem der ersten Grundlinie annähert. In der zweiten Interventionsphase müsste sich der Kurvenverlauf – bei einer entsprechend wirksamen Intervention - dann wiederum der Richtung des Kurvenverlaufs der ersten B-Phase annähern. Ist dies der Fall, handelt es sich mit hoher Wahrscheinlichkeit um eine kausale Beziehung.

Damit werden aus experimenteller Sicht die Schwächen der vorher besprochenen Designs überwunden und eine interne Validität hergestellt – was allerdings in der Praxis mit einem relativ hohen Aufwand verbunden ist (vgl. Julius, Schlosser & Goetze, 2000, S. 66 - 72).



**Abbildung 8:** Beispiel für einen hypothetischen A-B-A-B-Design mit 20 Messpunkten und den Auftretenshäufigkeiten im beobachteten Verhalten zwischen 0 und 10.

### Multiple-Treatment-Design (MTD)

Ein Multiple-Treatment-Design (MTD) stellt eine Erweiterung des ABAB-Designs dar.

Anstelle von nur einer Intervention werden zwei oder mehrere Interventionen oder auch Kombinationen mehrerer Interventionen eingeführt. Mögliche Designabfolgen wären z.B. A-B-A-B-C-B-C, A-B-A-B-BC-B-BC. Dadurch kann die Effektivität mehrerer Interventionen miteinander verglichen werden. Zu bedenken ist hier nach Julius, Schlosser & Goetze (2000, S. 74), dass durch die längere Versuchsdauer Reifungsvorgänge, Carry-over- oder Sequenzeffekte eintreten können.

### Multiple Baseline Design (MBD)

Ein multipler Baseline-Design (MBD) ist eine Kombination einzelner AB-Designs, die zeitlich zueinander versetzt werden. Dadurch unterscheiden sich die einzelnen AB-Designs jeweils in der Länge der A- und B-Phase. Ein MB-Design kann sich dabei über unterschiedliche Settings, Personen, Verhaltensweisen etc. erstrecken. Es gibt eine UV (die Interventionsmethode) und mehr als eine AV (z.B. Setting, Person, Verhaltensweisen). Die Intervention wird zuerst auf das Verhalten (bzw. Setting, Person) angewandt, welches die größte Stabilität während der anfänglichen Messung der Baseline widerspiegelt.

Für ein erstes Verhalten (erstes Setting bzw. erste Person) wird eine Baseline bestimmt, auf die anschließend eine Intervention erfolgt. Zeitgleich wird bei einer zweiten AV ebenfalls eine Baseline bestimmt, die jedoch länger andauert, als bei der ersten AV. Die gleiche UV wie im ersten Fall wird anschließend zeitversetzt auf die zweite AV angewandt usw. (vgl. Abb. 9). Je mehr AVs in die Untersuchung einbezogen werden, desto länger dauern die A-Phasen, d.h., desto länger muss abgewartet werden, bis für die weiteren Verhalten, Settings oder Personen die Intervention eingeführt werden kann.

Kausale Wirkungszusammenhänge können nur dann angenommen werden, wenn sich in allen B-Phasen Veränderungen zeigen.

Im Unterschied zum ABAB-Design können mit Hilfe eines MBD-Designs Ursache-Wirkungszusammenhänge wiederholt über *mehrere* Verhalten, Settings, Personen etc. erfolgen. Trotzdem eignet sich der ABAB-Design – wegen der wiederholten AB-Phasen an nur einem Verhalten zum Nachweis kausaler Wirkungszusammenhänge besser. Sofern allerdings eine Umkehr oder Ausblendung des einmal gelernten Verhaltens nicht möglich oder nicht gewünscht ist, bietet ein MBD-Design eine gute Alternative.

Abbildung 9 zeigt den hypothetischen Ablauf eines MBD-Designs über drei verschiedene Beobachtungssituationen, jeweils mit der gleichen Intervention (vgl. Julius, Schlosser & Goetze (2000, S. 77 – 87).

