Entstehungsbedingungen von Antwortverzerrungen durch soziale Erwünschtheit

Ein Vergleich der Prognosen der Rational-Choice Theorie und des Modells der Frame-Selektion

Determinants for Respondents' Susceptibility to Social Desirability Bias

A Comparison of Predictions from Rational Choice Theory and the Model of Frame-Selection

Volker Stocké*

Universität Mannheim, Sonderforschungsbereich 504, L13, 15, D-68131 Mannheim

Zusammenfassung: Die Rational-Choice Theorie und deren Generalisierung im Modell der Frame Selektion beanspruchen beide eine vollständige und die Interdependenz der Entstehungsfaktoren berücksichtigende Erklärung des sozial erwünschten Antwortverhaltens in Umfragen. Die hierbei vorhergesagten Erklärungsfaktoren wurden bisher nur unvollständig und oft mit widersprüchlichen Ergebnissen empirisch untersucht. Der vorliegende Beitrag überprüft daher theorievergleichend zentrale Vorhersagen der beiden Ansätze am Antwortverhalten über Einstellungen zu Ausländern. Die Ergebnisse bestätigen im ersten Schritt die Vorhersagen der „reinen" Rational-Choice Theorie, wonach mit sozial erwünschtem Antwortverhalten nur dann zu rechnen ist, wenn die motivationalen, kognitiven und sozialen Vorbedingungen gleichzeitig erfüllt sind. Diese bereits sehr differenzierte Prognose erweist sich allerdings bei der Analyse der im Modell der Frame-Selektion zusätzlich angenommenen kulturellen Faktoren als unvollständig. Mit diesen Determinanten einer „kooperativen" Rahmung der Interviews lässt sich die vollständige Ausschaltung instrumentell rationaler Anreize durch soziale Erwünschtheit vorhersagen. Die Gültigkeit der Rational-Choice Theorie ist demnach auf wichtige, aber eben nur spezielle und im Rahmen des Modells der Frame-Selektion vorhersagbare Bedingungen beschränkt.

1. Einleitung

Das sozialwissenschaftliche Interview ist bekanntlich kein neutraler sozialer Prozess, und eines seiner Hauptprobleme ist die Reaktivität der Befragten auf andere Umstände als auf den bloßen Fragestimulus. Hierbei ist besonders die Abgabe sozial erwünschter Antworten zu nennen, durch die die Gültigkeit der Antworten ernsthaft gefährdet werden kann. Dies trifft speziell für normativ geprägte Themen zu, bei denen die öffentliche Äußerung bestimmter „wahrer Werte" zu missbilligenden Reaktionen führen kann. Hierzu gehören Fragen zum sexuellen Verhalten oder dem Drogenkonsum der Befragten. Vor allem bei der Erfassung ethnischer Vorurteile muss mit einer Diskrepanz zwischen den „wahren" Einstellungen und den durch gesellschaftliche Normen als legitim definierten Ansichten gerechnet werden.

Da es sich beim „social desirability bias" (SD-Bias) um systematische Antwortverzerrungen in eine bestimmte Richtung handelt, kann vermutet werden, dass zumindest einige Ergebnisse über die Verteilung von Verhaltensweisen und Einstellungen verzerrt sind. Außerdem ist nicht ausgeschlossen, dass empirisch beobachtete Zusammenhänge zwischen dem betreffenden Erhebungsmerkmal und sozialstrukturellen Merkmalen der Befragten ein Artefakt gruppenspezifischer Unterschiede in der Stärke oder gar Richtung des SD-Bias sind.

Diese Antwortverzerrungen können nur während des Datenerhebungsprozesses vermieden oder nachträglich durch statistische Verfahren korrigiert werden, wenn ihre Stärke und inhaltliche Ausrichtung in Abhängigkeit gewisser Merkmale der Befragten oder der Interviewsituation genau vorhergesagt werden können. Das aber setzt eine explizite, möglichst vollständige und empirisch gut bestätigte Theorie zur Prognose des (sozial erwünschten) Befragtenverhaltens voraus. Es gibt inzwischen in der Tat eine Reihe von Studien, die Entstehungsbedingungen des SD-Bias untersuchen. Die dabei erzielten Ergebnisse zeigen jedoch ein teilweise sehr uneinheitliches und schwer interpretierbares Bild. Das liegt, so kann vermutet werden, an der Unvollständigkeit der dabei meist forschungsleitenden theoretischen Erklärungen des Befragtenverhaltens.

* Für wertvolle Hinweise zu einer früheren Version dieses Manuskriptes danke ich Stephan Ganter, drei anonymen Gutachtern und ganz besonders Hartmut Esser. Die vorliegende Arbeit wurde durch eine Förderung der Deutschen Forschungsgemeinschaft ermöglicht.

Hierbei werden die möglichen Determinanten des SD-Bias fast ausschließlich isoliert und ohne Berücksichtigung zentraler Moderatorvariablen für deren Erklärungskraft untersucht.

