

# **Einleitung**

## **Natur und Kultur**

### **1 Die Macht der Stereotypen**

#### **„Baby Sexing“**

Seit etlichen Jahren halte ich in regelmäßigen Turnus eine Vorlesung über die Entwicklung von Geschlechtsunterschieden, die damit zu beginnen pflegt, dass dem Auditorium eine Reihe von Videofilmen mit Kindern im Alter von eineinhalb bis zwei Jahren vorgeführt wird. Die Studierenden sollen bei jedem Kind raten, ob es sich um einen Jungen oder ein Mädchen handelt. Dies lässt sich bei dieser Altersklasse nicht einfach an Äußerlichkeiten feststellen, denn Mütter ziehen ihre kleinen Kinder längst nicht mehr geschlechtsrollenkonform an. So trifft man Jungen mit langen, wallenden Locken und goldenen Armbändchen ebenso wie Mädchen mit Kurzhaarfrisuren und in Höschen, beide sowieso noch mit Windelpaketen, die auch nicht gerade zwischen den Geschlechtern differenzieren. In Anlehnung an den angelsächsischen Sprachgebrauch, bei dem Geschlechtsbestimmung von Haustieren kurzerhand auf die dort üblich pragmatische Weise als „Sexing“ bezeichnet wird, bereicherte eine Studentin, die mir bei der Auswahl des Videomaterials zur Hand ging, unseren Laborjargon um den Begriff „Baby Sexing“. Die Studierenden sollen dabei ihren Eindruck zunächst spontan äußern und dann versuchen, sich Rechenschaft abzulegen, anhand welcher Merkmale sie zu ihrem Urteil kamen.

Da ich diese Übung nun schon öfter durchgeführt habe, geben mir einige Merkwürdigkeiten zu denken, die dabei regelmäßig aufraten.

Über die Jahre hinweg zeigte sich nämlich, dass die Beurteilungen zwar häufig danebengingen, dass die verschiedenen Gruppen aber in hohem Maße in der Wahl der Merkmale übereinstimmten, mit der sie ihre Zuordnungen begründeten, sodass ich dazu übergehen konnte, eine vorbereitete Folie mit einer Liste von Eigenschaften für den Hellraumprojektor mitzubringen, die mit jeweils nur geringfügigen Modifikationen dann die Grundlage für die Diskussion abgab. Die nachfolgende Tabelle informiert über die Einträge dieser Liste, soweit sie das Verhalten betreffen. Physiognomische Kriterien (Kopfform, Körperbau, Haltung und Bewegung, Zierlichkeit, Weichheit etc.) wurden auch genannt, sollen hier aber außer Betracht bleiben. Interessant ist auch, dass gewisse Merkmale sowohl für die Beurteilung als Junge als auch als Mädchen den Grund lieferten. So wurde z. B. einerseits der Eindruck, das Kind sei konzentriert und ausdauernd bei der Sache, als Indiz für Weiblichkeit genannt, dann aber wurde auch wieder geglaubt, Jungen an ihrer Beharrlichkeit zu erkennen.

Junge	Mädchen
demonstriert Stärke	behutsam
Imponiergehabe	vorsichtig
laut	zurückhaltend
angeberisch	geduldig
„Pascha“	sorgfältig
dominant	zaghaft
aggressiv	scheu
selbstständig	mutterorientiert
bestimmt	angewiesen auf
zielsicher	Ermutigung
initiativ	Hilfe
erfinderisch	Schutz
experimentierfreudig	kommunikativ
explorativ	kontaktbereit
ablenkbar	spricht viel
mehr am Spiel und an Objekten als an Personen interessiert	interaktives Spiel expressive Mimik emotional engagiert zeigt Mitgefühl schamhaft kokett

An diesem Ergebnis ist zunächst einmal nichts besonders Bemerkenswertes. Interessanter erscheint mir dagegen eine Verhaltenseigentümlichkeit, die sich während der Diskussion fast schon voraussagbar einstellte. Anfänglich gaben die Studierenden einige ganz unbefangene Urteile ab, etwa von der Art „Es ist ein Junge, denn er ist so draufgängerisch“, oder „Es ein Mädchen, denn es ist mehr am Kontakt als am Spielzeug interessiert“. Das rief die ersten verlegenen Lacher hervor und der Strom der Kommentare wurde daraufhin zähflüssiger. Inhaltlich tendierte man nun dazu, sich auf Merkmale der äußeren Erscheinung wie Haarlänge etc. zu beschränken, obwohl ich vorher betont hatte, dass diese irrelevant seien. Irgendetwas wurde den Studenten zunehmend peinlich, sie begannen sich offensichtlich klarzumachen, auf was sie sich bei ihrer Urteilsbildung einließen, und verloren ihre Unbefangenheit.

