

## Richter- und gerichtsbezogene Sanktionsdisparitäten in der deutschen Jugendstrafrechtspraxis

von Jörg Hupfeld

### Zusammenfassung

Leitthema des Beitrags sind regionale und richterbezogene Sanktionsdisparitäten in der Jugendstrafrechtspraxis. Die Darstellung beginnt mit einer Erläuterung der Hintergründe und Methoden der Sanktionsforschung. Als Hintergrundfolie für die weiteren Betrachtungen werden sodann Forschungsstrategien und zentrale Befunde zweier Untersuchungen zur Sanktionspraxis im Allgemeinen Strafrecht skizziert (Langer 1994; Oswald 1994). Hierauf aufbauend erfolgt eine Analyse der Grenzen und Möglichkeiten der Sanktionsforschung im Jugendstrafrecht sowie die Darstellung und Diskussion erster empirischer Befunde. Die vorgestellten Analysen beruhen auf 2470 im Erziehungsregister des Bundeszentralregisters (BZR) dokumentierten einzelrichterlichen Entscheidungen bei einfachen Diebstahlsdelikten gegenüber Jugendlichen. Es können systematische regionale und richterbezogene Sanktionsdisparitäten nachgewiesen werden, die zu einem erheblichen Teil auf Unterschiede im Umgang mit Mehrfachauffälligen zurückzuführen sind. Die Ergebnisse der Untersuchung werden im Vergleich zu Befunden zur Sanktionspraxis im Allgemeinen Strafrecht diskutiert. Den Abschluß bilden wiederum eher allgemeine Überlegungen zur spezial- und generalpräventiven Bedeutung von Sanktionsdisparitäten und Schlußfolgerungen für die weitere Forschung.

### Das Interesse an Sanktionsdisparitäten

Untersuchungen zu Ausmaß und Ursachen von Sanktionsdisparitäten in der deutschen Strafrechtspraxis haben eine lange Tradition. Die Forschungsbemühungen begannen Anfang dieses Jahrhunderts mit den Arbeiten von Woerner (1907) und Exner (1931) und reichen z. B. über solche von Schiel (1969), Schöch (1973) und Albrecht (1980) bis hin zu denjenigen von Pfeiffer & Savelsberg (1989), Albrecht (1994), Langer (1994) und Oswald (1994), um nur einige wenige zu nennen. Untersucht wurden und werden zum einen regionale (z. B. auf der Ebene von Bundesländern, Landgerichtsbezirken oder ausgewählten Amtsgerichten) Sanktionsdisparitäten sowie Sanktionsunterschiede zwischen einzelnen RichterInnen. Zum anderen widmete man sich wiederholt dem Vergleich der strafrechtlichen Reaktionen auf bestimmte Tätergruppen (z. B. Männer vs. Frauen, Ausländer vs. Deutsche).

Die Mehrzahl der bislang vorliegenden Untersuchungen scheint zumindest die Existenz sowohl regionaler als auch richterbezogener Sanktionsdisparitäten klar zu belegen. Über das genaue quantitative Ausmaß der Sanktionsdisparitäten können zwar bislang kaum eindeutige Aussagen gemacht werden. Insbesondere auf Juristen scheinen die Daten jedoch eine eher beunruhigende Wirkung auszuüben (vgl. Schöch 1989, S. 132). Eine ungleichmäßige Sanktionspraxis läßt zum einen fragen, inwieweit die Strafzumessungslehre überhaupt angeben kann, welche Bestrafung für unterschiedliche Strafzumessungssachverhalte richtig wäre (vgl. Giehring 1989, S. 80) und ob andererseits die Steuerungsinstrumente ausreichen, um eine gleichmäßige Auslegung und Anwendung des Strafrechts sicherzustellen (vgl. Pfeiffer & Savelsberg 1989, S. 17). Ferner stellt sich das Problem, daß systematische Ungleichbehandlungen möglicherweise nicht nur dem Gleichheitsgebot des Grundgesetzes (Artikel 3 Absatz 1) widersprechen, sondern zugleich den vom Strafrecht erhofften spezial- und generalpräventiven Wirkungen abträglich sein könnten (vgl. hierzu jedoch Krauss 1989).

Während systematische Sanktionsdisparitäten für Juristen eher problematisch sind, stellen sie in den Sozialwissenschaften in der Regel eine notwendige Voraussetzung für die empirische Überprüfung der jeweiligen fachspezifischen Theorien dar. Dementsprechend ist hier bereits die Freude über den bloßen Nachweis von Sanktionsdisparitäten meist deutlich größer.

### Methoden der Sanktionsforschung

Häufig werden bei Strafzumessungsuntersuchungen verkürzte oder simulierte Fallinformationen vorgelegt, auf deren Grundlage »fiktive« Strafurteile zu fällen sind. Vorteil dieser Methode ist, daß nachweisbare regionale oder richterbezogene Sanktionsunterschiede nicht auf eventuellen Unterschieden im Fallmaterial basieren. Ein häufig genannter Kritikpunkt betrifft jedoch die externe Validität solcher Simulationsstudien (vgl. Kette 1994, S. 128; Lemon & Bond 1987). Empirische Untersuchungen, in denen sowohl fiktive als auch reale Sanktionsentscheidungen erhoben wurden, belegen immerhin, daß generelle Zweifel an der Übertragbarkeit der Befunde unbegründet sind. Zwar kann die reale Strafentscheidung nicht immer genau vorhergesagt werden. Sofern fiktive und reale Fälle hinsichtlich der zentralen strafzumessungsrelevanten Merkmale vergleichbar sind, finden sich jedoch sehr deutliche Übereinstimmungen in den jeweiligen richterlichen Reaktionstendenzen (vgl. Hood 1972; Kapardis 1985).

De facto bleibt gleichwohl das Problem, daß die externe Validität solch experimenteller Untersuchungen von der Justizpraxis äußerst kritisch eingeschätzt wird (vgl. Oswald 1991) und die Befunde zumeist als »irrelevant« abgetan werden, insbesondere wenn sie auf die Existenz systematischer Sanktionsdisparitäten hindeuten (s. o.). Im Hinblick auf die Akzeptanz haben Forschungsbefunde, die auf realen gerichtlichen Strafurteilen basieren, offensichtlich deutliche Vorteile. Diesen Akzeptanzvorteilen stehen allerdings erhebliche methodische Probleme gegenüber. Sowohl bei Forschungsarbeiten, die auf der Beobachtung in Hauptverhandlungen als auch auf der Analyse der Strafakten beruhen, besteht das bereits erwähnte Problem der Variabilität des Fallmaterials. Selbst wenn man sich allein auf die Kontrolle der wesentlichsten Fallmerkmale beschränkt, müssen in der Regel sehr viele Akten analysiert bzw. Hauptverhandlungen beobachtet werden, und die Grenzen der Durchführbarkeit sind meist sehr schnell erreicht. Für den Bereich der deutschen Strafzumessungspraxis steht mit dem Bundeszentralregister (BZR) zur Zeit nur eine einzige Datenquelle zur Verfügung, die sowohl regionale als auch umfangreichere richterbezogene Strafzumessungsanalysen unter Wahrung der notwendigen methodischen Standards erlaubt (vgl. Oswald 1994, S. 88 ff.). Das BZR in Berlin besteht aus mehreren personenbezogenen Datenbanken (vgl. Uhlig 1989). Im eigentlichen Zentralregister werden alle strafrechtlichen Verurteilungen in einer Freitext-Datenbank erfasst, die gegenüber Erwachsenen aufgrund in Deutschland begangener Delikte ausgesprochen wurden. Die jeweiligen rechtskräftigen Verurteilungen werden unter dem Namen der betreffenden Person als einzelne Belege abgespeichert. Hier sind jeweils alle Paragraphen erfasst, gegen die die Person verstoßen hat, sowie alle Paragraphen, auf die sich das jeweilige Strafurteil begründet. Aus den Daten lassen sich somit neben personenbezogenen Informationen zu Alter, Geschlecht und Nationalität sowie der bisherigen »Legalkarriere« des Täters – allerdings mit Ausnahme der Angaben zu eventuellen Verfahrenseinstellungen wegen Gefährlichkeit und nach Erfüllung von Auflagen (§§ 153 Abs. 2 und 153 a Abs. 2 StPO) – vielfältige Informationen zu dem sanktionierten Delikt und den jeweiligen Rechtsfolgen ablesen. Von besonderem Interesse ist jedoch, daß sich anhand der Aktenzeichen das jeweils entscheidende Gericht sowie die betreffende Gerichtsabteilung erkennen läßt. Da in vielen Gerichten zwischen Abteilungsnummern und RichterIn bzw. Richter eine direkte

Zuordnung besteht, läßt sich in Verbindung mit den Geschäftsverteilungsplänen der Gerichte zumindest für Einzelrichterentscheidungen zugleich der individuelle Entscheidungsträger identifizieren. Somit besteht die prinzipielle Möglichkeit, sowohl regionale als auch richtervergleichende Sanktionsanalysen durchzuführen und hierbei viele der strafzumessungsrelevanten Fall- und Tätermerkmale zu kontrollieren. Eine effektive statistische Kontrolle solcher Störeinflüsse (z. B. in Form von Kovariaten in regressionsanalytischen Modellen) ist jedoch leider nur bei solchen Delikten möglich, über die die einzelnen Richterinnen und Richter sehr häufig zu entscheiden haben. Mögliche Untersuchungen beschränken sich infolgedessen leider auf sogenannte »Massendelikte« wie den einfachen Diebstahl.