Der folgende Beitrag versucht deshalb eine möglichst explizite und vollständige Erklärung des SD-Bias am Beispiel der Einstellungen zu Ausländern. Die theoretische Grundlage ist zunächst die (einfache) Rational-Choice Theorie (RCT) des Befragtenverhaltens (vgl. hierzu: Esser 1991). Schon dabei wird klar, dass starke Anreize durch soziale Erwünschtheit und damit *systematische* SD-Effekte nur bei einer sehr speziellen Ausprägungskonstellation der in der bisherigen Forschung meist nur isoliert untersuchten Bestimmungsfaktoren erwartet werden können. Anschließend wird untersucht, unter welchen Bedingungen sich auch starke Anreize zum „Lügen" *nicht* auf die Antworten auswirken. Für bestimmte Bedingungen wird die Orientierung an einer „kooperativen" Befragtenrolle und eine Rahmung der Interviewsituation unter dem „Code" einer Antwort nach dem „wahren Wert" erwartet, sodass alle anderen sozialen Bezüge ausgeblendet werden. Das Modell der Frame-Selektion (MdFS) ist die theoretische Grundlage für die Prognose dieser Bedingungen. Hierbei wird die in der RCT angenommene Bedeutung instrumenteller Anreize um die Wirksamkeit schemageleiteter Orientierungen erweitert und so eine übergreifende Erklärung des (sozialen) Handelns vorgeschlagen (Esser 2001: 259ff.).

Der folgende Beitrag verfolgt zwei eng miteinander verbundene Ziele. Es soll *erstens* die Brauchbarkeit des MdFS im Vergleich zur „einfachen" RCT des (Befragten-) Verhaltens empirisch überprüft werden. Beide Ansätze bieten eine Reihe von Hypothesen darüber, wann genau und durch welche Moderatorvariablen vermittelt, sich einzelne Determinanten des SD-Bias tatsächlich auf das Antwortverhalten auswirken. Durch die Überprüfung dieser Hypothesen, und darin besteht das *zweite* Ziel der Untersuchung, wird die genaue Konstellation jener Randbedingungen angebbar, unter der die Gültigkeit von Umfragedaten durch einen SD-Bias eingeschränkt ist. Mit diesen Erkenntnissen sind gezielte Maßnahmen zur Ausschaltung des SD-Bias in der Interviewsituation oder nachträgliche statistische Kontrollprozeduren möglich.

Im folgenden *zweiten* Abschnitt werden zuerst die Prognosen der RCT über die Determinanten des SD-Bias und dann deren weitere Differenzierung durch das MdFS dargestellt. Im *dritten* Abschnitt wird die verfügbare empirische Evidenz für die Gültigkeit dieser Vorhersagen skizziert. Die hierbei identifizierten Widersprüche und offenen Fragen werden im *vierten* Abschnitt auf die Vernachlässigung der theoretisch erwartbaren gegenseitigen Abhängigkeit der einzelnen Erklärungsfaktoren zurückgeführt. Im *fünften*, empirischen Teil werden zuerst die Prognosen der RCT und dann deren Differenzierung im Rahmen des MdFS am Gegenstand der Erfassung von Einstellungen zu Ausländern überprüft. Der *sechste* Abschnitt diskutiert die Bedeutung der gewonnenen Erkenntnisse für die theoretische Erklärung des Befragtenverhaltens und deren Konsequenzen für die Umfragepraxis.

2. Theoretische Vorhersagen über die Entstehungsbedingungen des SD-Bias

Die RCT zur Erklärung des (sozial erwünschten) Befragtenverhaltens prognostiziert die Antworten in Umfragen als das Ergebnis einer instrumentell rationalen Antwortselektion unter Maximierung der sozialen Anerkennung (Esser 1991, Stocké 2001). Der Ansatz ist wegen dieser engen Motivationsgrundlage und wegen der Annahme trotz begrenzter Rationalität jederzeit optimal reflektierten (Antwort-) Verhaltens vielfach kritisiert worden. Das MdFS hat diese Kritik aufgegriffen und die Theorie entsprechend modifiziert (Esser 2001: 259). Die Theorie gibt an, wann sich die im Rahmen der RCT definierten Anreize durch soziale Erwünschtheit auf die Antworten auswirken und wann *nicht,* sodass die Prognosen der RCT als ein möglicher Spezialfall im MdFS enthalten sind.

2.1 Die Erklärung des SD-Bias in der Theorie rationalen Handelns

Persönlichkeitstheoretische Erklärungen führen den SD-Bias auf das individuell unterschiedlich starke Anerkennungsbedürfnis der Befragten zurück.[1] Dieses Bedürfnis ist in der RCT der motivationale Kern und die „innere" Vorbedingung für den SD-Bias (Esser 1991). Die Erwartungen der Akteure über „äußere" Konsequenzen ihrer Antworten bilden dagegen

[1] Das Anerkennungsbedürfnis wird teilweise auf die an die äußere Umwelt gerichteten Selbstdarstellungsstrategien der Befragten („other deception") bezogen und teilweise als intrinsisch motivierte, völlig unbewusste Tendenz zur positiven Selbsttäuschung angesehen („self deception") (Paulhus 1984; für einen Überblick über diese und weitere Annahmen über die Ursachen des SD-Bias vgl. DeMaio 1984). Unsere Analyse bezieht sich ausschließlich auf die extrinsisch orientierte und strategisch motivierte Zielsetzung einer ‚Fremdtäuschung' der Befragten.

die kognitive Erklärungsgrundlage. Ein Vorzug der RCT besteht in der Möglichkeit genauer Vorhersagen über das Zusammenspiel der „inneren" Befragtenmerkmale und den „äußeren" Bedingungen der Interviewsituation bei der Entstehung des SD-Bias.

Soziale Anerkennung wird besonders effektiv in direkten Interaktionen durch den Ausdruck von Wertschätzung oder Missbilligung kommuniziert und so zur erlebbaren Realität. Rationale Befragte antizipieren solche Wertungen der sozialen Umwelt und schreiben sich durch ihr Antworten solche Merkmale zu, die die Chance positiver Reaktionen maximieren und das Risiko negativer Konsequenzen minimieren. Das setzt aber *erstens* voraus, dass die Befragten eine Sanktionierung ihres Antwortverhaltens prinzipiell für möglich halten. Und *zweitens*, dass die in der konkreten Befragungssituation gültigen Bewertungskriterien hinreichend klar bekannt sind.