Nun könnte das daran liegen, dass die meisten Studierenden in einer solchen Einstufung nicht geübt sind, denn eineinhalb bis zweijährige Kinder gehören im Allgemeinen nicht zu ihrem täglichen Umgang. Es wäre also denkbar, dass sie sich überfordert fühlten. Andererseits hing aber nichts für sie davon ab, ob sie richtig oder falsch urteilten. Die Ursachen lagen also wohl auf einem anderen Sektor. Wenn man sie direkt auf ihre Befangenheit hin ansprach, dann stellte sich regelmäßig heraus, dass sie sich zunehmend der Tatsache bewusst wurden, in ihren Äußerungen Inhalte zum Ausdruck zu bringen, die sich mit den gängigen *Geschlechtsrollenstereotypen* deckten, wie sie von Williams und Best kulturredigend festgestellt wurden: Männer gelten als durchsetzungsstärker, aggressiver, selbstbewusster und risikobereiter, Frauen als sensibler, fürsorglicher, vorsichtiger und nachgiebiger<sup>1</sup>.

<sup>1</sup> Williams & Best, 1990; Williams et al., 1999

## Stereotypen

Unter Stereotypen versteht man soziale Urteile, die eigentlich zutreffender als *Vorurteile* zu kennzeichnen sind, da sie die Tendenz haben, Personen grob vereinfachend und ohne Rücksicht auf ihre Individualität zu etikettieren. Solche Überzeugungen werden von einem großen Teil der Bevölkerung geteilt; sie bestimmen die Einstellung zu eigenen und zu fremden Gruppen und eben auch die Rollenerwartungen an die Geschlechter. Da den Studierenden in diversen Vorlesungen die Botschaft vermittelt wird, eine Urteilsbildung auf der Basis von Stereotypen sei fragwürdig, ja sogar verwerflich, gerieten sie beim „Baby Sexing“ offensichtlich in einen Konflikt, sobald sich zeigte, dass sie sich doch von solchen Überzeugungen leiten ließen – und dies dann gar noch öffentlich im Hörsaal. Eine Studentin brachte ihr Unbehagen einmal auf den Punkt, indem sie sich gerade heraus weigerte, überhaupt bei dieser Übung mitzumachen, mit der Begründung, ein solches Vorgehen zementiere die Diskriminierung von Frauen, man müsse doch endlich von diesen Stereotypen wegkommen und das Denken in Unterschieden überwinden.

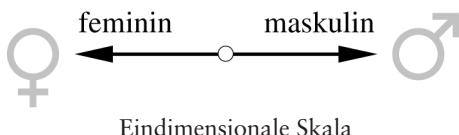
Nun trifft es ohne Zweifel zu, dass Stereotype die Eigenschaft haben, zu übertrieben und über einen Kamm zu scheren. Tatsächlich gingen die Studierenden bei ihrer Beurteilung teilweise von recht groben Klischees aus und trafen damit dann auch tüchtig daneben. Die krassesten Fehleinschätzungen kamen dort vor, wo allein schon das bloße Auftreten eines als „typisch“ männlich oder weiblich geltenden Verhaltens für die Zuordnung ausschlaggebend war, ohne dass dabei aber berücksichtigt wurde, *in welcher Art* dieses Verhalten ablief. Fußballspiel z. B. wurde wie selbstverständlich als Kennzeichen für Jungen gewertet, so als wäre es überhaupt nicht denkbar, dass auch einmal ein Mädchen Spaß daran findet.

Überhaupt wurden die Begründungen häufig so formuliert, als käme das betreffende Merkmal *ausschließlich* einem Geschlecht zu. Wurde also beispielsweise „Kontaktfähigkeit“ als Indiz für Weiblichkeit angeführt, so klang das zuweilen so, als sei das männliche Geschlecht in dieser Hinsicht schlechterdings inkompotent. Stereotype Urteile sind eben nicht nur pauschal, sondern auch ausgrenzend; wird eine Eigenschaft dem einen Geschlecht zugewiesen, so wird sie dem anderen ebenso unterschiedslos abgesprochen.

## Androgynie

Auch in den ersten Versuchen, Geschlechtsunterschieden wissenschaftlich durch Fragebogenerhebungen auf die Spur zu kommen<sup>2</sup>, war man davon ausgegangen, dass maskuline und feminine Merkmale bipolar auf einer einzigen Dimension liegen und sich daher gegenseitig ausschließen. Je weniger weibliche Eigenschaften eine Person aufweist, umso mehr männliche sollte sie auf sich vereinen und umgekehrt.