#### *Sanktionsdisparitäten im Bereich des Allgemeinen Strafrechts*

Bislang liegen zwei empirische Arbeiten vor, die sowohl lokale als auch richterbezogene Unterschiede im Sanktionsverhalten auf der Basis von Bundeszentralregisterauszügen untersuchen. Beide Arbeiten dokumentieren und analysieren die Sanktionspraxis bei einfachem Diebstahl (§ 242 StGB) in drei großstädtischen Amtsgerichten. In der Untersuchung von *Oswald* (1994) wird für den Vergleich der Sanktionen eine intervallskalierte Strafhärteskala genutzt. Diese Skala erlaubt die Abbildung unterschiedlicher Strafformen (Geldstrafe, bedingte und unbedingte Freiheitsstrafe) mit jeweils unterschiedlichen Strafhöhen (Tagessätze, Dauer der Freiheitsstrafe) auf einer einheitlichen Meßlatte. *Langer* (1994) hingegen untersucht unter Rückgriff auf weitere Datenquellen (staatsanwaltschaftliche Erledigungsstatistiken und Aktenhebungen) mittels äußerst elaborierter statistischer Verfahren, die zugleich staatsanwaltschaftliche Einflüsse zu berücksichtigen erlauben, u. a. Unterschiede in den Entscheidungen für oder gegen (bedingte und unbedingte) Freiheitsstrafen. Von den vielfältigen interessanten Befunden zu lokalen und richterbezogenen Sanktionspraktiken sollen hier nur einige wenige genannt werden (siehe hierzu auch *Löschper* 1996). Generell lassen sich die Sanktionsentscheidungen offensichtlich recht gut vorhersagen. Je nach Amtsgericht und Vorhersagemodell können zwischen 28 und 49 Prozent der Varianz der Strafzumessungsentscheidungen mittels der in die Untersuchungen einbezogenen Variablen erklärt werden. Den größten Erklärungswert haben hierbei legale Strafzumessungsmerkmale wie z. B. Vorstrafenbelastung und Schadenshöhe (Diebstahl geringwertiger Sachen gemäß §§ 242, 248 a StGB). Es lassen sich jedoch auch systematische gerichtliche- und richterbezogene Sanktionsdisparitäten nachweisen. Diese sind mit einem Erklärungsanteil von ca. 1 bis 4 Prozent der Gesamtvarianz relativ gleichbedeutend. Absolut gesehen sind sie aber von deutlich geringerer Bedeutung als die auf legale Merkmale rückführbaren Sanktionsunterschiede (vgl. *Oswald* 1994, S. 171; *Langer* 1994, S. 317, 325). Die Differenzen zwischen einzelnen Richterinnen und Richtern sind zwar absolut gesehen begrenzt, korrelieren aber in erheblichem Maße mit individuellen Unterschieden in zentralen strafbezogenen Attitüden (vgl. *Oswald* 1994, S. 176 ff.). Ferner kann *Langer* (1994, S. 320 ff.) zeigen, daß sich hinter den ebenfalls nicht allzu groß anmutenden regionalen Sanktionsdisparitäten dennoch deutliche Divergenzen in der Sanktionspraxis verbergen können. So ist die durchschnittliche Strafhärte in den drei Amtsgerichten über alle Fälle hinweg betrachtet zwar nicht allzu ungleichmäßig. Infolge deutlicher Unterschiede in der Gewichtung der verschiedenen Fall- und Tätermerkmale kann das Haftrisiko für einzelne Tätergruppen (z. B. mehrfach vorbestrafte männliche Ausländer) gleichwohl ganz erheblich variieren. Inwieweit auch bei der Interpretation der richterbezogenen Sanktionsdisparitäten Unterschiede in der Gewichtung einzelner Fallmerkmale zu berücksichtigen sind, ist indes weitgehend ungeklärt.

#### *Möglichkeiten und Probleme der Sanktionsforschung im Jugendstrafrechtsbereich*

Das Bundeszentralregister umfaßt neben dem eigentlichen Zentralregister (s. o.) mit dem Erziehungsregister zugleich eine Datenbank, in der die Reaktionen der Jugendstaatsanwaltschaften und Jugendgerichte auf Rechtsverstöße Jugendlicher und Heranwachsender dokumentiert werden. Erhebliche Schwierigkeiten im Hinblick auf die Skalierung der Rechtsfolgen erwachsen jedoch aus dem Umstand, daß im Bereich der jugendrichterlichen Tätigkeit zwei vollkommen unterschiedliche Regelsysteme zur Anwendung gelangen. Während bei Jugendlichen (14 bis 17 Jahre) der Fall im Allgemeinen nach den Regeln des Jugendstrafrechts verhandelt wird, entscheidet bei Heranwachsenden (18 bis 20 Jahre) das hier ebenfalls zuständige Jugendgericht, ob es die Regeln des Jugendstrafrechts oder diejenigen des Allgemeinen Strafrechts anwendet (§ 105 JGG). Problematisch ist dies deshalb, weil mit den verschiedenen Regelsystemen zugleich vollkommen unterschiedliche Rechtsfolgen zur Verfügung stehen.

Wird nach Allgemeinem Strafrecht sanktioniert, lassen sich die richterlichen Reaktionsalternativen im Wesentlichen auf Geldstrafen, Bewährungsstrafen und unbedingte Freiheitsstrafen begrenzen. Wie im vorigen Abschnitt skizziert, könnte man in diesem Falle zur Analyse von Sanktionsdisparitäten zum einen auf kategoriale Sanktionsskalen (z. B. Geldstrafe vs. Freiheitsstrafe) zurückgreifen. Zum anderen könnte man z. B. die von *Oswald* (1994) entwickelte intervallskalierte Strafhärteskala nutzen.

Welche Reaktionsalternativen stehen nun bei Anwendung der Regeln des Jugendstrafrechts zur Verfügung? Zum einen kann das Verfahren sowohl durch die Staatsanwaltschaft als auch von der Jugendrichterin bzw. dem Jugendrichter entweder folgenlos oder nach Erfüllung unterschiedlichster erzieherischer Maßnahmen eingestellt werden (§§ 45 und 47 JGG). Auch diese Entscheidungen werden – ebenso wie Freisprüche (wegen mangelnder Reife) – im Erziehungsregister vermerkt und können bei Sanktionsanalysen berücksichtigt werden. Sollte die Richterin bzw. der Richter hingegen eine förmliche Verurteilung für notwendig erachten, reichen die möglichen Rechtsfolgen über die sog. Erziehungsmaßregeln (z. B. die Erbringung einer Arbeitsleistung oder die Teilnahme an einem sozialen Trainingskurs etc.) und Zuchtmittel (z. B. eine förmliche Zurechtweisung, eine Schadenswiedergutmachung oder eine Arreststrafe, d. h. eine Art kurzer Freiheitsentzug) bis hin zur bedingten oder unbedingten Jugendstrafe. Bei der unbedingten Jugendstrafe handelt es sich um einen längeren Freiheitsentzug in einer Jugendstrafanstalt. Diese qualitativ sehr unterschiedlichen Reaktionen können in quantitativer Hinsicht jeweils äußerst variabel gestaltet und zudem bis auf einige Ausnahmen auf vielfältige Weise miteinander kombiniert werden.

Welche Schlußfolgerungen lassen sich aus den bisherigen Ausführungen ableiten? Ein grundsätzliches Problem bei der Analyse jugendrichterlichen Sanktionsverhaltens ergibt sich aus der Tatsache, daß Heranwachsende entweder nach den Richtlinien des JGG oder nach denen des Allgemeinen Strafrechts sanktioniert werden können. Die Entscheidung hierüber kann jedoch mit systematischen Unterschieden in der Sanktionshärte verbunden sein. Nun wäre es sicherlich interessant, solche Unterschiede aufzudecken. Hierzu müßten jedoch die beiden Rechtsfolgensysteme auf eine einheitliche Skala abgebildet werden. Aufgrund der nahezu unbegrenzten Kombinationsmöglichkeiten quantitativ und qualitativ variierender Reaktionsmöglichkeiten ist die Konstruktion einer intervallskalierten Skala für die Rechtsfolgen gemäß Jugendstrafrecht wahrscheinlich prinzipiell nicht erreichbar. So bliebe letztendlich wohl nur eine trichotome Reaktionsskalierung mit den drei umfassenden Kategorien (1) unbedingte Jugend-/Freiheitsstrafe, (2) bedingte Jugend-/Freiheitsstrafe und (3) Erziehungsmaßregeln bzw. Zuchtmittel und Geldstrafen als größter gemeinsamer Nenner. Auch wenn man einmal davon absieht, daß aufgrund der Unterschiede zwischen Zentralregister und Erziehungsregister Verfahrenseinstellungen bei einer gemeinsamen Skalierung der Rechtsfolgen leider unberücksichtigt bleiben müß-

ten, was zu systematischen Verzerrungen führen kann, handelt es sich dennoch nur um eine Scheinlösung. Der Bildung einer gemeinsamen kategorialen Sanktionsskala steht in erster Linie entgegen, daß viele der jugendstrafrechtlichen Rechtsfolgen (z. B. Erziehungsbeistandschaft, Arreststrafen etc.) prinzipiell kein Äquivalent im Allgemeinen Strafrecht haben. Am ehesten würde man sie wahrscheinlich den Geldstrafen im Allgemeinen Strafrecht gleichsetzen. Die Ähnlichkeit zwischen Geldstrafen und den »vergleichbaren« jugendstrafrechtlichen Rechtsfolgen besteht jedoch allein darin, daß es sich jeweils um die Alternativen zu »echten« Kriminalstrafen (Freiheits- bzw. Jugendstrafen) handelt. Die Wahl bzw. Bedeutung einer Jugendstrafe bzw. einer Freiheitsstrafe hängt jedoch nicht zuletzt davon ab, welche Alternativen im jeweiligen Regelsystem vorhanden sind. Da diese aber nicht vergleichbar sind, läuft man letztendlich Gefahr, im Hinblick auf alle Reaktionsalternativen immer jugendstrafrechtliche »Äpfel« mit allgemeinstrafrechtlichen »Birnen« gleichzusetzen.