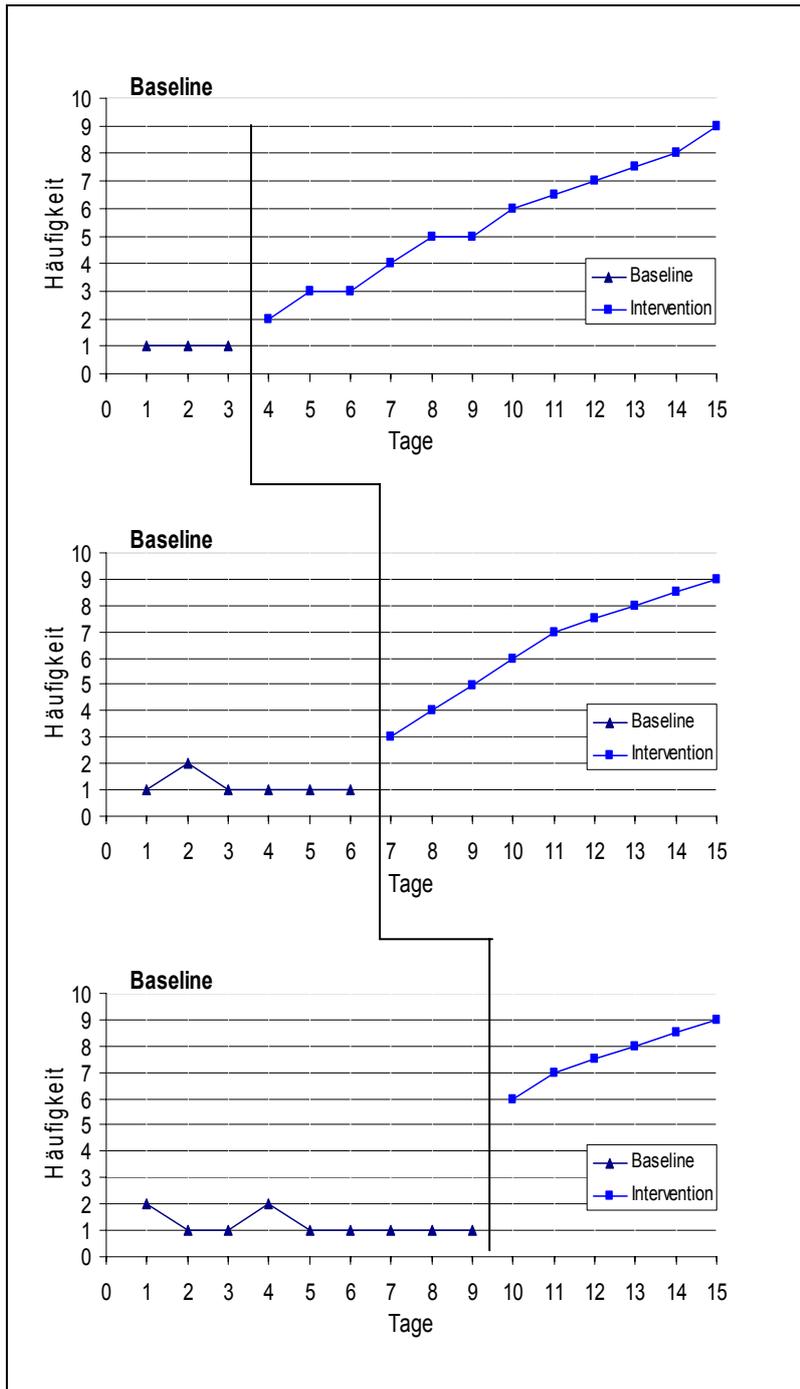


Abbildung 9: Geringfügig veränderter beispielhafter Aufbau eines MBD-Versuchsplans; Quelle: Julius, Schlosser & Goetze (2000, S. 78)

### Changing Criterion Design (CCD)

Nach der Logik des Changing Criterion Designs (CCD) sind Rückschlüsse auf kausale Wirkungszusammenhänge zwischen Intervention und Verhalten dann möglich, wenn sich das Zielverhalten sukzessive an die - zuvor mit dem Klienten – vereinbarten Lernziele annähert. Es trägt damit dem Umstand Rechnung, dass viele erwünschte Verhaltensziele nur schrittweise zu erreichen sind. Bei dem CCD steht die graduelle Veränderung von Verhaltensweisen und deren Auf- oder Abbau in eine erwünschte Richtung im Mittelpunkt.

Bei diesem Design handelt es sich im Prinzip um einen AB-Design mit dem Unterschied, dass die B-Phase eine längere Dauer aufweist und in mehrere Teilphasen gegliedert ist (B1, B2, B3 ...). Jede Teilphase stellt ein bestimmtes Niveau eines vereinbarten Lernzieles dar, das jeweils erreicht sein muss, bevor das nächste Niveau erklommen werden kann (vgl. Abb. 10). Ausschlaggebend für die Bestimmung des ersten zu erreichenden Kriteriums ist das gezeigte Verhalten in der A-Phase. Sofern die erste Teilphase (B1) mit einem über dem Mittelwert der Baseline liegenden Kriterium erfolgreich bewältigt wird, dient dieser Datenverlauf als neuer Standard für das nächste Kriterium, wobei ein Ausblenden der Intervention nicht zwingend erforderlich ist. (vgl. Julius, Schlosser & Goetze, 2000, S. 89 – 96).

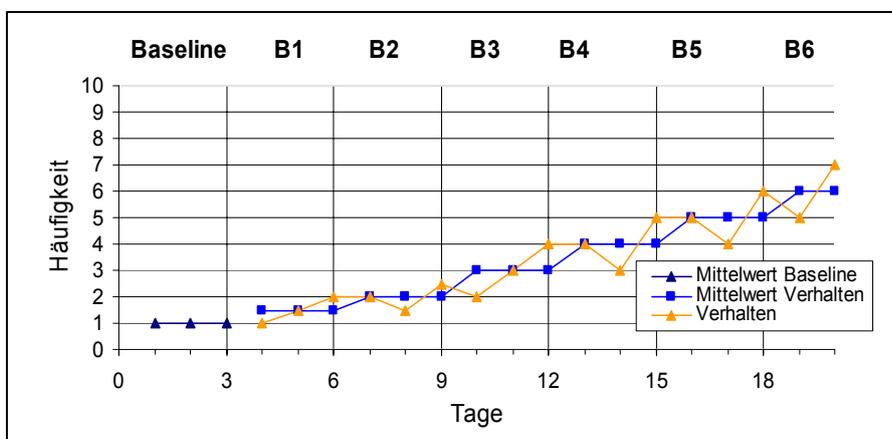


Abbildung 10: Beispiel für einen CCD-Design mit einem hypothetischen Interventionsverlauf

### Alternating Treatment Design (ATD)

Beim ATD-Design geht es – nach vorheriger Bestimmung einer Baseline – darum, alternierende Interventionen im schnellen systematischen Wechsel durchzuführen, um die Überprüfung der relativen Wirksamkeit von gleichzeitig oder kurz nacheinander eingeführten Interventionen herauszufinden und zu überprüfen, welche Intervention effektiver war. Der ATD-Design ist gegen Störfaktoren relativ unempfindlich: für alle möglichen Eventualitäten wurden Designvarianten entwickelt, um diese auszuschließen. Um beispielsweise die Sequenzefekte zu neutralisieren wird die Reihenfolge der Interventionen nach dem Zufallsprinzip ausgewählt (randomisiert) oder ausbalanciert. Mit der ausbalancierten Versuchsanordnung wird versucht, den möglichen Einfluss von Störfaktoren *gleichmäßig* auf die verschiedenen Interventionen zu verteilen.