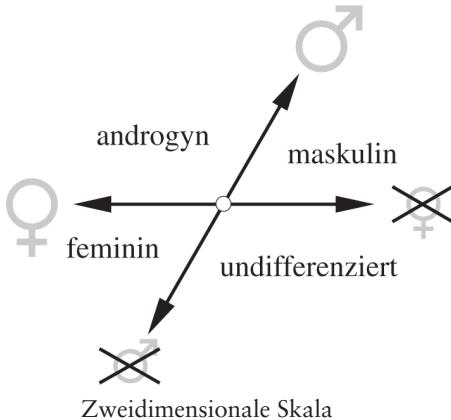
Diese Position wurde aber auf die Dauer unhaltbar. Es fanden sich zwar Merkmale, die ein Großteil der Frauen als typisch für sich reklamierten; die Männer fühlten sich dann aber keineswegs durch das Gegenteil dieser Eigenschaft adäquat charakterisiert, sondern besetzten mehr oder minder unbefangen die gesamte Bandbreite von „trifft auf mich genau zu“ bis „trifft überhaupt nicht zu“,



<sup>2</sup> Beispielsweise im Terman-Miles-Test of Masculinity and Femininity, 1936

mit einer natürlichen Häufung im Bereich mittlerer Ausprägung. Entsprechendes gilt für typisch männliche Merkmale.

Daraus ließ sich nur der Schluss ziehen, dass Maskulinität und Feminität unabhängige Merkmalsgruppen sind, die sich nicht strikt widersprechen, sondern kombiniert werden können. Statt einer einzigen hat man demnach zwei Achsen zu unterscheiden, deren eine durch die Pole „männlich-unmännlich“, die andere durch „weiblich-unweiblich“ zu charakterisieren sind. Ursprünglich wurde angenommen, dass diese Achsen völlig unabhängig voneinander sind, was dann graphisch durch ihre orthogonale Anordnung auszudrücken wäre. Inzwischen hat sich gezeigt, dass zwar keine Identität, aber wohl doch eine gewisse Verwandtschaft zwischen weiblich und unmännlich einerseits, männlich und unweiblich andererseits besteht, sodass die Achsen heute meist in obliquier Anordnung dargestellt werden (siehe Abbildung)<sup>3</sup>.



Zweidimensionale Skala

Auch in diesem Falle hat man statt einer bipolaren eine Vier-Felder-Anordnung. Personen rangieren nur noch dann als „feminin“, wenn sie viele feminine Merkmale und zugleich niedrige Werte in Maskulinität aufweisen. Entsprechendes gilt spiegelbildlich für die Zuweisung zur Gruppe „maskulin“. Zusätzlich gibt es nun noch Personen, die sich in Bezug auf sowohl Maskulinität als auch Feminität unternormal einschätzen. Diese werden meist, nicht sehr treffend, als „undifferenziert“ bezeichnet; niemand interessiert sich sonderlich für sie. Das Gegenteil gilt für die letzte Gruppe, die laut Selbstbeurteilung sowohl typisch männliche als auch typisch weibliche Merkmale in sich vereint. Diese wird nach einem Vorschlag von Sandra Bem als „androgyne“ bezeichnet<sup>4</sup>, und auf sie richtete sich ein gut Teil Emanzipationshoffnung, sind die betreffenden Personen doch gewissermaßen der Vortrupp auf dem Weg zur endgültigen Befreiung vom Diktat der Geschlechtsrollen.

Bem arbeitete mit einem Geschlechtsrolleninventar, das jeweils 20 eher männliche, eher weibliche und eher geschlechtsneutrale Merkmale enthielt. Die Probanden hatten sich selbst auf einer siebenstufigen Skala einzuschätzen, die von „nicht oder nahezu nie zutreffend“ bis zu „immer oder fast immer zutreffend“ reichte. Spence<sup>5</sup> konzipierten ungefähr um die gleiche Zeit ebenfalls einen Fragebogen zur Messung geschlechtsspezifischer Merkmale. Er unterschied nur zwei Skalen, die er mit den Symbolen F und M belegte. Zusätzlich versuchte er die beiden Klassen auch noch inhaltlich zu identifizieren, und zwar die maskuline Skala durch das Stichwort „Instrumentalität“, die feminine durch „Expressivität“. Diese Charakterisierung entspricht aber wohl mehr dem Bedürfnis, eine einigermaßen politisch korrekte, d.h. wertfreie, Etikettierung zu finden, als dass sie wirklich als inhaltlich adäquat überzeugt.