Viele der genannten Probleme lassen sich wahrscheinlich nur dadurch vermeiden, daß Strafzumessungsuntersuchungen im Bereich der Jugendstrafrechtspraxis auf Reaktionen gegenüber Jugendlichen beschränkt werden. Für diesen Fall lassen sich durchaus differenzierte Analysen jugendrichterlichen Sanktionsverhaltens vornehmen. So wurde z. B. von Hupfeld (1996, S. 16 f.) in Zusammenarbeit mit in der Jugendstrafrechtslehre tätigen JuristInnen und praktisch tätigen JugendstaatsanwältInnen und JugendrichternInnen eine mehrstufige Skala jugendstrafrechtlicher Rechtsfolgen entwickelt. Grundidee war hierbei, solche qualitativen Sprünge in den richterlichen Reaktionen abzubilden, die sowohl nach gesetzlicher Intention als auch nach Einschätzung der PraktikerInnen mit Unterschieden in der intendierten Eingriffsintensität einhergehen. Die entwickelte Rangskala jugendrichterlicher Reaktionsmöglichkeiten beinhaltet die folgenden sechs Stufen:

Die geringste intendierte Eingriffsintensität weist die folgenlose Verfahrenseinstellung auf (1). Deutlich eingriffsintensiver ist eine Verfahrenseinstellung nach Erteilung von Weisungen oder Auflagen (2). Erscheint auch dies der RichterIn bzw. dem Richter als »nicht ausreichend«, so erfolgt eine förmliche Verurteilung. Im günstigsten Fall handelt es sich bei den mit der Verurteilung einhergehenden Rechtsfolgen um ambulante Erziehungsmaßregeln oder Zuchtmittel (3). Die nächste Stufe der Eingriffsintensität bilden die verschiedenen Formen des stationären Zuchtmittels Jugendarrest (4). Arreststrafen werden häufig als letzte Warnung aufgefaßt, bevor zur Jugendstrafe, der »ultima ratio« des Jugendstrafrechts, gegriffen wird. Bei dieser ist von zentraler Bedeutung, ob ihre Verhängung zur Bewährung ausgesetzt wird (5) oder ob es sich um eine unbedingte Jugendstrafe handelt (6).

In weiteren Expertenbefragungen zeigte sich, daß diese sechsstufige Rangskala wahrscheinlich die maximal erreichbare Differenzierung darstellt, hinsichtlich derer zugleich weitestgehende Einigkeit erzielt werden kann. Die Uneinigkeit nimmt rapide zu, sobald die Einschätzung der Eingriffsintensität innerhalb einzelner Kategorien thematisiert wird (z. B. die Abgrenzbarkeit von Erziehungsmaßregeln und ambulanten Zuchtmitteln). Hingegen scheint die Zusammenfassung benachbarter Kategorien (insbes. bedingte und unbedingte Freiheitsstrafen sowie folgenlose und mit ambulanten Maßnahmen verbundene Verfahrenseinstellungen) weniger problematisch zu sein. Da viele der Rechtsfolgen miteinander kombinierbar sind, setzt die Anwendung des Klassifikationssystems jedoch in jedem Fall voraus, daß die jeweils eingriffsintensivste Rechtsfolge für die Einordnung der Gesamtreaktion ausschlaggebend ist. Mit dieser einfachen – maximal sechsstufigen – Skala der Eingriffsintensität lassen sich zwar nicht alle prinzipiell möglichen Rechtsfolgen (z. B. die Erziehungsmaßregel der Anordnung einer Heimunterbringung) eindeutig einordnen.

Die Summe der »problematischen« Fälle ist jedoch in der Praxis meist extrem klein. Generell zu bedenken ist ferner, daß sich die vorgeschlagene Skalierung der Eingriffsintensität an den Intentionen des Gesetzes und der RichterInnen orientiert. Die betroffenen Jugendlichen dürften nicht selten andere Abstufungen vornehmen.

#### *Sanktionsdisparitäten im Bereich des Jugendstrafrechts*

Grundlage der nachfolgend dargestellten Analysen bilden Erziehungsregisterauszüge aus vier großstädtischen Amtsgerichten, die dem Kriminologischen Forschungsinstitut Niedersachsen in weitgehend anonymisierter Form zur Verfügung gestellt wurden. Drei der vier Amtsgerichte sind identisch mit denjenigen der Untersuchungen von Langer (1994) und Oswald (1994). Die Analysen beschränken sich auf einzelrichterliche Entscheidungen gegenüber Jugendlichen, die wegen eines einfachen Diebstahls als schwerwiegendstem Rechtsnormverstoß in den Jahren 1987 bis 1990 vor Gericht standen. Die Ausdehnung des Entscheidungszeitraumes auf vier Jahre wurde notwendig, um eine für inferenzstatistische Analysen sinnvolle Mindestfallzahl pro Abteilung (Richter bzw. Richterin) zu erreichen. Hierzu wurden zudem die Entscheidungen derjenigen Richterinnen und Richter, die weniger als 20 den oben genannten Bedingungen genügenden Entscheidungen fällten, von den Analysen ausgeschlossen. Von den insgesamt 2602 Entscheidungen gegenüber Jugendlichen, die wegen eines einfachen Diebstahls als schwerwiegendstem Rechtsnormverstoß in den Jahren 1987 bis 1990 vor einem der vier Gerichte standen, konnten letztendlich ca. 95 Prozent (N = 2470) in die endgültigen Analysen einbezogen werden. Die Rahmendaten, auf denen die Analysen beruhen, sind in *Tabelle 1* zusammengefaßt.

*Tabelle 1: Verteilung der einzelrichterlichen Entscheidungen bei einfachem Diebstahl (§ 242 StGB) gegenüber Jugendlichen auf Gerichte und Abteilungen*

	Gericht A	Gericht B	Gericht C	Gericht D
In die Analysen einbezogene Abteilungen (RichterInnen)	31	8	9	7
Fälle pro Abteilung:				
Minimum	22	23	20	37
Maximum	86	83	87	60
Durchschnitt	46,7	45,7	37,1	45,9
Fälle pro Gericht	1449	366	334	321

Für jede einzelne Entscheidung wurden den Erziehungsregisterauszügen die in *Tabelle 2* aufgelisteten Informationen zu Fall- und Tätermerkmalen entnommen. Einer zusätzlichen Erläuterung bedürfen jedoch die zur Eingriffsintensität der jeweils gewählten Rechtsfolge entnommenen Angaben. Die nachfolgend dargestellten Befunde stellen einen kleinen Ausschnitt aus umfangreicheren Analysen zur jugendrichterlichen Sanktionspraxis dar (vgl. Hupfeld 1990; 1992). Aufgrund der je nach Analyse teilweise recht geringen Fallzahlen (z. B. bei der Vorhersage der Sanktionsentscheidungen einzelner Richter gegenüber Mehrfachauffälligen) wurden folgenlose Verfahrenseinstellungen und solche nach Erteilung von Weisungen und Auflagen zu einer Gesamtkategorie »Verfahrenseinstellungen« zusammengefaßt (vgl. *Tabelle 2*). Die Verwendung einer solchen fünfstufigen Skala beinhaltet zwar einen gewissen Informationsverlust, sie wurde jedoch notwendig, um eine

Vergleichbarkeit der vielfältigen im gesamten Forschungsprojekt durchgeführten Analysen zu gewährleisten.