Zur Verdeutlichung zeigt die nachfolgende Tabelle eine ausbalancierte Interventionsanordnung (vgl. Julius, Schlosser & Goetze, 2000, S. 97 – 114).

Tag Zeit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
T1	C	B	B	C	B	B	C	B	B	C	B	B	C	B	B	C	B	B	C	B
T2	B	C	C	B	C	C	B	C	C	B	C	C	B	C	C	B	C	C	B	C

Tabelle 1: Beispiel einer ausbalancierten Interventionsanordnung eines ATD mit 20 Messzeitpunkten

Eine Sonderform des ATD stellt der Adapted Alternating Treatments Design (AATD) dar. Im Unterschied zu einem Standard-ATD wird jeder Intervention ein spezifischer Aufgabensatz zugeordnet. Dabei wird die Strategie als überlegen angesehen, mit der das Lernkriterium als erstes erreicht wird (vgl. Julius, Schlosser & Goetze, 2000, 114). Der Graph in Abbildung 11 zeigt einen beispielhaften Interventionsverlauf an einem ATD-Design.

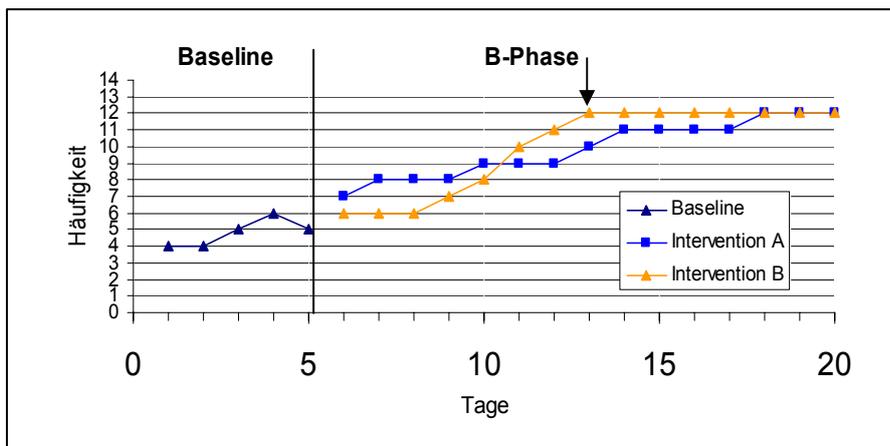


Abbildung 11: Beispiel eines hypothetischen Interventionsverlaufs an einem ATD-Design mit zwei zeitgleich einsetzenden Interventionen. Der Pfeil markiert das festgelegte Zielkriterium, das durch die Intervention B schneller erreicht wurde.

### 2.4 Replikation und Generalisierung

Ein besonderes Thema in der Single Subject Research ist die Frage nach der Generalisierung von Untersuchungsergebnissen, die nur bei einem einzelnen Klienten erhoben wurden (vgl. Hornung & Hinkel, 1998).

Durch Replikation lässt sich zwar aufzeigen, dass SSR dazu geeignet sein könnte, allgemein gültiges Wissen oder Ergebnisse mit externer Validität zu schaffen. Das gelingt allerdings nur, wenn eine Reihe von Studien dieselben abhängigen und unabhängigen Variablen und dieselben Designs - auf Teilnehmer mit demselben Verhalten oder Problemen - anwenden, um über den Einzelfall hinausgehende Erkenntnisse zu erhalten.

Der Mangel an externer Gültigkeit der meisten Ergebnisse ist die größte Beschränkung der Single Subject Forschung. Ohne sorgfältige Replikation kann SSR keinerlei Hinweise darauf liefern, mit welchen Personengruppen oder -kategorien eine Intervention wirksam sein kann oder nicht (Yegidis & Weinbach, 1996, S. 253).

Hingegen kann ein Kausalzusammenhang auf der Datenbasis eines Falles - und damit interne Validität - bereits mit einem A-B-A-B-Design hergestellt werden. In dem Maße, in dem der Interventionsprozess durch die Anordnung spezifischer Designs oder Anzahl von Studienphasen zunehmend strukturiert wird, erhöht sich die interne Validität.

### **3. Zur Beurteilung der Single Subject Research**

Die SSR bietet wie jedes andere Verfahren der (Selbst)Evaluierung eine Reihe von positiven Effekten, die v. Spiegel (1993,1998) mit Kontrolle, Aufklärung, Qualifizierung und Innovation beschrieben hat. Diese Auswirkungen werden im Folgenden im Hinblick auf die SSR zusammen mit deren spezifischen Stärken diskutiert. Kritikpunkte werden anschließend thematisiert. Eine Gesamteinschätzung schließt diesen Artikel ab.