<sup>3</sup> Reinisch et al., 1991

<sup>4</sup> Bem, 1974

<sup>5</sup> Spence et al., 1974

## Beispiele aus dem Geschlechtsrolleninventar von Bem

<i>ehler zu Männern passend</i>	<i>zu beiden Geschlechtern passend</i>	<i>ehler zu Frauen passend</i>
aggressiv	launisch	liebevoll
kompetitiv	freundlich	leidenschaftlich
dominant	theatralisch	warm
ehrgeizig	verlässlich	verständnisvoll
kraftvoll	ernsthaft	sanft
unabhängig	etc.	nachgiebig
risikobereit		scheu
etc.		etc.

## Beispiele aus dem Geschlechtsrolleninventar von Spence

<i>M-Skala (Instrumentalität)</i>	<i>F-Skala (Expressivität)</i>
aktiv	einfühlend
Druck standhaltend	auf andere eingehend
konkurrierend	freundlich
entscheidungsfähig	gefährsbetont
nicht leicht aufgabend	herzlich
selbstsicher	hilfreich
sich überlegen	verständnisvoll
fühlend	etc.
unabhängig	
etc.	

Die Bedeutung dieser Studien wird vornehmlich darin gesehen, dass sie durch einen eleganten Kunstgriff erlauben, den Kuchen gleichsam zu essen und zu behalten: Auf der einen Seite wird die im Volksmund unausrottbar verwurzelte Unterscheidung „typisch weiblicher“ von „typisch männlichen“ Eigenschaften aufgegriffen und beibehalten, auf der anderen aber von der schicksalhaften Bindung an das biologische Geschlecht gelöst. Wenn eine Person von Genetik und Anatomie her eine Frau ist, braucht sie deshalb noch längst nicht auch psychologisch feminin zu sein; ihr steht das ganze allgemeinmenschliche Wertspektrum offen. Wieso jene Merkmale aber überhaupt noch als „feminin“ und „maskulin“ apostrophiert werden, wieso man nicht wirklich konsequent nur noch etwa von „Instrumentalität“ und „Expressivität“ spricht, entzieht sich dann leicht der weiteren Reflexion.

**Die Studie von Maccoby und Jacklin**

Wie steht es nun aber wirklich mit der Verteilung „femininer“ und „maskuliner“ Eigenschaften auf die beiden Personengruppen, die wir im vorwissenschaftlichen Sprachgebrauch als „Mädchen“ und „Jungen“, als „Frauen“ und Männer“ zu unterscheiden pflegen? Werden die in den Stereotypen zum Ausdruck gebrachten

Differenzen wirklich nur herbeigeredet, sodass ihre terminologische Abkoppelung vom biologischen Geschlecht oder – wie im Falle der oben erwähnten Studentin – die strikte Weigerung, sie überhaupt beim Namen zu nennen, schon ausreichen könnte, um sie aus der Welt zu schaffen? Sind sie einfach nichts anderes als Vorurteile, die es endlich auszuräumen gilt?

Viele sind noch heute davon überzeugt, dass das in der Tat so ist. Die Annahme, dass die Geschlechter sich in Wirklichkeit überhaupt nicht unterscheiden, wird so selbstverständlich für zutreffend gehalten, dass man es gar nicht für notwendig hält, ihre Richtigkeit nachzuweisen. Sie wird vielmehr als das behandelt, was die Statistiker eine *Nullhypothese* nennen (siehe Kasten).

### Nullhypotesen

Wenn man dreimal hintereinander eine Sechs würfelt, muss das nicht heißen, dass der Würfel präpariert ist. Wenn also ein Forscher glaubt, irgendeinen interessanten, dem bisher verbindlichen Wissensstand widersprechenden Zusammenhang entdeckt zu haben, so muss er zunächst statistisch prüfen, ob der beobachtete Effekt nicht auch rein zufällig hätte zustande kommen können. Er muss, wie man es ausdrückt, eine *Arbeitshypothese* („der Würfel ist präpariert“) gegen die *Nullhypothese* („Der Würfel ist normal“) prüfen. Das Verfahren ist asymmetrisch: Die Arbeitshypothese muss man *sichern*, die Nullhypothese *widerlegen*. Die Arbeitshypothese trägt also die Beweislast.

Das bedeutet: Wenn jemand behauptet, Unterschiede im Verhalten und Erleben von Frauen und Männern existierten nur im Volksglauben oder gar in der Fantasie von Ideologen und Chauvinisten, während in Wirklichkeit doch alle Menschen gleich angelegt seien, dann darf er dies solange tun, bis irgendwer ihn unausweichlich widerlegt, während die These, Frauen seien anders als Männer, und dies womöglich auch noch von Natur aus, die volle Beweislast zu tragen hat. Offenbar wird es als „sparsamer“ empfunden, Geschlechtsunterschiede zu leugnen als sie zu akzeptieren. Die Berechtigung zu dieser asymmetrischen Betrachtungsweise wird freilich so gut wie nie reflektiert.