Ob die Befragten Konsequenzen ihres Antwortverhaltens erwarten, ergibt sich aus der Privatheit und der Anonymität der Antwortsituation. Bei eingeschränkter *Anonymität* halten die Befragten eine Identifizierung und Sanktionierung ihrer Angaben zu einem späteren Zeitpunkt für möglich. Mit solchen Befürchtungen und entsprechenden Auswirkungen auf das Antwortverhalten ist primär bei Befragungsthemen mit strafrechtlichem Sanktionspotential zu rechnen. Dagegen bezieht sich die *Privatheit* der Antwortsituation auf die Wahrnehmbarkeit der Angaben durch bei der Befragung anwesenden Personen. Die damit verbundene Möglichkeit unmittelbarer, informeller Sanktionen ist bei Fragen nach bloß sozial geächteten Merkmalen, etwa den xenophoben Einstellungen der Befragten, eine zentrale Vorbedingung des SD-Bias.

Mit welcher Sicherheit die Befragten bei der konkreten Ausgestaltung der Interviewsituation Konsequenzen ihres Antwortverhaltens erwarten, wird in der RCT entweder durch objektive oder subjektive Wahrscheinlichkeiten abgebildet. Die objektive Sanktionswahrscheinlichkeit liegt bei ausgeschlossener Wahrnehmbarkeit der Antwortinhalte bei null, wohingegen ein hiervon eventuell abweichendes Privatheitsgefühl durch subjektive Wahrscheinlichkeiten größer als null erfasst werden können. Mit einer solchen Diskrepanz muss bei Informationsmangel oder in unstrukturierten Befragungskontexten gerechnet werden. Beide Bedingungen sind für die Wahrnehmbarkeit der Antworten nur selten erfüllt. Daher erwarten die Befragten nur Konsequenzen ihrer Angaben für die Befriedigung des Anerkennungsbedürfnisses, wenn andere Personen anwesend *und* objektiv zur Wahrnehmung der Angaben in der Lage sind.

Systematische Antwortverzerrungen werden nach der RCT auch nur bei für die Befragten klar erkennbaren Bewertungskriterien der sozialen Umwelt erwartet (Esser 1991). Diese durch die Interviewer als typische Adressaten der Selbstdarstellungsstrategien im Befragungskontext verwendeten Kriterien sind den Befragten normalerweise aber unbekannt. Auf der Grundlage von Bewertungsreaktionen bei vorangegangenen Fragen und wahrnehmbarer Interviewermerkmale können indirekte, aber relativ zuverlässige Schlussfolgerungen über die aktuell gültigen Bewertungskriterien gezogen werden. Durch ein effektives Interviewertraining wird jedoch normalerweise versucht, bewertende Reaktionen der Interviewer zu minimieren. Außerdem wählen professionelle Feldinstitute die Interviewer danach aus, dass deren erkennbare Merkmale einen möglichst geringen Bezug zu den Erwünschtheitswahrnehmungen beim betreffenden Befragungsgegenstand aufweisen. Wenn beide Maßnahmen perfekt gelingen, steht den Befragten die betreffende Informationsbasis für die Bestimmung ihrer situationalen Erwünschtheitswahrnehmungen nicht zur Verfügung. Nach den Annahmen der RCT ziehen die Befragten dann das weniger zuverlässige, weil für die konkrete Interviewsituation unspezifischere Wissen über soziale Normen als Grundlage ihrer Sanktionserwartungen heran (Stocké 2001). Die wahrgenommenen Anreize durch soziale Erwünschtheit steigen mit den Unterschieden in der antizipierten Wertung der verfügbaren Antwortoptionen an. Möglichst klare Unterschiede in dieser Hinsicht sind somit eine weitere notwendige Bedingung für die Entstehung des SD-Bias.

Zusammenfassend gesagt prognostiziert die RCT eine Abhängigkeit der Erklärungskraft jeder Einzeldeterminante des SD-Bias von der Ausprägungskonstellation der anderen Faktoren. Antwortverzerrungen werden nur bei klaren Erwünschtheitsunterschieden zwischen den Antwortoptionen, einer mangelnden Privatheit der Antwortsituation und hinreichend starkem Anerkennungsbedürfnis der Probanden erwartet. Ist auch nur *eine* dieser Bedingungen nicht erfüllt, so prognostiziert die RCT die Nennung des „wahren" Wertes und eine Ausschaltung auch der Erklärungskraft der anderen Determinanten des SD-Bias. Die Gesamtstärke und Richtung der wirksamen Anreize durch soziale Erwünschtheit ergeben sich aus dem nicht-kompensatorischen und daher multiplikativen Zusammenspiel der Einzeldeterminanten (vgl. Gleichung 1).

$$SEU(SD) = U_{SD} \bullet w_P \bullet \Delta w_{TD} \qquad (1)$$

Der *erste* Parameter U_{SD} repräsentiert das Befriedigungspotential durch anerkennende Reaktionen der sozialen Umwelt und steigt proportional mit dem Anerkennungsbedürfnis der Befragten an. Der Parameter nimmt bei maximaler Bedürfnisstärke den Wert eins und bei Befragten ohne Anerkennungsbedürfnis den Wert null an. Durch den *zweiten* Parameter w_P wird die situational geprägte Möglichkeit sozialer Sanktionen erfasst, die bei vollständiger Privatheit null und bei Wahrnehmbarkeit der Angaben eins ist. Der *dritte* Parameter Δw_{TD} bildet die Erwünschtheitswahrnehmungen der Befragten ab. Bei gleicher wahrgenommener Bewertung aller Antwortoptionen ist dieser Parameter null und nimmt bei maximalen Bewertungsunterschieden, je nach Richtung der Differenz, Werte von plus oder minus eins an.