#### **3.1 Möglichkeiten und Stärken**

##### **Zielsetzung, Mittelauswahl und Klärungsprozesse**

Fachkräfte in der Sozialen Arbeit werden häufig mit dem Vorwurf mangelnder Klarheit in der Zielformulierung konfrontiert. Übersehen wird hierbei, dass neben der Komplexität des Gegenstandsbereichs Sozialer Arbeit immer auch die Besonderheiten des Einzelfalles beachtet werden müssen, wobei vor allem auch der Prozesscharakter des Interventionsverlaufs eine wichtige Rolle spielt (vgl. Heiner, 1988, S. 13).

In der SSR konzentrieren sich die Fachkräfte besonders auf den Prozessverlauf mit Fokus auf Intervention und Zielvereinbarung.

Zwar löst die Methodik von SSR die Komplexität des Gegenstandsbereichs Sozialer Arbeit nicht auf, aber es macht den Prozesscharakter des Interventionsverlaufs deutlich sichtbar. Als Teil des Interventionsprozesses unterstützt SSR die schrittweise Problemlösung. Es begleitet den Problemlöser durch Versuch und Irrtum – nur systematischer als in der „üblichen“ Praxis.

Wie alle Evaluationsverfahren, fördert die SSR grundsätzlich das methodisch-systematische Vorgehen. Sie kann den Klärungsprozess in der Auswahl der Interventionen und Festlegung von Zielvereinbarungen auf Seiten der Fachkräfte und der Klientel unterstützen – und beschleunigen. Ihre Methodik – die Bestimmung der Designs – zwingt zu methodischer Vorgehensweise unter der zentralen Fragestellung, was mit einer bestimmten Intervention wie erreicht werden soll. Die Fachkraft und der Klient müssen sich festlegen, woran sie erkennen werden, ob eine Intervention wirkt oder nicht.

Dies lässt sich bereits mit einem einfachen B-Design darstellen, weil sich mit diesem zumindest der Prozesscharakter des Interventionsverlaufs klar abbilden lässt.

Jede, beispielsweise durch einen im Team erarbeiteten Beobachtungsbogen erfasste Verhaltensänderung – ob erwünscht oder unerwünscht – lässt sich damit ohne viele Worte veranschaulichen und aufgrund der systematisch-methodischen Messung des Verhaltens auch empirisch belegen.

SSR erhebt nicht den Anspruch, die Komplexität des Gegenstandsbereichs Sozialer Arbeit als Ganzes abbilden zu wollen. Ihre Qualität liegt vielmehr in der vertieften Betrachtung von fest umrissenen, klar begrenzten Einzelthemen, die einer genaueren methodisch-systematischen Beobachtung unterzogen werden.

Durch die klaren Vorgaben der SSR birgt das Verfahren die Chance, die Handlungsketten durch das Zerlegen in kleinste Einzelschritte beobachtbar und messbar zu machen. So werden nach und nach auch komplexe Strukturen überschaubar und damit steuerbarer.

SSR erhebt auch nicht den Anspruch, die Komplexität schlecht strukturierter Probleme mit denen sich die Akteure Sozialer Arbeit konfrontiert sehen, abbilden zu wollen – oder zu können.

In der Praxis Sozialer Arbeit steht hinter der Einführung einer Intervention immer die Annahme (Hoffnung), dass sich das Verhalten des Klienten in die beabsichtigte Richtung entwickelt und das Interventionsziel erreicht wird.

In der SSR ist das ebenso. Im Unterschied zur „üblichen“ Praxis wird das beobachtete Verhalten genau aufgezeichnet. Die Aufzeichnungen haben nichts mit Aktenführung zu tun, sondern versetzen den Beobachter in die Lage, empirisch nachweisen zu können, welche Wirkung eine Intervention auf das beobachtete Verhalten hat.

##### **Fachlichkeit, Qualifizierung und Motivation**

Das forschungsorientierte Evaluationsverfahren macht nicht nur die Auseinandersetzung mit aktueller Fachliteratur notwendig, es fördert darüber hinaus die fachliche Diskussion mit Kollegen im Team von der Entwicklung der spezifischen Designs über den Interventionsprozess bis zu dessen Auswertung. Die Auswahl bestimmter Methoden sollte dabei durch die Fachkräfte gegenüber Klienten und Kollegen offengelegt sowie im Hinblick auf die Zielsetzung begründet und diskutiert werden.

Die visuelle Analyse der Einzelergebnisse ermöglicht auch Teammitgliedern, die nicht in einen bestimmten Prozessverlauf involviert waren - einen relativ schnellen Überblick – nicht nur über einen einzigen Einzelfall. Unter bestimmten Fragestellungen lassen sich in einem Arbeitsfeld Auffälligkeiten, Abweichungen oder Ähnlichkeiten über mehrere Interventionsverläufe diskutieren und neue Perspektiven in der konkreten Fallarbeit entwickeln (vgl. Vestner, 2004, S. 95).

Bei der Auswertung der Einzelfälle anhand der visuellen Analyse sollte jeder Einzelfall durch die Fachkräfte – trotz Operationalisierung des Zielverhaltens – ganzheitlich betrachtet werden. Die visuelle Analyse schärft dabei den Blick für das Wesentliche, weil Schwankungen im Interventionsverlauf sofort „mit einem Blick“ deutlich und damit einer kritischen Reflexion eher zugänglich werden.

Während der Anwendung von SSR wird die „Kontrolle“, d.h. die Rückbesinnung auf die in der Konzeption formulierten Ziele bzw. anderen externen Erfolgskriterien notwendig. Die SSR fördert dabei besonders die Reflexion auf die Beziehung zwischen Interventionen und dadurch erreichten Veränderungen. Insofern werden Hilfesettings und Interventionsverläufe („Aufklärung“) besonders fokussiert.