Eleanor Maccoby und Carolin Jacklin von der Stanford University, die gegenwärtig immer noch zu den prominentesten Gewährsleuten in der Frage der Geschlechtsunterschiede zählen, gingen in ihrem Buch „The Psychology of sex differences“ ebenfalls von der gerade formulierten Nullhypothese aus und prüften nach, ob die empirische Befundlage an irgendwelchen Stellen dazu nötigen würde, sie zu verwerfen. Sie kamen dabei zu dem Schluss, dass „viele populäre Überzeugungen über die psychologische Eigenart der beiden Geschlechter erwiesenmaßen eine geringe oder überhaupt keine Grundlage haben. Die Ursache, warum solche Mythen dennoch aufrechterhalten werden, liegt darin, dass Stereotype eine so machtvolle Wirkung haben“.<sup>6</sup>

Die beiden Autorinnen haben etwa 1.600 empirische Arbeiten aus allen erdenklichen Gebieten der Psychologie daraufhin analysiert, ob männliche und weibliche Versuchspersonen dabei irgendwie abweichende Resultate erzielt hatten. Sie kamen zu dem Ergebnis, dass Unterschiede in lediglich vier Bereichen angenommen werden müssten. Drei davon betreffen den kognitiven Stil:

Männer hätten bessere Fähigkeiten im mathematischen Denken und in der visuellräumlichen Vorstellung, während das weibliche Geschlecht eher sprachbegabt sei.

---

<sup>6</sup> Maccoby & Jacklin, 1974, S. 355, Übersetzung von D.B.-K.

Auf dem emotionalen Sektor gäbe es überhaupt nur einen belegbaren Unterschied, nämlich in der Aggressivität. Alle übrigen üblicherweise postulierten Unterschiede dagegen seien ohne empirische Basis.

Als Resultat ihrer Studie sahen sich Maccoby und Jacklin in ihrer Vermutung bestätigt, Stereotypen seien nicht viel mehr als willkürliche Setzungen der Gesellschaft, die ihre Durchschlagkraft lediglich der Tatsache verdanken, dass sie von Generation zu Generation weitergegeben und insbesondere von den Männern bereitwillig immer wieder aufgegriffen werden, um die eigene Vorherrschaft zu sichern. Als engagierte Feministinnen folgerten die Autorinnen weiter, ein Abbau der Diskriminierung sei am ehesten dadurch einzuleiten, dass man die Stereotypen als das entlarve, was sie eigentlich sind, nämlich „Mythen“ ohne Wahrheitsgehalt.

Das passt zunächst gut zu der Feststellung, dass die gängigen Stereotypen wenig geeignet waren, meinen Studierenden zu wirklich qualifizierten Urteilen zu verhelfen. Aber ganz so einfach ist es auch wieder nicht.

## Feinschlägige Nuancen

Das Bild sieht nämlich anders aus, wenn man einige Zusatzerfahrungen mitberücksichtigt, die sich in unserer Arbeit inzwischen eingestellt haben. So zeigt sich etwa, dass die Trefferquote erheblich ansteigt, wenn die Beurteiler über mehr Erfahrung mit der betreffenden Altersgruppe verfügen.

In einer Studie an unserem Institut hatte eine Studentin den Auftrag, die Verhaltensweisen 16-monatiger Kinder beiderlei Geschlechts, die auf Video aufgenommen waren, genau zu analysieren. Die Kinder trugen einheitlich neutrale Kleidung, und Namensnennungen sowie sonstige Hinweise auf das Geschlecht waren auf den Videobändern gelöscht. Die Studentin konnte also nicht wissen, ob es sich bei einem Kind um einen Jungen oder um ein Mädchen handelte; vorgefasste Meinungen über Geschlechtsunterschiede hätten die Auswertung daher nicht beeinflussen sollen.

Dennoch stellte sich heraus, dass es der Untersucherin gar nicht möglich war, *keine* Mutmaßungen über das Geschlecht der Kinder anzustellen. Wir ließen sie, eher der Vollständigkeit halber, nach Abschluss der Auswertung ihre diesbezüglichen Einschätzungen notieren. Und diese erwiesen sich dann ausnahmslos als zutreffend – die Studentin hatte sich nicht bei einem einzigen der 25 Kinder im Geschlecht getäuscht!