2.2 Erklärung sozial erwünschten Antwortverhaltens im Modell der Frame-Selektion

Die RCT geht von in jeder Einzelsituation sorgfältig abwägenden und die Gesamtheit der verfügbaren Informationen nutzenden Akteuren aus. Verhalten auf der Grundlage generalisierter Einstellungen widerspricht dieser theoretischen Annahme, weil die Akteure hierbei ganze Klassen von Einstellungsobjekten zusammenfassen, diese gleichförmig beurteilen und den Implikationen dieser Wertungen oft ganz unreflektiert folgen. Dabei setzt sich die mit Einstellungen verbundene evaluative Prädisposition oft auch gegen das Selbstinteresse der Akteure durch (vgl. etwa Diekmann 1996). Das Modell der Frame-Selektion (MdFS) integriert systematisch beide Arten der Handlungsselektion und prognostiziert, wann instrumentell rationales und wann einstellungsbasiertes Verhalten erwartbar ist (Esser 2001: 259). Eine Ausschaltung der Anreize wird erwartet, wenn inhaltlich positive und kognitiv stark verankerte Einstellungen die Interviewsituation unter einem nicht der instrumentellen Abwägung unterworfenen und daher unhinterfragt gültigen Oberziel der Kooperation rahmen.[2]

In Anknüpfung an rollentheoretische Konzeptionen des Befragtenverhaltens wird im MdFS alternativ von der Einnahme einer kooperativen oder konformistischen Befragtenrolle durch die Probanden und der Dominanz einer damit jeweils verbundenen spezifischen Zielorientierung ausgegangen. So streben Befragte mit einer kooperativen Rollenorientierung eine bedingungslose Unterstützung der aktuellen Umfrage durch möglichst korrekte Antworten an. Eine solche Rahmung der Interviewsituation bewirkt die Ausschaltung auch starker Anreize durch soziale Erwünschtheit. Beherrscht andererseits eine konformistische, an der sozialen Realität im Interview orientierte Befragtenrolle die Situationsdefinition der Befragten, so zielen deren Antworten auf möglichst vorteilhafte Bewertungsreaktionen. Hier stimmen dann die Vorhersagen über die Determinanten des SD-Bias perfekt mit denen der RCT überein.

Die Qualität der von den Befragten eingenommenen Befragtenrolle ist das Ergebnis eines unreflektierten Einstellungsaktivierungsprozesses. Demnach bewirkt alleine schon die Teilnahme an einer Umfrage eine spontane Aktivierung der mit diesem Objekt assoziierten Bewertungen. Bei einer positiven Bewertung nehmen die Probanden ebenso spontan eine Rolle als kooperativer Befragter ein. Die hiermit verbundene Selbstverpflichtung zur bedingungslosen Unterstützung der Befragung setzt sich aber nur dann gegen die Versuchung zur positiven Selbstdarstellung durch, wenn es sich bei den zugrunde liegenden Einstellungen nicht um „Non Attitudes“ handelt. Solche „Non Attitudes“ liegen dann vor, wenn die betreffenden Urteile nicht hinreichend in der Identität der Probanden verankert und automatisch abrufbar im Gedächtnis gespeichert sind. Das Konzept hat somit enge Bezüge zu den Dimensionen der Einstellungsstärke, Einstellungskristallisation und Einstellungsverfügbarkeit (vgl. als Überblick: Wegener et al. 1995). Nach den Vorhersagen des MdFS wirken sich Anreize durch soziale Erwünschtheit *nicht* auf das Antwortverhalten aus, wenn die Befragten über positive *und* gleichzeitig kognitiv stark verankerte Einstellungen zu Umfragen verfügen. Die Befragten sehen dann eine kooperative Befragtenrolle M_{KOOP} als vollständig angemessen an. In welchem Ausmaß dies bei einem konkreten Befragten der Fall ist, wird durch den Parameter „m“ erfasst, während dessen Komplementärwert „1-m“ die Dominanz einer konformistischen Orientierung M_{KONF} abbildet. Die Ausprägung des Parameters ist das Ergebnis der multiplikativen Verknüpfung des durch den Parameter m_{KOOP} erfassten *Inhaltes* und der durch den Parameter m_{VANK} abgebildeten *Stärke* der Umfrageeinstellung ($m = m_{KOOP} \bullet m_{VANK}$). Da m_{KOOP} bei negativen Einstellungen und m_{VANK} bei „Non Attitudes“ jeweils null sind, ist in beiden Fällen die Do-

[2] Neben diesen Determinanten der Modell-Selektion werden im MdFS auch Erklärungsfaktoren des Informationsverarbeitungsmodus als relevant angesehen (Esser 2001: 271). Dieses relativ eigenständige theoretische Modul ist nicht Gegenstand der vorliegenden Untersuchung.

minanz „m“ einer kooperativen Orientierung ebenfalls null und eine konformistische Befragtenrolle gilt mit einer Sicherheit von eins.