Einen hohen Stellenwert dürfte die Anwendung von SSR für die „Qualifizierung“ haben, da besonders fachliche Diskussionen über angemessenere und damit zielgerichtete methodische Vorgehensweisen angeregt werden (kollegiale Beratung). Damit einhergehend werden neue Interventionen diskutiert und evtl. eingeführt.

Durch die SSR können - und das ist ein wesentlicher Unterschied zur *normalen* Praxis - selbst kleine Veränderungen, die im Alltagsgeschehen untergehen, mit einem Blick wahrgenommen werden. Möglicherweise sind diese ersten Hinweise Grund dafür, sich noch kleinere Teilbereiche genauer anzusehen, um sie entweder zum Gegenstand einer Intervention zu machen, oder diesen Aspekt aus dem Problemlösungsprozess ausklammern zu können.

Gerade in schwierigen Fällen können diese – erst durch die visuelle Analyse sichtbar gewordenen, graduellen - Veränderungen einen Motivationsschub leisten und Fachkräfte motivieren, „an dem Fall dranzubleiben“. SSR stärkt damit die Eigenmotivation und beugt der Einkehr von unangemessener Routine in der täglichen Fallarbeit vor.

### **3.2 Grenzen und Schwächen**

#### **Kausalität und Übertragbarkeit**

Wie bei allen Evaluierungen ist der Nachweis kausaler Ursache-Wirkungs-Zusammenhänge nicht ohne weiteres zu führen. Selbst ein differenzierter SSR-Versuchsplan ist nur eine grobe Annäherung an die Erfordernisse eines Experiments, das letztlich über kausale Einflussfaktoren informieren könnte.

Voraussetzung für die Vermutung eines kausalen Zusammenhang ist bei der SSR, dass ein Verhalten im Lauf der Zeit als dauerhaft auftretend beobachtet worden ist (stabil, zunehmend, abnehmend, verschwindend, in wiederholter Weise). Dies hat zur Folge, dass u.U. eine relativ hohe Anzahl von Messpunkten erhoben werden muss. Die angestrebte Veränderung ist dabei genauso zu spezifizieren, wie die Intervention, mit der diese Veränderung erzielt werden soll. Je nach Problemsituation und Zieldefinition – kann dies einen relativ langen Erhebungszeitraum über Monate oder sogar Jahre erforderlich machen. Denn komplexe Verhaltens- oder Einstellungsänderungen passieren nicht einfach über Nacht, sondern werden oft erst im Zeitverlauf und möglicherweise erst lange Zeit nach einer erfolgten Intervention beobachtbar und damit messbar. Dies erfordert den Einsatz komplizierter Designs, um relevante Störvariablen oder Reifungsprozesse als Ursache der Verhaltens- bzw. Einstellungsänderung weitgehendst ausschließen zu können. Dies ist in der Praxis Sozialer Arbeit kaum realisierbar. Hinzu kommen die manchmal überaus schwierig zu interpretierenden Verläufe der beobachteten Variablen.

Die Kritik ist berechtigt. Dennoch könnte von der Vermutung ausgegangen werden, dass je schneller sich das Verhalten eines Klienten nach Einführung einer Intervention in die angestrebte Richtung entwickelt, desto mehr ist dies auf die Intervention zurückzuführen. Dies gilt bereits bei einem basalen B-Design. Letztlich kann jedoch einer möglicherweise verzerrenden Interpretation aufgrund subjektiver und/oder sozialer Erwünschtheit durch die Diskussion der Ergebnisse mit Kollegen und externen Experten entgegengewirkt werden.

Als empirisches hypothesengenerierendes Evaluations-Verfahren betrachtet, könnten mit SSR Designs - durch eine schrittweise Erweiterung der Fallbasis - Vergleichsfälle untersucht, Extremfälle gegenübergestellt oder rein *theoretisch* interessante Fälle einer genaueren Evaluation zugeführt werden, die das Handlungssystem Sozialer Arbeit gleichsam von innen beleuchtet und gestaltet. Ob und inwieweit sich die mit SSR untersuchten Ergebnisse auf andere Einzelfälle übertragen lassen (externe Validität), bleibt dennoch diskussionsbedürftig

und bedarf zusätzlicher Anstrengungen. Julius, Schlosser & Goetze (2000, S. 155 ff.) schlagen hier die Metaanalyse als „ein ganzes Bündel von Maßnahmen“ (vor), um die (...) Validität von zusammenfassenden Aussagen in einem interessierenden Forschungsgebiet zu erhöhen.“

Herkömmliche (Gruppen)Evaluierungen können ebenso die Übertragbarkeit analysieren und so die individuellen Einzelfalldarstellungen ergänzen.

### **Ethische Bedenken**

Ein weiterer Kritikpunkt an der SSR liegt darin, dass die forschungsorientierten Designs den Moralkodex der Fachkräfte Sozialer Arbeit verletzen, weil diese ihre Aufmerksamkeit zu sehr auf das Evaluationsverfahren lenken und sich nicht mehr ausschließlich auf den Hilfeprozess konzentrieren könnten.