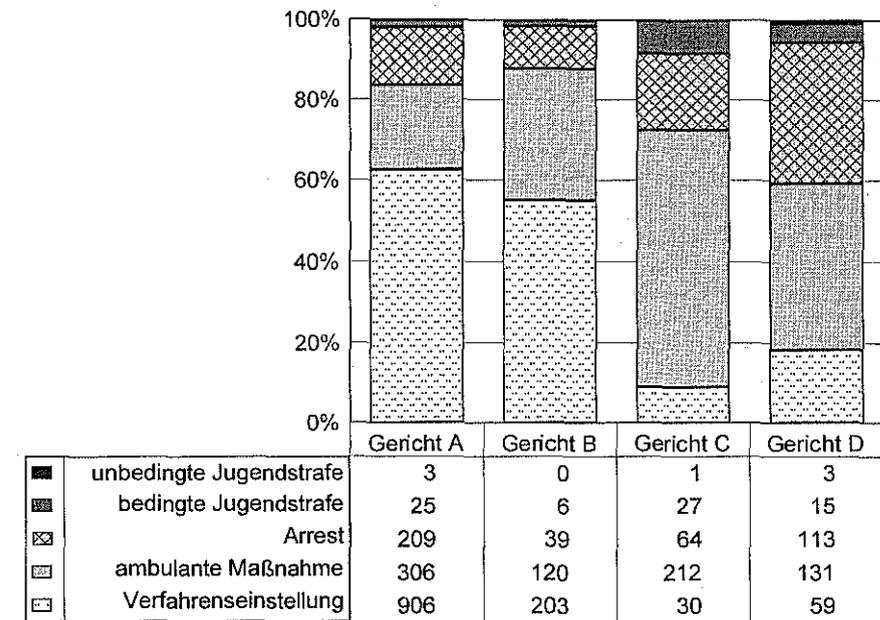
Tabelle 2: Liste der bei den Analysen berücksichtigten Fallmerkmale

Merkmal	Inhaltliche Bedeutung
Tateinheit	Wurde im Rahmen der Tatbegehung gegen eine oder gegen mehrere Rechtsnormen verstoßen (§ 52 StGB)?
Tatmehrheit	Steht der Täter wegen nur einer Straftat oder wegen mehrerer Straftaten vor Gericht (§ 53 StGB)?
Geringwertigkeit	Handelt es sich um einen Diebstahl geringwertiger Sachen (§§ 242, 248 a StGB)?
Vorbelastung	Wie häufig hatte der Täter bzw. die Täterin zuvor bereits wegen strafrechtswidriger Verhaltensweisen Kontakte mit der Jugendstaatsanwaltschaft bzw. dem Jugendgericht, die nicht mit einem Freispruch endeten?
Alter	Wie alt war der Täter bzw. die Täterin zur Zeit der Tat?
Geschlecht	Handelt es sich um einen Täter oder um eine Täterin?
Ausländer	Handelt es sich bei dem Täter bzw. der Täterin um einen Ausländer bzw. eine Ausländerin?
Gericht	In welchem Gericht wurde über den Fall entschieden?
Abteilung	Welche Richterin bzw. welcher Richter hat über den Fall entschieden?
Rechtsfolge	Welche Rechtsfolge wurde gewählt: Verfahrenseinstellung, ambulante Maßnahme, Arrest, bedingte Jugendstrafe, unbedingte Jugendstrafe?

Betrachten wir nun als erstes die Sanktionspraxis in den vier untersuchten Gerichten. Wie in *Abbildung 1* ersichtlich, ist in den Gerichten A und B der Anteil der Verfahrenseinstellungen deutlich größer und der Anteil von Arrest und Jugendstrafen deutlich geringer als in den Gerichten C und D. Es besteht jedoch die Möglichkeit, daß diese gerichtsbezogenen Sanktionsdisparitäten weitgehend auf systematischen Unterschieden in den jeweils sanktionierten Fällen beruhen. Beispielsweise könnten die Staatsanwaltschaften in den Gerichten C und D sehr viel häufiger Verfahren gemäß § 45 JGG einstellen als die Staatsanwaltschaften in den Gerichten A und B. In diesem Falle wäre zu erwarten, daß diejenigen Fälle, über die in den Städten C und D noch von JugendeinzelrichterInnen entschieden wird, im Durchschnitt schwerwiegender sind bzw. die Täter bereits häufiger mit dem Gesetz in Konflikt geraten sind als in den Städten A und B.

Die Frage, ob Sanktionsdisparitäten auch noch nach Berücksichtigung eventueller Unterschiede im Fallmaterial nachweisbar sind, kann z. B. mit Hilfe multivariater Regressionsverfahren beantwortet werden (vgl. z. B. Bortz 1993, S. 415 ff.). Hierbei werden möglichst alle Faktoren, die einen Einfluß auf die Eingriffsintensität der gewählten Rechtsfolgen haben können, gleichzeitig als Prädiktorvariablen in die statistischen Analysen einbezogen.

Abbildung 1: Absolute und prozentuale Verteilung der Entscheidungen gegenüber Jugendlichen bei einfachem Diebstahl



Bei der vorliegenden Untersuchung ist die interessierende abhängige Variable (Eingriffsintensität) jedoch nur rangskaliert. Einige der zentralen Voraussetzungen zur Anwendung der verbreiteten linearen Regressionsmodelle sind somit nicht erfüllt (vgl. Ludwig-Mayerhofer 1990, S. 82; Maddala 1983). Aus diesem Grund wird bei den statistischen Analysen jugendrichterlichen Entscheidungsverhaltens auf den von McKelvey & Zavoina (1975) vorgestellten Ordered-Probit-Ansatz zurückgegriffen.

Alle nachfolgenden Angaben über die statistische Signifikanz sowie über die praktische Bedeutung der einzelnen Prädiktoren<sup>1</sup> beruhen auf hierarchischen Modelltests (vgl. Cohen & Cohen 1983). Hierbei wird jeweils ein Regressionsmodell, das alle relevanten Prädiktoren enthält, mit einem Modell verglichen, das den speziellen Prädiktor nicht enthält, jedoch alle übrigen relevanten Prädiktoren. Führt die Weglassung des speziellen Prädiktors zu einer signifikanten Verschlechterung der Vorhersage, kommt diesem offensichtlich ein überzufälliger eigenständiger Erklärungsbeitrag zu. Der Absolutbetrag dieser Vorhersageverschlechterung ( $\Delta R^2$ ) wird als Indikator für die praktische Bedeutung des speziellen Prädiktors gewertet. Die Angaben über die Güte der Vorhersage der Regressionsmodelle ( $R^2$ -Werte) basieren hierbei auf den von McKelvey & Zavoina (1975) vorgeschlagenen Formeln<sup>2</sup>. Die wichtigsten Ergebnisse der gerichtsübergreifenden Analysen sind in *Tabelle 3* zusammengefaßt.

- 1 Bei allen nominalen Prädiktoren (z.B. Tateinheit, Gericht) wurden aus formalen Gründen sogenannte Indikatorvariablen gebildet, die alle Informationen des nominalskalierten Merkmals in einer speziell kodierten Form enthalten (vgl. Bortz 1993, S. 447 ff.).
- 2 Alle Berechnungen wurden mit Hilfe des Programms NPC (vgl. Hupfeld & Grotten 1992), einer für große Datensätze angepaßten Version des Originalprogramms NPROBIT von McKelvey & Zavoina (1975), durchgeführt.

Tabelle 3: Eigenständiger Erklärungswert der Prädiktorvariablen sowie Güte der Gesamtvorhersage bei gerichtübergreifenden Ordered-Probit-Analysen (abhängige Variable: richterliche Eingriffsintensität)

Prädiktoren	Erklärungswert
Tateinheit	$\Delta R^2 = 0.01^{**}$
Tatmehrheit	$\Delta R^2 = 0.02^{**}$
Geringwertigkeit	$\Delta R^2 = 0.02^{**}$
Vorbelastung	$\Delta R^2 = 0.09^{**}$
Alter	$\Delta R^2 = 0.02^{**}$
Geschlecht	$\Delta R^2 = 0.00$
Ausländer	$\Delta R^2 = 0.01^{**}$
Gericht	$\Delta R^2 = 0.09^{**}$
Gesamtmodell mit allen Prädiktoren (N = 2470 Einzelrichterentscheidungen)	$R^2 = 0.32^{**}$

\*\* p < 0.01

Alles in allem zeigt sich, daß die Eingriffsintensität der von den JugendrichterInnen gewählten Rechtsfolgen recht gut vorhergesagt werden kann. Insgesamt 32 Prozent der Sanktionsvarianz ist mittels der einbezogenen Prädiktoren »aufklärbar«. Die relativ geringen eigenständigen Erklärungsbeiträge der einzelnen Prädiktoren (zwischen 0 und 9 Prozent) sind nicht zuletzt auf teilweise erhebliche systematische Kovariationen zurückzuführen (z. B. zwischen Alter und Vorbelastung). Bis auf das Tätergeschlecht trägt jedoch jeder Prädiktor auch einen signifikanten eigenständigen Beitrag zur Vorhersage der Eingriffsintensität bei. Dies gilt zum einen für legale Merkmale wie das Vorliegen von Tateinheit und Tatmehrheit sowie bei Geringwertigkeit des gestohlenen Gutes. Zum anderen gilt dies aber auch für ein extralegales Merkmal wie den Ausländerstatus. Die Rechtsfolgen, die bei Ausländern gewählt werden, sind signifikant eingriffsintensiver als bei Deutschen, deren Fälle ansonsten vergleichbare Merkmale aufweisen. Nähere Analysen belegen, daß ausländische Täter bei Verfahrenseinstellungen und ambulanten Maßnahmen unterrepräsentiert und bei Arreststrafen überrepräsentiert sind. Von der insgesamt beobachtbaren Sanktionsvarianz »erklärt« der Ausländerstatus jedoch nur ca. 1 Prozent<sup>3</sup>. Die größte Bedeutung bei den Fall- und Tätermerkmalen kommt der Tätervorbelastung zu (9 Prozent »erklärte« Varianz). Interessant ist aber, daß auch die Frage, in welchem Gericht über den Fall entschieden wurde, von ähnlich großer Bedeutung ist. Dies gilt, wie aus Tabelle 3 ersichtlich, selbst nach statistischer Kontrolle der übrigen Fall- und Tätermerkmale.

Neben der hiermit angesprochenen Bedeutung genereller gerichtbezogener Sanktionsdisparitäten wurde in einem weiteren Schritt der Frage nachgegangen, ob zudem

3 Da der Ausländeranteil in den vier Gerichten im Durchschnitt immerhin 30 Prozent beträgt, kann der recht geringe Vorhersagewert nicht allein darauf zurückgeführt werden, daß das Merkmal »Ausländerstatus« selbst nur wenig Varianz aufweist.