Mit der inneren „Selektion“ einer kooperativen Befragtenrolle ist der Verzicht auf die durch sozial erwünschtes Antwortverhalten realisierbare Befriedigung des Anerkennungsbedürfnisses verbunden. Dieser Verzicht ist der Anreiz für einen Wechsel zu einer konformistischen Situationsorientierung und wird durch den Nutzenparameter U_{KONF} in die Gesamtprognose integriert. Die Stärke dieser Versuchung ergibt sich aus den im Rahmen der RCT prognostizierten Gesamtanreizen durch soziale Erwünschtheit ($U_{KONF} = SEU(SD) = U_{SD} \bullet w_P \bullet \Delta w_{TD}$; vgl. Gleichung 1 im vorherigen Abschnitt). Durch den Parameter U_{KONF} werden somit die Vorhersagen der RCT über die instrumentell rationalen Entstehungsbedingungen des SD-Bias in das MdFS integriert. Die Anreize durch soziale Erwünschtheit wirken sich mit einer zunehmend konformistischen Befragtenrolle immer stärker auf das Antwortverhalten aus. Die Gesamtdominanz $D(M_{KONF})$ einer konformistischen Situationsorientierung ergibt sich aus dem Zusammenspiel der *inneren* Einstellung der Befragten zu Umfragen und den *äußeren* Anreizen durch soziale Erwünschtheit. Eine solche Orientierung ist dann ausgeschlossen, wenn die Befragten eine Unterstützung von Umfragen ohne Zweifel als geboten ansehen ($1\text{-}m = 0.0$) oder keine Anreize zur positiven Selbstdarstellung wahrnehmen ($U_{KONF} = 0.0$). Da jede dieser Bedingungen alleine schon zur Ausschaltung des SD-Bias führt, werden beide Teildeterminanten multiplikativ in Beziehung gesetzt:

$$D(M_{KONF}) = (1\text{-}m) \bullet U_{KONF} \qquad (2)$$

Das MdFS integriert so die *instrumentell rationalen Anreize* für Antwortverzerrungen durch soziale Erwünschtheit mit den *einstellungsbasierten Determinanten* einer kooperativen Situationsorientierung zu einer Gesamtprognose der Antwortqualität. Dabei gelten die Vorhersagen der RCT wenn der Parameter „m“ oder auch nur eine seiner Teilkomponenten null ist. Dagegen lassen sich die Antwortinhalte mit steigenden Werten des Parameters „m“ immer weniger durch die im Parameter U_{KONF} zusammengefassten Anreize durch soziale Erwünschtheit erklären.

2.3 Prognosen der Rational-Choice Theorie und des Modells der Frame-Selektion

Nach der RCT ist das Antwortverhalten, soweit es durch soziale Erwünschtheit bedingt ist, durch das interdependente Zusammenspiel von drei Erklärungsfaktoren prognostizierbar. Demnach lassen sich die Angaben nur bei mangelnder, nicht aber bei sichergestellter Privatheit der Interviewsituation durch die von den Befragten wahrgenommene Erwünschtheit der Antwortoptionen vorhersagen. Die Erklärungskraft dieses Interaktionseffektes ist bei fehlendem Anerkennungsbedürfnis null und steigt mit zunehmender Bedürfnisstärke an.

Nach dem MdFS gelten diese Prognosen für Probanden mit einer konformistischen Befragtenrolle: Die in der RCT definierten Anreize durch soziale Erwünschtheit wirken sich bei Befragten mit einer negativen Umfrageeinstellung und bei solchen mit positiven „Non Attitudes“ auf die Antworten aus. Befragte mit einer positiven und gleichzeitig kognitiv starken Einstellung werden dagegen nicht durch soziale Erwünschtheit beeinflusst: Den in der RCT definierten Determinanten des SD-Bias kommt dann keine Erklärungskraft für das Antwortverhalten zu.

3. Forschungsstand über die vorhergesagten Determinanten des SD-Bias

Die Erklärungskraft eines Teils der von der RCT prognostizierten *Einzeldeterminanten* des SD-Bias wurde bereits intensiv empirisch untersucht. So haben Forscher die Stärke des Anerkennungsbedürfnisses durch „Social Desirability“ (SD)-Skalen erfasst und mit den Angaben der Probanden bei sensiblen Befragungsthemen in Beziehung gesetzt.[3] Hierbei wird die vergleichsweise undifferenzierte Hypothese geprüft, dass sich Befragte mit höheren SD-Werten mehr erwünschte Eigenschaften zuschreiben und unerwünschte leugnen (Crowne 1979: 153). Die empirische Evidenz für die generelle Gültigkeit dieser Hypothese ist außerordentlich inkonsistent.

Untersuchungsergebnisse haben etwa gezeigt, dass Probanden mit hohen SD-Werten nur selten gesellschaftlich stigmatisierte Symptome psychischer Erkrankungen oder ein hohes Ausmaß an Ängstlichkeit zugeben (Davies et al. 1998). Mit zunehmendem Anerkennungsbedürfnis steigt auch die Wahrscheinlichkeit für die Nennung positiv be-

[3] Alle SD-Skalen erfassen das Anerkennungsbedürfnis durch Fragen über positiv und negativ bewertete Eigenschaften der Probanden. Die positiv bewerteten Merkmale treffen praktisch für niemanden und die sozial unerwünschten Eigenschaften wahrscheinlich für jede Person zu. Die Probanden erhalten umso höhere SD-Werte, je mehr Merkmale sie sich vom ersten Typ dennoch zuschreiben und je mehr sie vom zweiten Typ leugnen.

werteter Werthaltungen an (Fisher/Katz 2000). In anderen Studien lassen sich diese Korrelationen jedoch nicht feststellen. So erweist sich die Stärke des Selbstdarstellungsmotivs als unabhängig vom Grad der selbstberichteten Lebenszufriedenheit (Vella-Broderick/White 1997). Auch die Wahrscheinlichkeit einer Selbstzuschreibung traumatischer Lebensereignisse steht nicht mit den SD-Werten in Zusammenhang (Brunet et al. 1996).