Jedwede Praxis Sozialer Arbeit basiert auf der Annahme, dass bestimmte Methoden - in sensibler aber systematischer Weise angewandt – problemlösend und zielführend sind. Evaluationsverfahren wie die SSR versuchen sich genau mit diesen Fragestellungen auseinanderzusetzen. Zugespielt formuliert würde es sogar ethisch bedenklich sein, wenn man nicht systematisch überprüft oder - zumindest nach gängiger fachlicher Auffassung - nicht davon ausgehen kann, dass die Interventionen positive Auswirkungen auf Klienten haben. Wenn Evaluationen als nicht hilfreich, sondern sogar als verletzend oder peinlich von Klienten erlebt werden, dann sollten sie auch nicht angewandt werden. Klienten sollten bei der SSR in die Problem- und Zieldefinition - unter Berücksichtigung individueller Wertvorstellungen und Handlungsmöglichkeiten - auch in die Interpretation und Entscheidung über die weitere Verwendung der erhobenen Daten einbezogen werden. Sie müssen die Evaluation jederzeit unterbrechen oder beenden können, wenn sie sich durch die Evaluation verletzt fühlen oder einen Nachteil befürchten.

Damit kann auch der Vorbehalt von Praktikern entkräftet werden, die Bedenken äußern, wenn sie eine offensichtlich wirksame Intervention stoppen müssen, nur weil der gewählte Design dies zu einem bestimmten Zeitpunkt erfordert. Zudem birgt die Rücknahme einer Intervention auch die Chance, frühzeitig eine mögliche – unbeabsichtigte - Interventions-Abhängigkeit des Klienten zu erkennen und dieser entgegen zu wirken.

Weitere ethische Bedenken werden von Fachkräften genannt, wenn ein SSR –Design eine Baseline oder A-Phase vorsieht, da sie sich aus berufsethischen Gründen verpflichtet sehen, bei Hilfebedarf von Personen auch Hilfe anzubieten.

Selbstverständlich sind diese Bedenken zu beachten, könnten aber durch geeignete Versuchspläne z.B. B-A-B umgangen werden.

### **Künstlichkeit des Evaluationsverfahrens**

Evaluation mit SSR-Designs erfordert das Festlegen bestimmter Verfahrensschritte im Interesse eines zielorientierten Problemlösungsprozesses. Vielen Praktikern erscheinen diese Bedingungen als zu starr, zu künstlich. Sie sehen sich aufgrund der experimentellen Designs zu Handlungen, bzw. Interventionen gezwungen, die es ohne Evaluation mit SSR nicht bräuchte. Die Vorbehalte erstrecken sich dabei auf die Notwendigkeit der Datenerhebung über lange Baselines, das Erfordernis der Operationalisierung und Reduzierung komplexer Verhaltensweisen auf möglichst einfach quantifizierbare Variablen, die den Gesamtzusammenhang nicht mehr adäquat abbilden könnten, etc..

Diese Vorbehalte dürften jedoch nicht zutreffen, denn Sozialarbeiter und Klient müssen übereinstimmend der Meinung sein, dass durch die Evaluation des Prozessverlaufs mit SSR im individuellen Kontext ein unmittelbares Feedback auf problemorientierte Interventionen hilfreich für den weiteren Prozessverlauf sein kann. SSR bietet beispielsweise durch einen B-A-B-Design die Möglichkeit zur Evaluation auch dann, wenn eine Hilfeleistung sofort erforderlich ist und keine Zeit bleibt, um eine Baseline zu erfassen.

Im Mittelpunkt der Überlegungen muss auch in der SSR immer die Ausgewogenheit von Kosten und Nutzen für den Klienten stehen. Denn in der SSR geht es um die Beantwortung der Frage, ob die gewählte Intervention bei einem bestimmten Klienten eine Veränderung bewirkt und wenn ja, ob der Unterschied mit den Interventionszielen übereinstimmt.

Dem Vorbehalt der scheinbaren „Künstlichkeit“ des Evaluationsverfahrens kann jedoch durch inhaltliche Auseinandersetzungen mit der spezifischen Methodik der SSR und/oder durch kollegiale oder externe Beratung entgegengewirkt werden.

### 3.3 Zusammenfassende Einschätzungen

SSR könnte in seiner Begrenzung auf kleine Teilausschnitte für die Soziale Arbeit eine Chance darstellen, die Komplexität des Gegenstandsbereiches in seinen Teilthemen zu erfassen, um sie für den Problemlösungsprozess der Akteure Sozialer Arbeit zu strukturieren und zu systematisieren.

Ihre Stärke liegt in der Beleuchtung der Einzelthemen, indem es die Wirkungen der eingeführten oder abgezogenen Interventionen in den Verhaltensweisen der beobachteten Subjekte in seiner Prozesshaftigkeit begleitend dokumentiert. SSR wendet sich damit an die Praktiker Sozialer Arbeit und will diese durch die Reduktion der Komplexität auf kleine Einheiten in ihrer Entscheidungsfindung unterstützen.

Deshalb könnte überall dort, wo die selbstinitiierte systematische Beobachtung der eigenen Handlungszusammenhänge (oder Teile davon) mit dem Ziel der Feststellung des erreichten oder potentiellen Problemlösungsgrades der untersuchten Handlungselemente als Chance für eine weitere Professionalisierung gesehen wird, der Einsatz der SSR sinnvoll erfolgen. (vgl. Vock, 1999, S 45).

Darüber hinaus ist es grundsätzlich möglich, auch kompliziertere Verhaltenskomplexe mit Hilfe von Skalen und Indizes zu erfassen.

Wenn das Arbeitsfeld – ganz oder in Teilen – mit SSR evaluiert werden soll, erfordert dies eine entsprechende Operationalisierung. Dadurch können zwangsläufig zunächst nur die wesentlichen „Standardsituationen“ in ihrer Breite evaluiert werden.