Es muss also bereits in diesem Alter Differenzierungen geben, die dem Experten eine zuverlässige Zuordnung ermöglichen, auch wenn die üblichen Stereotypen vielleicht nicht zutreffen. Bei den Verhaltensunterschieden, die sich dann bei der Auswertung dieser Studie objektiv herausstellten, handelte es sich weder um die von Maccoby und Jacklin konstatierten Abweichungen in den kognitiven Stilen – diese manifestieren sich erst viel später im Verhalten –, noch lag es daran, dass die Buben jetzt schon „aggressiver“ waren als die Mädchen. Es handelte sich vielmehr um feinschlägige Unterschiede in der Art und Weise, wie die Babys mit Spielsachen umgingen, wie sie diese manipulierten und untersuchten und wofür sie sich besonders interessierten. Auch in der Interaktion mit der Mutter und in der Reaktion auf eine kurzfristige Trennung von ihr unterschieden sich die Geschlechter.

Ich werde auf diese Ergebnisse noch genauer zu sprechen kommen und habe hier nur auf sie verwiesen, um deutlich zu machen, dass Maccoby und Jacklin in ihrer zugegebenermaßen anerkennenswerten Absicht, Vorurteile abzubauen, offensichtlich über das Ziel hinausgeschossen sind.

## Methodische Probleme

Zu derselben Ansicht war Jeanne Block, eine Kollegin der genannten Autorinnen, bereits im Jahre 1976 gekommen, als sie das Buch von Maccoby und Jacklin bald nach seinem Erscheinen einer Revision unterzog. Sie analysierte das darin zugrunde gelegte Material ein zweites Mal und kam zu recht deutlich abweichen den Befunden<sup>7</sup>.

Vor allem machte sie auf methodische Schwachpunkte des Buches aufmerksam. Ein großer Prozentsatz der von Maccoby und Jacklin berücksichtigten Arbeiten hatte gar nicht primär die Frage nach Geschlechterunterschieden zum Thema gehabt. Die Autoren waren im Gegenteil sogar ausdrücklich daran interessiert gewesen, möglichst nicht auf solche zu stoßen, und zwar gar nicht unbedingt aus Gründen der politischen Korrektheit, sondern einfach deshalb, weil diese die Stichprobengröße halbiert und damit ihre eigentlich erwarteten Ergebnisse weniger signifikant gemacht hätten.

Signifikanz ist ein mathematischer Ausdruck für die Sicherheit, mit der man ein zufälliges Zustandekommen des Ergebnisses einer empirischen Untersuchung ausschließen kann. Sie hängt wesentlich von der Größe der Stichprobe ab (siehe Kasten).

In den von Maccoby und Jacklin untersuchten Fällen waren nun oft genug durchaus Geschlechtsunterschiede aufgetreten, diese hatten aber wegen der zu kleinen Stichprobengröße keine ausreichende Signifikanz erreicht. Im Text der Veröffentlichung wurde dies dann als „keine Unterschiede“ verzeichnet. Bloße Tendenzen blieben unerwähnt. Wenn nun aber in mehreren Untersuchungen, die unabhängig voneinander demselben Thema gewidmet sind, überall dieselbe Tendenz auftauchte, dann beginnen sich die Befunde gegenseitig zu stützen und können gemeinsam durchaus ein passables Signifikanzniveau erreichen.

Gravierender ist allerdings Blocks Feststellung, dass Maccoby und Jacklin eine Reihe von Untersuchungen zwar angeführt, bei der statistischen Gesamtevaluation dann aber überhaupt nicht berücksichtigt haben; man hatte sie offensichtlich vergessen. Dabei handelte es sich nun ausgerechnet um Befunde, die weitere Geschlechtsunterschiede gerade auch im motivational-emotionalen Bereich eindeutig belegten, nämlich höhere *Explorativität*, *dominanteres Verhalten* und *stärkeres Selbstvertrauen* bei Jungen, und erhöhte *Ängstlichkeit*, *Neigung zu Gehorsam* und *größere Nähe im persönlichen Kontakt* bei Mädchen.

### Signifikanz

Wenn in einer empirischen Untersuchung, sagen wir, zwölf von 20 Jungen, aber nur neun von 20 Mädchen eine bestimmte Leistung erbrachten, so erlaubt dieses Ergebnis für sich noch nicht die Behauptung, es bestehe ein Geschlechtsunterschied. Es könnte immer noch leicht durch Zufall zustandegekommen sein; es ist nicht „signifikant“, sondern höchstens „tendentiell“, was soviel heißt wie, dass man die Untersuchung noch einmal mit einer größeren Stichprobe wiederholen sollte, bevor man ihr wirklich trauen kann. Wenn jedoch von jedem Geschlecht die zehnfache Zahl, also je 200 Kinder, an dem Versuch teilgenommen und dann 120 Jungen und 90 Mädchen positiv abgeschnitten hätten, so wäre ein zufälliges Zustandekommen des Ergebnisses so unwahrscheinlich, dass man es ernst nehmen müsste.