überzufällige gerichtbezogene Unterschiede in der Gewichtung von Fall- und Tätermerkmalen bestehen. Zu diesem Zweck wurde mittels hierarchischer Modellvergleiche überprüft, inwieweit die Hinzufügung von Interaktionstermen zu einem Modell, das bereits alle oben aufgeführten Haupteffekte enthält, eine bessere Sanktionsvorhersage erlaubt. Als signifikant erwiesen sich hierbei die Interaktion zwischen Gericht und Anzahl bisheriger Auffälligkeiten ( $\Delta R^2 = 0.02$ ), Gericht und Alter zur Tatzeit ( $\Delta R^2 = 0.007$ ), Gericht und Ausländerstatus ( $\Delta R^2 = 0.006$ ) sowie Gericht und Geringwertigkeit des gestohlenen Gutes ( $\Delta R^2 = 0.007$ ).

Diese Unterschiede in der Gewichtung von Fall- und Tätermerkmalen werden durch die nachfolgende Gegenüberstellung der gericht-internen Sanktionsanalysen verdeutlicht (vgl. Tabelle 4). Das methodische Vorgehen bei diesen Analysen ist weitgehend mit dem oben beschriebenen identisch. Die hierarchischen Modellvergleiche wurden jedoch getrennt für jedes Gericht durchgeführt. Einige der zentralen Befunde seien nachfolgend skizziert. So ist beispielsweise die Anzahl bisher vermerkter Rechtsnormverstöße (Vorbelastung) in den Gerichten A und B der jeweils varianzstärkste Prädiktor. In Gericht D, in dem die Vorhersage der Eingriffsintensität jedoch generell eher mäßig gut gelingt, ist dieser Prädiktor zwar signifikant, jedoch deutlich weniger bedeutsam, und in Gericht C scheint diese Variable die Eingriffsintensität der Rechtsfolgen kaum in systematischer Weise zu beeinflussen. Von besonderem Interesse ist ferner, daß in drei Gerichten signifikante richterbezogene Sanktionsdisparitäten nachweisbar sind. Mit einem Erklärungsbeitrag zwischen 4 und 8 Prozent sind sie zudem von nicht unerheblicher praktischer Bedeutung. In Gericht D sind hingegen keine systematischen richterbezogenen Sanktionsunterschiede nachweisbar.

Tabelle 4: Erklärungswert der Prädiktorvariablen sowie Güte der Gesamtvorhersage bei gericht-internen Ordered-Probit-Analysen (abhängige Variable: richterliche Eingriffsintensität)

Prädiktoren	eigenständiger Erklärungswert			
	Gericht A ( $\Delta R^2$ )	Gericht B ( $\Delta R^2$ )	Gericht C ( $\Delta R^2$ )	Gericht D ( $\Delta R^2$ )
Tateinheit	0.01 **	0.02 **	0.00	0.00
Tatmehrheit	0.01 **	0.03 **	0.03 **	0.03 **
Geringwertigkeit	0.01 **	0.03 **	0.04 **	0.02 **
Vorbelastung	0.16 **	0.11 **	0.01	0.04 **
Alter	0.01 **	0.05 **	0.05 **	0.01 *
Geschlecht	0.00	0.00	0.00	0.00
Ausländer	0.02 **	0.00	0.03 **	0.00
Abteilung	0.08 **	0.04 **	0.07 **	0.01
Gesamtmodell mit allen Prädiktoren	$R^2 = 0.34^{**}$ (N= 1449)	$R^2 = 0.44^{**}$ (N= 366)	$R^2 = 0.27^{**}$ (N= 334)	$R^2 = 0.14^{**}$ (N= 321)

\* p < 0.05;

\*\* p < 0.01

Zusätzliche Analysen zum Nachweis von richterbezogenen Unterschieden in der Gewichtung von Fall- und Tätermerkmalen zeigen, daß in den Gerichten A, B und C signifikante und praktisch bedeutsame Interaktionen zwischen den Variablen »Abteilung« und

»Vorbelastung« nachweisbar sind. Die Berücksichtigung dieser Interaktion erlaubt für Gericht A eine Verbesserung der Vorhersage um weitere 4 Prozent, in Gericht B um 2 Prozent und in Gericht C ebenfalls um ca. 4 Prozent. Andere richterbezogene Unterschiede in der Gewichtung von Fall- oder Tätermerkmalen sind nicht nachweisbar. Dies dürfte u. a. auch darauf zurückzuführen sein, daß der gezielte Nachweis von Interaktionseffekten generell mit erheblichen methodischen Problemen behaftet ist (vgl. *Evans* 1991). Da bei der Testung richterbezogener Effekte die Fallzahlen bereits recht niedrig sind, werden nur sehr starke Interaktionseffekte als signifikant ausgewiesen. So zeigt sich denn auch bei genauerer Analyse der in dieser Untersuchung vorliegenden Haupt- und Interaktionseffekte, daß offensichtlich der größte Teil der zwischen den einzelnen Richterinnen und Richtern bestehenden Sanktionsunterschiede letztendlich auf eine unterschiedliche Gewichtung der Tätervorbelastung zurückzuführen ist. Ähnliches gilt im Übrigen – wengleich in etwas geringerem Umfang – auch für die gerichtszugehörigen Sanktionsunterschiede.

#### *Konsistenzen und Disparitäten in der Jugendstrafrechtspraxis: beruhigend oder alarmierend?*

Wie lassen sich die für den Bereich der deutschen Jugendstrafrechtspraxis berichteten Befunde zusammenfassen und welche Bedeutung kommt ihnen zu? Die Antworten auf diese Fragen dürften je nach Betrachterperspektive sehr unterschiedlich ausfallen. Gemäß der fachlichen Ausrichtung des Autors beschränkt sich die nachfolgende Diskussion primär auf einige methodische und psychologische Aspekte.

Wie bereits erläutert, bestehen bei der Skalierung jugendrichterlicher Handlungsalternativen erhebliche Probleme. Den vorliegenden Analysen liegt eine fünfstufige ordinale Skala der Rechtsfolgen zugrunde. Zu bedenken ist, daß bei Verwendung dieser Skala die erheblichen quantitativen Unterschiede, die innerhalb der einzelnen Rangstufen (Verfahrenseinstellungen, ambulante Maßnahmen, Arrest, bedingte und unbedingte Jugendstrafe) bestehen können, prinzipiell unberücksichtigt bleiben. Für diese Art der Skalierung spricht jedoch, daß sie die zentralen qualitativen »Sprünge« in der Eingriffsintensität jugendstrafrechtlicher Rechtsfolgen abzubilden erlaubt. Bei Anwendung neuerer multivariater Analyseverfahren (z. B. Ordered Probit oder Ordered Logit Regression, vgl. *Greene* 1990) bietet diese einfache Skala einen interessanten Zugang für die empirische Sanktionsforschung. In der vorliegenden Untersuchung konnte gezeigt werden, daß auf diesem Wege sowohl viele der fallbezogenen Hintergründe für die Wahl eingriffsintensiver Rechtsfolgen als auch regionale und richterbezogene Sanktionsdisparitäten analysierbar sind.

Basis der vorgestellten Untersuchungen sind jugendeinzelrichterliche Entscheidungen gegenüber Jugendlichen, die wegen eines einfachen Diebstahls als schwerwiegendstem Rechtsnormverstoß zwischen 1987 und 1990 vor einem von vier großstädtischen Gerichten standen. Die entsprechenden Daten wurden dem Kriminologischen Forschungsinstitut Niedersachsen vom Bundeszentralregister in Berlin in weitgehend anonymisierter Form zur Verfügung gestellt. Allerdings erwies sich bereits die Aufarbeitung der Daten in ein Format, welches die berichteten statistischen Analysen erlaubte, als extrem aufwendig (ausführlich hierzu: *Langer* 1994; *Oswald* 1994). Dieser Umstand stellt ein großes Hindernis für ähnliche Forschungsvorhaben dar. Folgeuntersuchungen wären allerdings sehr wünschenswert, denn die bislang analysierten Daten stellen einen sehr selektiven Ausschnitt aus der Sanktionspraxis dar. Die nicht geklärte Frage der Repräsentativität sollte darum bei der Interpretation der vorliegenden Befunde unbedingt berücksichtigt werden.