Allerdings ist hierbei festzuhalten, dass mit zunehmender Komplexität der zu beobachtenden Zusammenhänge bzw. der zu beobachtenden Verhaltenskomplexe, auch der Umfang und der Anspruch an die Erhebungsinstrumente steigen. Entsprechend aufwändig und schwierig gestaltet sich die kontinuierliche Messung des beobachteten Verhaltens und dessen Auswertung anhand von Skalen und Indizes.

Mit zunehmender Komplexität der Operationalisierungen und der Designs wird es zudem immer schwieriger, die daraus resultierenden Ergebnisse angemessen zu interpretieren und sie auf die insgesamt verfolgten Ziele zu beziehen. Eine kollegiale Diskussion z.B. im Arbeitsteam könnte dazu verhelfen, diese Ergebnisse sinnvoll zu interpretieren und sie auf die zu erreichenden Ziele zu beziehen.

Bei der SSR geht es nicht um Grundlagenforschung. Die Auswahl von Problemverhalten im SSR geschieht wegen ihrer sozialen Relevanz – zumindest für das zu beforschende Subjekt - und ist problemorientiert. Es ist dabei immer zu fragen, wie wichtig die anzustrebenden Verhaltensänderungen für das betreffende Individuum, die zugehörige Gruppe und die Gesellschaft sind. (Vgl. Julius, Schlosser & Goetze, 2000, S. 14).

Es wird aber bei der Einschätzung von SSR ebenso zu fragen sein, wie wichtig die verstärkte Evaluationsorientierung Sozialer Arbeit für ihre Akteure, ihre Institutionen und Träger und die Gesellschaft ist. Schließlich ist bei der kritischen Betrachtung von SSR noch zu fragen, welche Möglichkeiten und Alternativen es grundsätzlich gibt, die Wirksamkeit von Interventionen im Zeitverlauf empirisch so zu bestimmen, dass sie in der Praxis der Sozialen Arbeit angewendet werden können.

Praktische Hindernisse für den Einsatz von SSR können jedoch sein, wenn eine unmittelbare Krisenintervention durchzuführen ist, wenn die Arbeitsbelastung von SozialarbeiterInnen zu hoch ist, wenn dieses Verfahren in der eigenen Einrichtung nicht gewürdigt bzw. anerkannt, aber auch, wenn es vom Klienten abgelehnt wird (vgl. Rubin & Babbie, 1997, S. 314).

### Resümee

Die SSR ist ein pragmatisches empirisches Verfahren um die Wirkungen professionellen Handelns auf Einzelsysteme im Zeitverlauf zu analysieren. Sie hat eine bestechende Attraktivität: Sie ist relativ einfach und ohne größeren Zeitaufwand durchführbar und auch die (optische) Auswertung kann ohne Spezialkenntnisse vorgenommen werden. Sie bezieht sich unmittelbar auf die interessierenden Interventionen bei den einzelnen Klienten und versucht die Praxis-Wirksamkeit des professionellen Handelns zu bestimmen. Insofern greift sie unmittelbar die Anliegen der Fachkräfte auf, die Wirksamkeit ihrer Arbeit zu analysieren und in der Konsequenz nach effektiveren Interventionensmethoden zu suchen und anzuwenden (vgl. Wüstendörfer, 2005). Sie kann somit dazu beitragen, das Handeln der Akteure in der Sozialen Arbeit zu professionalisieren.

Ernstzunehmenden Kritikpunkte sind jedoch zu beachten: Wie bei allen quantitativ orientierten Verfahren ist zu fragen, ob die Operationalisierungen valide und reliabel sind, in welcher Art und Weise dadurch auf die zugrundeliegenden (komplexen) Zielvariablen geschlossen werden kann und ob und inwieweit die Ergebnisse

verallgemeinerbar sind. Ein schematisierend-ordnendes Evaluationsverfahren wie SSR muss dabei nicht im Widerspruch zur situativen Offenheit methodischen professionellen Handelns stehen. Professionelles Handeln unterliegt einer gewissen „Künstlichkeit“ in der Anwendung spezifischer Methoden. Diese sind zunächst nicht mehr als ein Grundmuster, welches für unterschiedliche Aufgaben unterschiedlich akzentuiert und konkretisiert wird. Damit löst sich der scheinbare Widerspruch auf (vgl. Thiersch, 2000, S. 24)

Wie jedes andere Evaluierungsverfahren beansprucht die SSR jedoch Arbeitszeit, die dann möglicherweise unter dem zunehmenden Kosten- und Erfolgsdruck in der Sozialen Arbeit in der direkten Arbeit mit Klienten fehlt. Dies gilt jedoch gleichermaßen für Fortbildungsmaßnahmen, Qualitätssicherung u.ä. Deshalb sind zunächst grundsätzliche Fragen zu klären, welchen Stellenwert Evaluierung in der praktischen Arbeit haben soll und wieviel Arbeitszeit dafür bereitgestellt werden muss. In Anlehnung an Rubin & Babbie (1997) stellt sich dabei nicht die Frage, wie hoch der Zeitaufwand für die Durchführung des SSR ist sondern vielmehr, ob wir es uns leisten können auf dieses Verfahren zu verzichten.

Insgesamt betrachtet erscheint die SSR als eine gute Möglichkeit, sich mit der Frage auseinanderzusetzen, ob die (pädagogischen) Interventionen bei Klienten etwas bewirken und zielt insbesondere auf den prozessualen Ablauf. Sie könnte dadurch eine gute Ergänzung zu anderen (Selbst-)Evaluierungen bilden, jedoch auch als eigenständiges Verfahren eingesetzt werden.