---

<sup>7</sup> Block, 1976

Man wird einer erwiesenermaßen integren Wissenschaftlerin vom Format Eleanor Maccobys sicher keine bewusste Täuschungsabsicht unterstellen können. Arbeiten von diesem Umfang lassen sich aber nun einmal nicht ohne einen Stab von Mitarbeitern bewältigen, und da können sich, wie man auch hier wieder sieht, allerlei Pannen einschleichen.

Block zieht ihrerseits aus der intensiven Beschäftigung mit dem Datenmaterial das Fazit, dass die Stereotypen doch einen Wahrheitsgehalt haben könnten. Sie schreibt:

„Meiner Meinung nach kann man den ‚Volksglauben‘ und die ‚Mythen‘ über Geschlechtsunterschiede, auf die sich Maccoby und Jacklin beziehen, nicht einfach dadurch wegerklären, dass man nur auf die Verbreitung und Überzeugungsmacht von Stereotypen verweist. ... Selbst wenn Stereotypen nur ‚Mythen‘ zur Grundlage haben sollten, so könnten darin dennoch auch gewisse kulturell scharfsichtig wahrgenommene und vielfach bestätigte Wahrheiten zum Ausdruck kommen.“<sup>8</sup>

Die Kontroverse hatte in jüngster Zeit noch ein bedenkenswertes Nachspiel. In einer Buchveröffentlichung von 1998 hat sich Eleanor Maccoby, inzwischen emeritiert, noch einmal ausführlich auf der Basis der neuesten Befundlage mit dem Thema befasst und ihre Annahmen von 1974 erheblich revidiert. Von ihrer damals vertretenen Ansicht, die Geschlechter würden sich nur unwesentlich unterscheiden, ist dabei nicht mehr viel übriggeblieben. Stattdessen befasst sie sich nun eingehend mit den Unterschieden, die sie vor allem in der Gruppenstruktur, in den Interaktionsmustern, in Konfliktbewältigungsstrategien und im Spielverhalten lokalisiert. Sie sieht in ihnen die Erklärung für die durchgängig beobachtbare und im frühen Alter bereits einsetzende Geschlechtertrennung, einem Phänomen, dem sie weite Partien ihres Buches widmet<sup>9</sup>. Was die Ursachenanalyse betrifft, so problematisiert sie die Annahme einer ausschließlichen Sozialisiertheit von Geschlechtsunterschieden und zieht neben kognitiven Faktoren auch eine biologische Mitverursachung ernsthaft in Erwägung, wobei man allerdings ihrer Diktion anmerkt, wie schwer ihr dieser Schritt fällt. Umso bemerkenswerter sind ihr Versuch eines Neuzugangs und die Erweiterung ihrer Perspektive.

Mit diesem Perspektivenwechsel steht Maccoby übrigens nicht allein. Auch andere, zum Teil durchaus feministisch orientierte Forschende, die ursprünglich davon ausgegangen waren, die Annahme geschlechtstypischer Verhaltensunterschiede hätte keine reale Basis, haben inzwischen ihre Einstellung revidiert. Alice Eagly, eine der prominentesten amerikanischen Genderforscherinnen, kennzeichnet 1995 die in den 70er Jahren auf feministischer Seite dominierende Überzeugung, Geschlechtsunterschiede seien so unbedeutsam, dass man sie vergessen könne, als unzutreffend und kontraproduktiv<sup>10</sup>. John Williams und Deborah Best, deren Forschungsabsicht es ursprünglich gewesen war, die Nicht-Existenz von Geschlechtsunterschieden nachzuweisen, gelangten aufgrund einer kulturvergleichenden Studie über Geschlechtsstereotype zu der Einsicht, dass diese nichtwegzuleugnen seien, und diskutieren in diesem Zusammenhang durchaus auch biologische Ursachen<sup>11</sup>. Diane Halpern, die davon ausgegangen war, dass kognitive Geschlechtsunterschiede nicht existierten und diesbezügliche Berichte auf Untersuchungartefakten, Fehlern und Vorurteilen beruhten, kam nach Sichtung der gesamten einschlägigen Litera-