Bevor jedoch eine inhaltliche Interpretation der vorliegenden Befunde vorgenommen wird, bedarf es wiederum einiger methodischer Vorüberlegungen. Im Rahmen der empirischen Sanktionsforschung wird zumeist versucht, die Bedeutung unterschiedlichster Prädiktoren für die Vorhersage der richterlichen Eingriffsintensität bzw. der Sanktionshärte zu bestimmen. Als Indikator für die praktische Bedeutung eines einzelnen Prädiktors wird in den meisten Fällen derjenige Anteil der Sanktionsvarianz angesehen, der in systematischer Beziehung zum einzelnen Prädiktor steht. Vereinfacht

wird dieser Varianzanteil mit dem Begriff der erklärten oder aufgeklärten Varianz gekennzeichnet<sup>4</sup>. Die Gesamtheit der Sanktionsvarianz, die mittels aller einbezogenen Prädiktoren aufgeklärt wird, kann man folglich als systematische Varianz bezeichnen. In den vorgestellten Analysen jugendrichterlichen Sanktionsverhaltens können unter Berücksichtigung aller Haupteffekte gerichtszugehörig 32 Prozent der Gesamtvarianz erklärt werden. Bei den gerichtszugehörigen Analysen beträgt der Anteil systematischer Varianz zwischen 14 und 44 Prozent. Welche Bedeutung hat aber der restliche, nicht vorhersagbare Anteil an Sanktionsvarianz? Dieser unsystematische Varianzanteil beträgt immerhin zwischen 86 und 56 Prozent. Hinter diesem recht großen Anteil verbirgt sich zum einen tatsächlich Regellosigkeit. Diese könnte man mit dem Begriff der richterlichen »Tagesform« umschreiben (vgl. *McFatter* 1989). Zum anderen verbirgen sich in diesem Varianzanteil Regelmäßigkeiten, die auf bislang nicht berücksichtigte Einflußgrößen zurückzuführen sind. In welchem Verhältnis die unterschiedlichen Varianzanteile zueinander stehen, läßt sich kaum abschätzen. Sicherlich wäre es begründenswert, auch nicht im Erziehungsregister enthaltene Informationen, wie z. B. solche über die genaue Schadenshöhe, die Erscheinungsform des Diebstahls (z. B. Ladendiebstahl), den genauen Tathergang, die private und berufliche Situation des Angeklagten sowie die Einstellung des Angeklagten zur Tat, in die Analysen einzubeziehen. Hierdurch ließe sich besser abschätzen, wie groß der tatsächliche Anteil unsystematischer Varianz ist. In der Praxis dürften entsprechende Versuche jedoch in aller Regel bereits an dem damit verbundenen Erhebungsaufwand scheitern. Dies gilt zumindest dann, wenn inferenzstatistisch abgesicherte Aussagen über gerichtszugehörige Unterschiede in der Gewichtung all dieser Merkmale getroffen werden sollen.

Kommen wir nun zur Interpretation der systematischen Varianzanteile. Prinzipiell sollte dabei die Möglichkeit bedacht werden, daß auch die hier nachgewiesenen Einflüsse und Disparitäten auf systematische Kovariationen mit bislang nicht berücksichtigten Fallmerkmalen (s. o.) zurückzuführen sein könnten. Ebenso besteht aber die Möglichkeit, daß sich die Effekte nach einer Kontrolle weiterer Drittvariablen sogar noch vergrößern. Ob und welche Konfundierungen vorliegen, ist leider nicht eruierbar. Die nachfolgenden Interpretationen beziehen sich darum auf die Befundlage, so wie sie sich in der vorliegenden Untersuchung darstellt.

Erklärte Varianzanteile, die zwischen 14 und 44 Prozent der Gesamtvarianz liegen, lassen zumindest die Schlussfolgerung zu, daß die Regelmäßigkeit insgesamt ähnliche Ausprägungen erreicht wie im Bereich des Allgemeinen Strafrechts. Zudem scheint auch die relative Bedeutung der einzelnen Fall- und Tätermerkmale durchaus vergleichbar zu sein (vgl. *Langer* 1994; *Oswald* 1994). Ähnlich wie bei *Langer* (1994) erweist sich die Nationalität des Täters als bedeutsam. Dies gilt zumindest für zwei der vier untersuchten Gerichte. Der Befund, daß ausländische Täter bei ambulanten Maßnahmen unterrepräsentiert sind, deckt sich zudem mit Ergebnissen von *Drewniak* (1996). Neben diesem extralegalen Merkmal erweisen sich jedoch primär legale Fall- und Tätermerkmale als relevante Prädiktoren. Dies gilt, wiederum ähnlich wie im Bereich des Allgemeinen Strafrechts, in erster Linie für die Vorbelastung des Täters. Hingegen scheinen die systematischen gerichtszugehörigen Disparitäten deutlich höhere Maximalwerte anzunehmen. Ohne Einbeziehung von Interaktionseffekten lassen sich 9 Prozent der Sanktionsvarianz allein auf systematische Unterschiede zwischen den vier untersuchten Gerichten zurückführen. Innerhalb der Gerichte lassen sich wiederum bis zu 8 Prozent der Gesamtvarianz durch systematische Unterschiede zwischen den Richterinnen und Richtern erklären. Die größten gerichtszugehörigen und richterbezogenen Unterschiede finden sich bei der Bereitschaft zur Verfahrenseinstellung sowie der Häufigkeit einer mit einem förmlichen Urteil verbundenen Anordnung ambulanter Maßnahmen. Im Hinblick auf die Bereitschaft zur Verhängung bedingter und unbedingter Jugendstrafen sind die Disparitäten deutlich geringer.

4 Zur Frage der formalen Angemessenheit dieser Bezeichnung vgl. *Guttman* 1977 und *Stelzl* 1982.

Man könnte aus der Tatsache, daß im Erziehungsregister auch Verfahrenseinstellungen registriert werden, einen glücklichen Umstand sehen, erlaubt er doch u. a. die Aufdeckung erheblicher Sanktionsdisparitäten. Für die betroffenen Jugendlichen dürfte die Tatsache der vollständigen Registrierung der Legalkarriere allerdings oftmals nachteilige Folgen haben, insbesondere wenn man bedenkt, welch zentrale Bedeutung der Vorbelastung im Hinblick auf die jugendrichterliche Sanktionswahl im Allgemeinen zukommt (vgl. hierzu auch *Gerken & Berlitz* 1988).

Die bisher zusammengefaßten Analysen könnten die Schlußfolgerung nahelegen, zwischen den Gerichten bzw. den JugendeinzelrichterInnen bestünden primär generelle Unterschiede in der Punitivität. Durch weitergehende Untersuchungen konnte jedoch gezeigt werden, daß die gerichtlichen und richterbezogenen Sanktionsdisparitäten offensichtlich zu einem beträchtlichen Anteil auf Unterschieden in der Gewichtung der Tätervorbelastung beruhen. Der Nachweis solch erheblicher Unterschiede im Umgang mit Mehrfachauffälligkeit zeigt, daß bereits gegenwärtig Alternativen zur vielfach angeprangerten automatischen Sanktionseskalation gegenüber Mehrfachauffälligen praktiziert werden. In Untersuchungen mit simuliertem Fallmaterial konnte zudem gezeigt werden, daß sich recht gut erklären und vorhersagen läßt, welche richterlichen Überlegungen und Attitüden mit Unterschieden im Ausmaß der Sanktionseskalation assoziiert sind (vgl. *Hupfeld* 1996).

Die vorliegenden Untersuchungen zeigen gleichwohl, daß die Sanktionsforschung speziell bei der Untersuchung richterbezogener Unterschiede in der Gewichtung einzelner Fallmerkmale rasch an ihre Grenzen stößt. So besteht bei der Analyse realer Sanktionsentscheidungen ein wesentliches Problem darin, daß für statistische Analysen hinreichende Fallzahlen nur im Bereich der Bagatelldelikte erreichbar und zudem die jeweiligen Fallinformationen recht beschränkt sind. Die Grenzen des prinzipiell Machbaren im Hinblick auf die Aufdeckung tatsächlich bestehender Sanktionsdisparitäten dürften mit den jetzt vorliegenden Untersuchungen bereits weitgehend erreicht sein. Zusätzliches Untersuchungen könnten wahrscheinlich primär die bislang unbeantwortete Frage der Repräsentativität der Befunde fokussieren.

Bei der Verwendung fiktiven Fallmaterials bestehen zwar neben den angesprochenen Validitätsproblemen ebenfalls erhebliche Schwierigkeiten bei einem statistischen Nachweis von Interaktionseffekten (vgl. *Hupfeld* 1996, S. 220 ff.). Diese sind jedoch prinzipiell lösbar. Problematischer erscheint dem Autor vielmehr der Umstand, daß die Variation einzelner Fallmerkmale oftmals relativ beliebig erscheint und für die Befragten zudem äußerst ermüdend ist. Bereits jetzt dürfte es schwierig sein, ein Fallmerkmal ausfindig zu machen, das noch nicht in irgendeiner Untersuchung experimentell variiert wurde. Interessiert man sich zudem dafür, welche Fallmerkmale von welchen RichterInnen bei der Sanktionsentscheidung berücksichtigt werden, stellt sich die zusätzliche Frage, welche Personenmerkmale auf Seiten der RichterInnen in die Untersuchung einbezogen werden sollen. Auch im Hinblick auf den dominierenden Umgang mit diesem Problem kann man sich des Eindrucks einer gewissen Beliebigkeit in weiten Bereichen der Forschung kaum erwehren. Das hier angesprochene doppelte Auswahlproblem ließe sich vermutlich nur durch systematischen Bezug auf (interaktionale oder transaktionale) Theorieansätze beheben. Entsprechende theoretische Modelle liegen jedoch bislang kaum vor. Erst wenn die Forschung in diesem Punkt Fortschritte macht, könnte die bereits jetzt vorliegende Fülle von Einzelbefunden zu einem Gesamtbild zusammengesetzt werden. Mit dieser Kritik soll keinesfalls der Eindruck erweckt werden, die bisherige Forschung habe keinerlei Erkenntnisse zu der Frage liefern können, welche Überlegungen RichterInnen in bestimmten Einzelfällen anstellen und welche Konsequenzen dies für ihre Handlungen hat. Aufgezeigt werden sollen einzig einige Grenzen und Desiderata, die zumindest aus der Sicht des Autors bestehen.