Auch im Bereich der Einstellungen zu ethnischen Minderheiten sind keine klaren Schlussfolgerungen über die Erklärungskraft des Anerkennungsbedürfnisses für den SD-Bias möglich. So geben in einer Studie von Mielke (1995) Befragte mit höheren Marlowe-Crowne SD-Werten weniger ethnische Vorurteile zu. Dieser Zusammenhang lässt sich in anderen Untersuchungen trotz Verwendung der gleichen SD-Skala nicht nachweisen (Duck/Hunsberger 1999).

Ob eine mangelnde Privatheit der Antwortsituation *ohne weitere Vorbedingungen* die Wahrscheinlichkeit eines SD-Bias erhöht, lässt sich an den Ergebnissen von Modus-Experimenten ablesen. Es wurde geprüft, ob die bei interviewer-administrierten Befragungen bestehende Wahrnehmbarkeit der Antwortinhalte zu mehr sozial erwünschten Angaben führt. Nach den vorliegenden Ergebnissen ist dies in vielen, aber keineswegs allen Studien der Fall (vgl. als Überblick: Tourangeau et al. 2000: 295).

In einer Reihe von „Split-Ballot“ Experimenten finden sich bei gesicherter Privatheit in selbst-administrierten Interviews mehr gesellschaftlich negativ bewertete Antworten: Die Befragten schreiben sich mehr Symptome psychischer Störungen sowie einen höheren Drogenkonsum zu (Aquilino 1994, Jobe et al. 1997). Im Widerspruch hierzu finden andere Studien beim Vergleich von selbst- und interviewer-administrierten Befragungen keine Unterschiede im selbstberichteten Alkohol- und Haschischkonsum sowie der Häufigkeit der in der Vergangenheit begangenen Ladendiebstähle (Mangione et al. 1982, Reuband/Blasius 1996).

Einige Studien finden in selbst-administrierten Befragungen durchschnittlich negativere Angaben weißer Befragter über ihre Haltung zu Afroamerikanern, als dies in interviewer-administrierten Befragungen der Fall ist (Krysan et al. 1994, Krysan 1998). Eine differenzierte Betrachtung auf der Ebene der den Gesamtskalen zugrunde liegenden Einzelitems zeigt aber, dass die private Befragungsart nur bei neun der insgesamt neunzehn untersuchten Items signifikant negativere Antworten und bei zwei Items sogar eine Tendenz in die entgegengesetzte Richtung bewirkt. Dies legt die Vermutung einer bei den Einzelfragen unterschiedlichen Stärke und Richtung der Anreizwahrnehmungen nahe.

Untersuchungen über die antizipierte Bewertung der Merkmale, die sich die Befragten durch ihre Antworten selbst zuschreiben, finden, dass die hierbei beobachteten individuellen Unterschiede immer mit den Angaben korreliert sind: Die Wahrscheinlichkeit für die Nennung einer Antwortoption steigt mit der wahrgenommenen Erwünschtheit der dadurch offenbarten Eigenschaft an.

Befragte charakterisieren sich durch ihre Antworten eher als argumentative Persönlichkeiten und schreiben sich Symptome psychischer Störungen sowie unterschiedliche Persönlichkeitseigenschaften zu, wenn sie das betreffende Merkmal als gesellschaftlich erwünscht einschätzen (Huang et al. 1998, Nicotera 1996, Phillips/Clancy 1970). Eine Studie von Wyner (1980) zeigt durch die Validierung der Angaben über die Anzahl vergangener Verhaftungen mit Polizeiakten, dass die Erwünschtheitswahrnehmungen nicht nur mit den Antwortinhalten korreliert sind, sondern zur Prognose des SD-Bias in der Lage sind. Die von den Befragten wahrgenommene Bewertung des erfragten Merkmals sagt vorher, wie stark die Anzahl vergangener Verhaftungen über- oder untertrieben wird.

Nur eine Forschungsarbeit hat bisher die Erwünschtheitswahrnehmungen unterschiedlicher Einstellungen zu ethnischen Fremdgruppen untersucht und gezeigt, dass die Befragten durchschnittlich solche Iteminhalte der Gesamtskala als erwünschter beurteilen, in denen sich eine positive Haltung zu Ausländern ausdrückt (Reinecke 1991: 152). Die beobachteten Bewertungen variieren jedoch deutlich zwischen den konkreten Items und unterscheiden sich nach dem Alter sowie dem sozio-ökonomischen Status der Informanten. Es bleibt eine offene Frage, ob sich die Erwünschtheitswahrnehmungen auch auf die Angaben der Befragten auswirken.

Die von der RCT prognostizierte *gegenseitige Abhängigkeit* der Einzeldeterminanten des SD-Bias wurde bisher kaum und wenn, mit inkonsistenten Ergebnissen untersucht. So hat sich gezeigt, dass die Antworten der Befragten über ihre positive Affektivität mit zunehmendem Anerkennungsbedürfnis immer stärker mit den Erwünschtheitswahrnehmungen übereinstimmen (Chen et al. 1997). In einer anderen Studie wirken sich die beiden Faktoren allerdings unabhängig voneinander auf die selbstberichtete Lebenszufriedenheit und die Angaben über ethnische Vorurteile aus (Phillips/Clancy 1972).