<sup>8</sup> Block, 1976, S. 285 und 295

<sup>9</sup> Maccoby, 2000

<sup>10</sup> Eagly, 1995

<sup>11</sup> Williams & Best, 1990; siehe auch Lippa, 2010

tur zu dem Ergebnis, „es gibt tatsächlich und in manchen Fällen sogar erhebliche Geschlechtsunterschiede in einigen kognitiven Fähigkeiten. Sozialisationspraktiken sind dafür zweifelsohne wichtig, aber es gibt auch gute Hinweise, dass biologische Geschlechtsunterschiede eine Rolle spielen.“<sup>12</sup>

Solche Einsichten haben inzwischen in Form populärer Bestseller auch die breite Öffentlichkeit erreicht und um Erkenntnisse des Kalibers bereichert, Frauen „können nicht einparken“ und Männer seien „vom Mars“. Nun erweisen sich diese Elaborate in ihrer kruden und wenig fundierten Argumentation aber als ausgesprochen kontraproduktiv. Dies nicht nur, weil sie unzutreffende Vorstellungen über Ausmaß und Ursachen von Geschlechtsunterschieden in ihrer Leserschaft hervorrufen, sondern vor allem auch, weil sie durch undifferenzierte und einseitige Annahmen über die Wirkung der Veranlagung antibiologischen Vorurteilen erneut Nahrung geben. So beklagt Sylvia Jahnke-Klein in einem Artikel im jüngst erschienenen Handbuch zur Mädchenpädagogik, in dem es um Mädchen und Naturwissenschaften geht, dass *biologische Begründungsansätze* (kursiv von D.B.-K.) in der öffentlichen Diskussion immer wieder viel Aufmerksamkeit erfahren. Und dann führt sie als einziges Beispiel eben das Buch von Pearse an („Warum Männer nicht zuhören und Frauen schlecht einparken“), das sie offensichtlich als repräsentativ für den biologischen Ansatz erachtet, denn sie kommt ohne Umschweife zur einer pauschalen Aburteilung: „Wissenschaftlich haltbar ist meines Erachtens keiner dieser biologischen Erklärungsansätze“<sup>13</sup> (man beachte die Mehrzahl!). Wenn sie anschließend dem „biologischen Begründungsmuster“ einen „gravierenden Denkfehler“ unterstellt, weil es der überlappenden Verteilung von Fähigkeitskurven zwischen den Geschlechtern angeblich nicht Rechnung trage, dann beruht der gravierende Denkfehler ausschließlich darauf, dass sie sich offensichtlich nie ernsthaft mit seriösen biologischen Ansätzen befasst hat.

Bei dieser Grundstimmung verwundert es nicht, dass das Pendel neuerdings wieder in die andere Richtung ausschlägt. Eine Studie von Janet Hyde trägt denn auch den bezeichnenden Titel „Die Hypothese der *Geschlechterähnlichkeit*“<sup>14</sup>. Hydes Zusammenstellung von 46 Übersichtsarbeiten zu verschiedenen psychologischen Bereichen läuft auf die These hinaus, dass die meisten in der Populärliteratur behaupteten Geschlechtsunterschiede keine oder nur eine schwache empirische Basis hätten. So haben vor ihrem kritischen Blick lediglich Unterschiede in der Aggression, der Sexualität (Männer masturbieren mehr!) und der Motorik Bestand. Wenn man ihre Übersichtstabellen allerdings genauer anschaut, stellt man fest, dass dort noch etliche weitere Unterschiede wie beispielsweise in Fürsorglichkeit, Durchsetzungsorientiertheit und räumlichem Vorstellungsvermögen angeführt sind, auf die dann aber in der zusammenfassenden Diskussion nicht Bezug genommen wird, obwohl sie ebenfalls ein deutliches statistisches Gewicht aufweisen. Hyde hat Untersuchungen mit eindeutigen Unterschieden bei ihrer Bilanz offensichtlich nicht berücksichtigt, wenn nicht *sämtliche* Studien in diesem Bereich zu einem vergleichbaren Ergebnis kamen. Dies trägt meines Erachtens der Tatsache nicht genügend Rechnung, dass Ergebnisse je nach der verwendeten Methodik variieren können. Es macht zum Beispiel einen erheblichen Unterschied, ob ein Befund durch direkte Beobachtung oder aber nur über eine Befragung erhoben wurde, und durch den Einsatz bestimmter Tests können unterschiedliche Aspekte einer Fähigkeit angesprochen werden. In solchen Fällen ergibt die Mittelung der Er-

---

<sup>12</sup> Halpern, 2000 (Zitat aus Rhoads, 2004); Halpern et al., 2007

<sup>13</sup> Jahnke-Klein, 2010, S. 248

<sup>14</sup> Hyde, 2005