Wie aber stellt sich die Bewertung der vorliegenden Forschungsbefunde dar, wenn man nach den psychischen und verhaltensmäßigen Auswirkungen und der gesellschaftspolitischen Bedeutung fragt, die Sanktionsdisparitäten haben? Wirkt sich z. B. eine milde, nachsichtige Sanktionspraxis in spezial- oder generalpräventiver Hinsicht anders – möglicherweise gar nachteilig – aus? Werfen wir zunächst einen kurzen Blick auf die Abschreckungswirkung staatlicher Strafen. Eine merkbliche Wirkung staatlicher Strafen könnte man dann vermuten, wenn man zum einen bei (potentiellen) Tätern ein rationales Kosten-Nutzen-Kalkül voraussetzt und zum anderen annimmt, daß die staatliche Strafe im Rahmen dieses Kosten-Nutzen-Kalküls überhaupt ins Gewicht fällt. Ein Vermeidungsverhalten aufgrund sanktionsbezogener Überlegungen scheint jedoch nur bei wenigen Jugendlichen und Heranwachsenden nachweisbar zu sein. Bei denjenigen, die generell risikobehaftete Verhaltensweisen bevorzugen, kann sich die Strafandrohung sogar delinquenzfördernd auswirken (vgl. *Lösel* 1975, S. 228; *Lyng* 1993). Die Androhung härterer Sanktionen kann für diese Personen den Reiz der Übertretung gesetzlicher Vorschriften vergrößern. Eine tatsächliche Sanktionierung kann zu einer Staturerhöhung des Bestraften beitragen und die Vorbildfunktion innerhalb der peer group steigern. Bei den meisten Jugendlichen und Heranwachsenden hat die subjektive Einschätzung der Sanktionsrisiken jedoch praktisch keinen Effekt auf kriminelles Verhalten. In aller Regel haben formelle Abschreckungsfaktoren gegenüber den sozialen Kosten bzw. sozialer Mißbilligung nur eine marginale Bedeutung (vgl. *Diekmann* 1980; *Schumann, Berlitz, Guth & Kaulitzki* 1987). Eine eindeutig abschreckende Wirkung staatlicher Strafen auf (potentielle) Täter ist empirisch kaum zu belegen.

Wie sieht es jedoch im Hinblick auf die »sittenbildende Kraft« (vgl. *Mayer* 1936, S. 33), die Stärkung der Wertordnung (vgl. *Streng* 1987, S. 50) bzw. den Einfluß auf die moralische Orientierung (vgl. *Roxin* 1978, S. 306) und Normakzeptanz (vgl. *Jakobs* 1983, S. 9) aus? Für die reine Bestätigung der Normgeltung dürfte allein die Tatsache der staatlichen Sanktionierung wichtiger sein als die absolute Sanktionshöhe. Im Hinblick auf andere Aspekte der positiven (General-)Prävention wurde hingegen vielfach vermutet, daß die erwünschte Wirkung weniger eine Frage der absoluten Härte oder Milde als vielmehr der relativen »Angemessenheit« der Sanktion ist. Doch selbst wenn die Frage der Angemessenheit auf die der »Schuldangemessenheit« reduziert wird, sind die Auswirkungen von Sanktionsdisparitäten nicht vorhersagbar. Dies gilt auch dann, wenn man nicht die Bewertungen der konkreten Täter oder Opfer, sondern ausschließlich diejenigen der »Allgemeinheit« fokussiert. In der Bevölkerung besteht offensichtlich weder im Hinblick auf die Bewertung der Schwere von Straftaten Einigkeit, noch im Hinblick auf die Härte der meisten strafrechtlichen Sanktionen (vgl. *Haley* 1984; *Smaus* 1985). Zudem hängt die von »Normalbürgern« jeweils bevorzugte Sanktion in erheblichem Maße davon ab, wie viele Informationen den urteilenden Personen über den konkreten Fall vorliegen. Je mehr Fallinformationen den Beurteilern vorliegen, um so mehr erweist sich beispielsweise die Annahme einer »punitiven Bevölkerung« als Mythos (vgl. *Cumberland & Zamble* 1992; *Zamble & Kalm* 1990). In der Praxis dürfte es die gerechte Sanktion also gar nicht geben. Und welche, wie viele oder wie gut informierte Personen die Strafe als (noch) gerecht empfinden sollen, wird selbst von Vertretern der positiven Generalprävention nicht beantwortet. Zudem ist zu bedenken, daß die staatliche Sanktionspraxis nur eine von sehr vielen Quellen darstellt, die einem »Normalbürger« Informationen über die Geltung von Normen geben können. Aus all diesen Gründen verwundert es kaum, daß selbst Einflüsse erheblicher Veränderungen der Strafrechtspraxis auf Normakzeptanz und Moral, sofern sie überhaupt nachweisbar sind, oftmals komplex und widersprüchlich sind (vgl. *Karstedt*

1993; Schumann 1989). Die in der vorliegenden Arbeit nachgewiesenen Sanktionsdisparitäten dürften daher kaum eine gefährliche Wirkung entfalten.

Die bisherigen Ausführungen beschäftigten sich allein mit der Frage der Auswirkung von »milden«, »harten« oder von unterschiedlichen Personen als »gerecht« empfundenen staatlichen Sanktionen. Hierbei wurde nicht davon ausgegangen, daß jemand die tatsächlich vorhandene Sanktionsvarianz bemerkt. Es wurde primär danach gefragt, wie sich die unterschiedlichen Praktiken jeweils auswirken. Wie aber muß man allein die Existenz nicht unerheblicher Sanktionsdisparitäten beurteilen? Auch hier lassen sich nicht nur negative, sondern auch einige positive Aspekte nennen. Werfen wir zuerst einen Blick auf eventuelle negative Auswirkungen. In einer Reihe von Untersuchungen konnte gezeigt werden, daß sowohl die Bewertung der Ergebnisse außergerichtlicher Mediationsbemühungen als auch diejenige von gerichtlichen Sanktionsentscheidungen nicht nur von individuellen Erwartungen, sondern auch von interpersonalen Vergleichen beeinflusst wird (vgl. Casper, Tyler & Fisher 1988; Hassebrauck 1988; Tyler 1990). Bemerkenswert ist, daß ein Täter, der für seine Tat härter bestraft wurde als eine andere Person in einem vergleichbaren Fall, beeinflusst dies in der Regel die Akzeptanz der konkreten Sanktion negativ. Die hiermit angesprochene Frage des gerechten Strafmaßes stellt den zentralen Aspekt der retributiven Gerechtigkeit dar (vgl. Haley 1984). Es ließe sich somit vermuten, daß bei Bekanntwerden der systematischen gerichtlichen und richterbezogenen Sanktionsdisparitäten ebenso wie bei Bekanntwerden von unsystematischen Disparitäten (Tagesform) die Akzeptanz richterlicher Entscheidungen drastisch sinkt. Neben Aspekten der retributiven Gerechtigkeit spielen jedoch auch andere Faktoren nachweislich eine große Rolle. So konnte wiederholt gezeigt werden, daß prozedurale Aspekte, wie z. B. die Durchschaubarkeit des Entscheidungsverfahrens und der Entscheidungsregeln, die Beteiligungs- und Einspruchsmöglichkeiten der Betroffenen, die Unvoreingenommenheit, Freundlichkeit und Höflichkeit der RichterInnen für die Akzeptanz der konkreten Sanktionsentscheidung sowohl für Täter als auch für Opfer von großer Bedeutung sind. Im Hinblick auf die allgemeine Bewertung der Entscheidungsträger, das generelle Vertrauen in den Rechtsstaat und die längerfristige Bindung an die Rechtsnormen kommt der prozeduralen Fairness nach bisherigem Kenntnisstand eine deutlich größere Bedeutung zu als Aspekten der retributiven Gerechtigkeit (vgl. Tyler 1990).

Die Erforschung von Sanktionsdisparitäten ist sicherlich auch für die Zukunft eine wichtige Aufgabe. Ihr Vorkommen ist gleichwohl nicht nur negativ zu bewerten. Zum einen stellt sich die Frage, ob in einer pluralistischen modernen Gesellschaft, in der man sich bereits über den Sinn der Strafe kaum einigen kann, Sanktionsdisparitäten tatsächlich unerwünscht sind. Sie zeigen zum einen an, daß sowohl der Zweck des Strafens als auch die jeweils genutzten Mittel diskussionswürdig und diskussionsfähig sind. Zum anderen lassen sie erkennen, »daß Entwicklungsprozesse in Gang sind, daß Innovationen möglich sind« (Krauss 1989, S. 137). Zudem relativieren sich Befürchtungen hinsichtlich ihrer negativen psychologischen Konsequenzen, sobald neben den tatsächlich gewählten Rechtsfolgen auch die Prozesse der Entscheidungsfindung und Entscheidungsvermittlung berücksichtigt werden. Zukünftige Forschung sollte darum nicht allein Unterschiede im Strafmaß und ihre Auswirkungen fokussieren, sondern vermehrt prozedurale Aspekte (z. B. Kommunikationsmöglichkeiten, Durchschaubarkeit der Verfahrensregeln und Vermittlung der zur Anwendung gelangenden Regeln bei der Findung einer »gerechten« Entscheidung) berücksichtigen.