## Literatur

Biermann, Adrienne (2004):

Interventionsforschung im Bereich von Körper- und Geistigbehindertenpädagogik am Beispiel der Kontrollierten Einzelfallforschung. In: Sonderpädagogik, 34.Jg., H. 1, S. 22 - 33.

Bloom, Martin; Fischer, Joel; Orme, John G. (2002):

Evaluating Practice: Guidelines for the Accountable Professional, 4. Aufl., Boston u.a.: Allyn & Bacon.

Braun, Frank (2003):

Konzeption der wissenschaftlichen Begleitung des Modellprogramms „Kompetenzagenturen“. Unveröff. Manuskript. München: Deutsches Jugendinstitut

Braun, Frank; Schaub, Günther (2002):

Wissenschaftliche Begleitung des Modellprogramms „Kompetenzagenturen zur sozialen und beruflichen Integration von Jugendlichen 2002 – 2006“. München: DJI. 01.08.2002

Heiner, Maja (1988): Von der forschungsorientierten zur praxisorientierten Selbstevaluation. Entwurf eines Konzeptes. In: Heiner, Maja: Selbstevaluation in der sozialen Arbeit. Freiburg im Breisgau: Lambertus. S. 7 – 40.

Heiner, Maja (2002):

Evaluation. In: Deutscher Verein für öffentliche und private Fürsorge (Hrsg.): Fachlexikon der sozialen Arbeit. 5. aktual. Auflage. Frankfurt/Main: Eigenverlag. S. 301 - 302.

Hornung, Joachim; Hinkel, Jana (1998):

Single-Case Studies. Eine kleine systematische Übersicht.

[http://www.datadiwan.de/netzwerk/index.htm?/evaluation/hr\\_023d\\_.htm](http://www.datadiwan.de/netzwerk/index.htm?/evaluation/hr_023d_.htm), 03.12.1998

Julius, Henri; Schlosser, R.W.; Goetze, Herbert (2000):

Kontrollierte Einzelfallstudien. Göttingen: Hogrefe.

Kern, Horst, J. (1997):

Einzelfallforschung. Eine Einführung für Studierende und Praktiker. Weinheim: Psychologie Verlags Union.

Lauth, Gerhard W.; Freese, Stephan 2003:

Effekt einer schulbasierten Behandlung von ADHD in der Bewertung von Lehrern und Eltern – eine Einzelfallstudie an 4 Kindern. In: Sonderpädagogik, 34. Jg. S.

Mühl, H. (2002): Einzelfallstudien zum pädagogischen Umgang mit Verhaltensstörungen bei Menschen mit geistiger Beeinträchtigung. Oldenburg: Didaktisches Zentrum (diz) der Carl von Ossietzky Universität

Petermann, Franz (1989):

Einzelfallanalyse. 2., völlig überarb. Aufl., München, Wien: Oldenbourg

Rubin, Allen; Babbie, Earl (1997):

Research Methods for Social Work. 3. Aufl., Pacific Grove u. a.: Brooks/Cole Publ. Co.

Spiegel, Hiltrud von (1993):

Aus Erfahrung lernen: Qualifizierung durch Selbstevaluation., Münster: Votum

Spiegel, Hiltrud von 1998:

Arbeitshilfen für das methodische Handeln. In: Heiner Maja; Meinhold, Marianne; Spiegel, Hiltrud von; Staub-Bernasconi, Silvia: Methodisches Handeln in der Sozialen Arbeit. 4. erw. Aufl., Freiburg im Breisgau: Lambertus

Spiegel, Hiltrud von(2001):

Perspektive der Selbstevaluation – veranschaulicht am Beispiel Qualitätsentwicklung in der Kinder- und Jugendhilfe. In: Bundesministerium für Familie, Frauen, Senioren und Jugend (Hrsg.): Materialien zur Qualitätssicherung in der Kinder- und Jugendarbeit „QS 35“. S 85 – 88. Berlin 2001.

Thiersch, H. (2000):

Lebensweltorientierte Soziale Arbeit. Aufgaben der Praxis im sozialen Wandel. 4. Aufl.. Weinheim: Beltz

Vock, Rainer (1999):

Selbstevaluation in Beschäftigungs- und Qualifizierungsprojekten. Teil 1. Theoretische und methodische Grundlagen. In: heidelberger institut beruf und arbeit: hiba-Weiterbildung. Band 20/04. Lübeck: hiba verlag

Vestner, Bruni (2004):

Single Subject Research – ein Forschungsverfahren zur Evaluation der Wirksamkeit von Sozialer Arbeit. Eine Darstellung des Verfahrens und seiner exemplarischen Anwendung im Rahmen des bundesweiten Modellprogramms „Arbeitsweltorientierte Jugendsozialarbeit“ in der Modellphase „Kompetenzagentur“. Unveröff. Diplomarbeit, FB Sozialwesen der G.S.O-FH Nürnberg vom SS 2004.

Wüstendörfer, Werner (2005):

Erprobung der „Single Subject Research“ im Modellprogramm „Kompetenzagenturen“. München/Halle: DJI, Wissenschaftliche Texte 5/2005 ([www.dji.de/abt\\_fsp1](http://www.dji.de/abt_fsp1))

Yegidis, Bonnie L.; Weinbach, Robert.W. (1996):

Research methods for social workers. 2. Aufl., Needham Heights: Allyn and Bacon

Verf.: Dipl. Sozialpäd. Bruni Vestner, Geschäftsführerin des KJR Nürnberg-Land;

Prof. Dr. Werner Wüstendörfer, Georg-Simon-Ohm Fachhochschule Nürnberg, Fachbereich Sozialwesen, Bahnhofstr. 87, 90402 Nürnberg