Die vom MdFS zusätzlich angenommenen Determinanten des SD-Bias wurden in einer Studie untersucht, in der die generalisierte Umfrageeinstellung erfasst und mit der Beeinflussbarkeit der Befragten durch soziale Erwünschtheit in Beziehung gesetzt wird (Rogelberg et al. 2001). Als Kriterium für diese Beeinflussbarkeit werden die Angaben bei einer

SD-Skala, dem sogenannten „Balanced Inventory of Desirable Responding (BIDR)“ herangezogen (Paulhus 1984). Die durch die SD-Werte erfasste Neigung zu sozial erwünschtem Antwortverhalten unterscheidet sich nach keiner der analysierten Teilskalen der Umfrageeinstellung.

Ob der kognitiven Verankerung der Umfrageeinstellung die ihr im MdFS zugeschriebene Bedeutung für die Entstehung des SD-Bias zukommt, ist derzeit völlig ungeklärt. Ergebnisse aus anderen Bereichen zeigen aber, dass die Prägekraft von Einstellungen für das Verhalten durch ein Bündel verschiedener Indikatoren der Einstellungsstärke vorhergesagt werden kann.

Das Verhalten ist substantiell besser durch Einstellungsangaben vorhersagbar, wenn es sich um extreme, weniger ambivalente, subjektiv sichere oder kognitiv leichter zugängliche Bewertungen handelt (Bassili/Bors 1997). Die direkte Erfahrung mit dem Einstellungsgegenstand erweist sich hierbei als wichtige Determinante der Einstellungsstärke (Schuette/Fazio 1995).

4. Mögliche Ursachen der inkonsistenten Ergebnislage

Die Angaben der Befragten werden in manchen Studien durch das Anerkennungsbedürfnis und die Antwortprivatheit beeinflusst, während diese Zusammenhänge in anderen Untersuchungen scheinbar grundlos ausbleiben. Diese inkonsistente Ergebnislage ist nach der RCT die Folge der isolierten Analyse der Einzelfaktoren. Hierbei bleibt nämlich unberücksichtigt, dass schon das Fehlen einer der Entstehungsbedingungen des SD-Bias die Wirksamkeit aller anderen Determinanten neutralisiert. So ist etwa nur bei fehlender Privatheit mit einem Anstieg des SD-Bias mit zunehmendem Anerkennungsbedürfnis zu rechnen. In der Tat bleiben diese Zusammenhänge in solchen Studien aus, die beide Faktoren durch schriftliche Befragungen und damit bei gesicherter Privatheit erfassen (Brunet et al. 1996, Duck/Hunsberger 1999, Vella-Broderick/White 1997).

Eine *zweite* theoretisch ableitbare Erklärung für die inkonsistente Ergebnislage bezieht sich auf eine mögliche Heterogenität bezüglich der *Richtung* wahrgenommener Anreize durch soziale Erwünschtheit innerhalb der befragten Stichproben. So konnte etwa die Existenz geschlechtsspezifischer Normen über das gebotene Ausmaß sexueller Promiskuität belegt werden, die bei Männern zur Übertreibung und bei Frauen zur Untertreibung der Anzahl ihrer Sexualpartner führen (vgl. Tourangeau et al. 2000: 295). Antwortverzerrungen in unterschiedliche Richtungen heben sich dann auf der Aggregatebene der Angaben gegenseitig auf und es lassen sich trotz Erklärungskraft der Entstehungsbedingungen des SD-Bias keine Einflüsse auf das Antwortverhalten feststellen.

Auch die negativen Ergebnisse über die vom MdFS vorhergesagte Erklärungskraft der Umfrageeinstellung sind wegen der Vernachlässigung anderer Entstehungsbedingungen des SD-Bias wenig aussagekräftig. So ist in der Studie von Rogelberg und Kollegen (2001) nicht klar, ob in dieser Studie überhaupt ein SD-Bias vorgelegen hat. Da die Datenerhebung durch eine schriftliche Befragung und damit bei sichergestellter Privatheit durchgeführt wurde, muss dies aus theoretischen Gründen angezweifelt werden.

5. Empirische Untersuchung

Die RCT macht charakteristische Vorhersagen über die gegenseitige Abhängigkeit der Determinanten des SD-Bias bei der Erklärung des Antwortverhaltens und darüber, bei welcher Ausprägungskonstellation sich die Einzelfaktoren *nicht* auf die Befragtenangaben auswirken. Diese Prognosen werden im *ersten* Teil der folgenden empirischen Analyse getestet. Im *zweiten* Teil wird der Frage nachgegangen, ob mit der im MdFS vorgeschlagenen theoretischen Erweiterung der RCT und der damit verbundenen weiteren Differenzierung der Entstehungsbedingungen des SD-Bias eine verbesserte Vorhersage des Befragtenverhaltens erzielt werden kann.

5.1 Stichprobe und Vorgehensweise

Die insgesamt 150 Befragten entstammen einer lokalen, mehrstufigen Zufallsstichprobe der volljährigen Wohnbevölkerung mit deutscher Staatsbürgerschaft einer süddeutschen Großstadt. Alle Befragungen wurden computergestützt und in der Form persönlicher Interviews bei den Befragten zu Hause durchgeführt.[4] Die Haushaltsstichprobe