#### Summary

The present study focuses on sentencing disparities between different juvenile courts and juvenile court judges. We start with some general comments upon the background and methods of sentencing research. Subsequently the paper sketches the research strategies and central findings of two recently published studies dealing with court sentences passed on adult offenders (Langer 1994; Oswald 1994). In a next step we discuss differences between research strategies concerning the analyses of juvenile- and adult-court cases and present some first empirical results. The analyses are based upon 2470 sentences passed on juvenile offenders whose major offence was petty theft (§ 242 StGB). All cases are gathered from the correctional register of juvenile delinquents, a special register of the Federal Central Register (FCR) in Berlin. The analyses reveal substantial disparities between courts and between judges. Most of these disparities can be reduced to differences in reactions to recidivism. The results are discussed in comparison with court sentences passed on adult offenders. In a final part we generally discuss psychological and behavioural effects of sentencing disparities on offenders and the public and end with some considerations about future research.

#### Literatur:

- Albrecht, H.-J.: Strafzumessung und Vollstreckung bei Geldstrafen unter Berücksichtigung des Tagessatzsystems, Berlin 1980. – Albrecht, H.-J.: Strafzumessung bei schwerer Kriminalität, Berlin 1994. – Bortz, J.: Lehrbuch der Statistik, Berlin 1993; Casper, J.D., Tyler, T.T. & Fisher, B.: Procedural Justice in Felony Cases. *Law and Society Review* 1988, 483-507. – Cohen, J. & Cohen, P.: Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, Hillsdale 1983. – Cumberland, J. & Zamble, E.: General and Specific Measures of Attitudes Toward Early Release of Criminal Offenders. *Canadian Journal of Behavioural Science* 1992. – Diekmann, A.: Die Befolgung von Gesetzen. Empirische Untersuchungen zu einer rechtssoziologischen Theorie, Berlin 1980. – Dzwoniat, R.: Ambulante Maßnahmen für junge Straffällige, Baden-Baden 1996. – Evans, M.G.: The Problem of Analyzing Multiplicative Composites. *American Psychologist* 1991, 6-15. – Exner, F.: Studien über die Strafzumessung der deutschen Gerichte, Leipzig 1931. – Gerken, J. & Berlitz, C.: Sanktionseskalation. Zum fatalen Zusammenspiel von Erziehungsideologie und Registerrecht, in: Gerken, J. & Schumann, K.F., Ein trojanisches Pferd im Rechtsstaat, Pfaffenweiler 1988, 11-39. – Giebring, H.: Ungleichheiten in der Strafzumessungspraxis und die Strafzumessungslehre – Versuch einer Analyse aus der Sicht eines Strafrechtswissenschaftlers, in: Pfeiffer, C. & Oswald, M.E., Strafzumessung. Empirische Forschung und Strafrechtsdogmatik im Dialog, Stuttgart 1989, 77-125. – Greene, W.H.: Econometric Analysis, New York 1990. – Guttman, L.: What is not What in Statistics. *The Statistician* 1977, 81-107. – Haley, H.J.: Retribution and the Definition of Just Measures of Pain, in: Müller, D.J., Blackman, D.E. & Chapman, A.J., Psychology and Law, Chichester 1984, 393-405. – Hassebrauck, M.: Die Bewertung von Gerichtsurteilen: Der Einfluß intra- und interpersonaler Vergleiche. *Zeitschrift für Sozialpsychologie* 1988, 131-138. – Hood, R.: Sentencing the Motorist Offender. A Study of Magistrates' Views and Practices, London 1972. – Hupfeld, J.: Auswirkungen der wiederholten Zuständigkeit auf das jugendrichterliche Sanktionsverhalten (unveröff. Manuskript), Hannover 1990. – Hupfeld, J.: Gerichts- und richterbezogene Sanktionsunterschiede in der Jugendstrafrechtspraxis. Eine empirische Analyse anhand von Daten des Bundeszentralregisters (unveröff. Manuskript), Hannover 1992. – Hupfeld, J.: Jugendrichterliches Handeln, Baden-Baden 1996. – Hupfeld, J. & Groten, L.: NPC: Ein PC-Programm für multivariate Regressionsanalysen bei ordinalen abhängigen Variablen, Hannover, Ulm 1992. – Jakobs, G.: Strafrecht, Allgemeiner Teil, Berlin 1983. – Kapardis, A.: Sentencing by English Magistrates as a Human Process, Nicosia 1985. – Karstedt, S.: Normbindung und Sanktionsdrohung, Frankfurt 1993. – Kette, G.: Urteilsbildung bei Laienrichtern, Berlin 1994. – Krauss, D.: Trifft Strafungleichheit das Strafrecht an seinem Nerv?, in: Pfeiffer, C. & Oswald, M.E., Strafzumessung. Empirische Forschung und Strafrechtsdogmatik im Dialog, Stuttgart 1989, 134-137. – Langer, W.: Staatsanwälte und Richter. Justizielles Entscheidungsverhalten zwischen Sachzwang und lokaler Justizkultur, Stuttgart 1994. – Lemon, N. & Bond, R.: Some Methodological Problems in Sentencing Research, in: Pennington, D.C. & Lloyd-Bostock, S., The Psychology of Sentencing, Oxford 1987, 46-54. – Lösel, F.: Handlungskontrolle und Jugenddelinquenz, Stuttgart 1975. – Ludwig-Mayerhofer, W.: Multivariate Logit-Modelle für ordinalskalierte abhängige Variablen. *ZA-Information* 1990, 62-88. – Löschper, G.: Neues aus der deutschsprachigen Rechtspsychologie: Die Analyse des richterlichen Strafens. *Kriminologisches*

Journal 1996, 198–219. – *Lyng, S.*: Dysfunctional Risk Taking: Criminal Behavior as Edgework, in: *Bell, N.J. & Bell, R.W.*, Adolescent Risk Taking, Newbury Park, 107–130. – *Maddala, G.S.*: Limited-dependent and Qualitative Variables in Econometrics, Cambridge 1983. – *Mayer, H.*: Das Strafrecht des Deutschen Volkes, Stuttgart 1936. – *McFatter, R.M.*: Ungleichheit in der Strafzumessung und Zweck der Strafe, in: *Pfeiffer, C. & Oswald, M.E.*, Strafzumessung. Empirische Forschung und Strafrechtsdogmatik im Dialog, Stuttgart 1989, 183–195. – *McKelvey, R.D. & Zavoina, W.*: A Statistical Model for the Analysis of Ordinal Level Dependent Variables. Journal of Mathematical Sociology 1975, 103–120. – *Oswald, M.*: Wie reagieren Richter auf Befunde der Strafzumessungsforschung?, in: *Greive, W.*, Mehr Transparenz in der Strafjustiz. Der Prozeß des Strafens im Blickpunkt kriminologischer Forschung, Loccum 1991, 160–167. – *Oswald, M.E.*: Psychologie des richterlichen Strafens, Stuttgart 1994. – *Pfeiffer, C. & Savelsberg, J.J.*: Regionale und altersgruppenbezogene Unterschiede der Strafzumessung, in: *Pfeiffer, C. & Oswald, M.E.*, Strafzumessung. Empirische Forschung und Strafrechtsdogmatik im Dialog, Stuttgart 1989, 17–41. – *Roxin, C.*: Zur jüngsten Diskussion über Schuld, Prävention und Verantwortlichkeit im Strafrecht, in: Festschrift für Bockelmann, Frankfurt 1978. – *Schiel, J.*: Unterschiede in der deutschen Strafrechtsprechung. Eine Untersuchung anhand von Urteilen der Schöffengerichte und Strafkammern in Koblenz, Frankfurt und München aus den Jahren 1962/63, Hamburg 1969. – *Schöch, H.*: Kriminologische Aspekte der Strafzumessungslehre, in: *Pfeiffer, C. & Oswald, M.E.*, Strafzumessung. Empirische Forschung und Strafrechtsdogmatik im Dialog, Stuttgart 1989, 132–134. – *Schumann, K.F.*: Positive Generalprävention. Ergebnisse und Chancen der Forschung, Heidelberg 1989. – *Schumann, K.F., Berlitz, C., Guth, H.W. & Kaulitzki, R.*: Jugendkriminalität und die Grenzen der Generalprävention. Neuwied 1987. – *Smaus, G.*: Das Strafrecht und die Kriminalität in der Alltagssprache der Bevölkerung, Opladen 1985. – *Stelzl, I.*: Fehler und Fallen in der Statistik, Bern 1982. – *Streng, F.*: Tiefenpsychologie und Generalprävention – ein Diskussionsbeitrag. Kriminologisches Journal 1987, 48–54. – *Tyler, T.R.*: Why People Obey the Law: Procedural Justice, Legitimacy, and Compliance, New Haven 1990. – *Uhlig, S.*: Das Bundeszentralregister und andere Zentralregister. Aufgaben und Nutzungsmöglichkeiten, in: *Jehle, J.M.*, Datensammlung und Akten in der Strafrechtspflege, Wiesbaden 1989, 45–66. – *Woerner, O.*: Die Frage der Gleichmäßigkeit der Strafzumessung im Deutschen Reich, München 1907. – *Zamble, E. & Kalm, K.L.*: General and Specific Measures of Public Attitudes Toward Sentencing. Canadian Journal of Behavioural Science 1990.

(Ansch. d. Verf.: Dr. Jörg Hupfeld, Universität Bern, Institut für Psychologie, Lehrstuhl für Sozial- und Rechtspsychologie, Muesmattstraße 45, CH-3000 Bern 9)