[4] Die Befragungen wurden von insgesamt 14 Interviewern durchgeführt, die für Ihre Arbeit entlohnt wurden. Bei der Auswahl der Kandidaten wurden ausschließlich Bewerber berücksichtigt, deren sichtbare Merkmale nicht auf einen Migrationshintergrund hindeuteten und die keinen ausländischen Akzent aufwiesen. Die Interviewer waren zu 42.9 % männlich, hatten zu 92.9 % mindestens das Abitur und waren im Durchschnitt 27.4 Jahre alt. Da es sich so-

wurde durch eine Zufallsbegehung und die Zielperson im Haushalt durch die „Geburtstagsmethode" ausgewählt. An den Interviews nahmen 55 % Frauen und 45 % Männer mit einem Durchschnittsalter von 46.6 Jahren teil. Die Befragten haben zu 32 % einen Hauptschulabschluss, zu 25 % die mittlere Reife und zu 43 % das Abitur. Von den 46 % erwerbstätigen Befragten sind 9 % Arbeiter, 67 % Angestellte oder Beamte und 25 % selbstständig. Die Ausschöpfungsquote beträgt 31.3 %. Der erste Teil der Befragung wurde mit allen Probanden interviewer-administriert durchgeführt (CAPI-Modus). Vor dem zweiten Interviewteil wurden die Teilnehmer, gesteuert durch eine Zufallsprozedur des Befragungsprogramms, einer von zwei Experimentalbedingungen zugeteilt. Bei der ersten Bedingung wurde die Befragung weiter durch die Interviewer durchgeführt, während bei Probanden des zweiten Experimentalsplits ein Wechsel zu einem selbst-administrierten Befragungsmodus stattfand (CASI-Modus).

5.2 Operationalisierung der theoretischen Parameter

Als Anwendungsgegenstand der empirischen Untersuchung werden die Antworten der Befragten bei 10 Items zur Erfassung der Einstellung zu Ausländern herangezogen (vgl. Tabelle 1 für die Item-Formulierung).[5] Die Angaben werden auf siebenstufigen Antwortskalen erfasst und für die folgenden Analysen so rekodiert, dass hohe Skalenwerte bei allen Items eine positive Haltung zu Ausländern ausdrücken. Die insgesamt fünf theoretischen Parameter der beiden Erklärungsansätze werden wie folgt operationalisiert:

– *Bedürfnis nach sozialer Anerkennung (U_{SD}):* Unterschiede im Bedürfnis nach sozialer Anerkennung zwischen den Probanden werden durch die Angaben bei der deutschen Kurzversion der Marlowe-Crowne SD-Skala erfasst (vgl. Stocké 2003). Als Indikator für die Stärke des Anerkennungsbedürfnisses wird ein additiver Index der insgesamt 10 gleichgerichteten Antworten dieses Messinstrumentes herangezogen, wobei höhere Werte ein stärkeres Anerkennungsbedürfnis ausdrücken. Da die Validität des Instrumentes auf einer maximalen Umsetzung des latenten Anerkennungsbedürfnisses in das Antwortverhalten beruht, wird die SD-Skala für alle Befragten im zuerst stattfindenden CAPI-Teil der Interviews erhoben.

– *Privatheit der Antwortsituation (w_P):* Die Wahrnehmbarkeit der berichteten Einstellungen zu Ausländern wird durch den im zweiten Teil der Interviews variierten Befragungsmodus operationalisiert. Die Interviewer lesen die Einstellungsfragen beim CAPI-Modus vor und geben die von den Befragten verbal geäußerten Antworten in den Laptop-Computer ein. Im Gegensatz zu dieser öffentlichen Interviewart lesen die Probanden die Fragen beim privaten CASI-Modus auf dem Bildschirm und geben ihre Antworten selbst in den Computer ein. Die Interviewer sind auch bei CASI-Interviews anwesend und beantworten Verständnisfragen, ziehen sich aber soweit zurück, dass sie keinen Einblick in die Antwortinhalte nehmen können.

– *Relative soziale Erwünschtheit der Einstellungspositionen (Δw_{TD}):* Die Erwünschtheitswahrnehmungen werden in der Form „extrinsisch evaluierter", von den Befragten in der sozialen Umwelt durchschnittlich erwarteten Bewertungsreaktionen operationalisiert (Edwards 1957). Den Befragten wird eine dem Interview nachgestellte Gesprächssituation zwischen zwei Fremden geschildert und die Einschätzung erfragt, wie peinlich den Gesprächspartnern bestimmte Meinungsäußerungen über Ausländer wären (vgl. für die genaue Darstellung des Messinstrumentes: Stocké 2004). Die dabei bewerteten 20 Ansichten sind mit jenen Standpunkten identisch, die sich in einer vollständigen Zustimmung oder Ablehnung jedes der 10 zur Erfassung der Ausländereinstellung verwendeten Iteminhalte ausdrücken. Diese Erwünschtheitswahrnehmungen werden für jedes Items so rekodiert und dann voneinander abgezogen, dass positive Differenzwerte eine größere Erwünschtheit positiver und negative Werte eine höhere Erwünschtheit negativer Einstellungsantworten anzeigen. Bei einem Wert von null wird eine Zustimmung und Ablehnung des betreffenden Item-Inhaltes als gleich erwünscht angesehen.

– *Generalisierte Umfrageeinstellung (m_{KOOP}):* Die Einstellungen zu Umfragen wurden mit 16 Einstellungsfragen erfasst, die sich auf die wahrgenommene Nützlichkeit und Verlässlichkeit von Umfragen sowie auf die angenommenen Belastungen bei

---

mit um einen relativ jungen und hoch gebildeten Interviewerstab handelte, können, zusätzlich zu den in unserer Studie analysierten Einflüssen sozialer Erwünschtheit, Interviewereinflüsse nicht ausgeschlossen werden. Zusätzliche Analysen haben aber gezeigt, dass sich die genannten Interviewermerkmale nicht signifikant auf das Antwortverhalten ausgewirkt haben (Analyseergebnisse nicht berichtet).

[5] Alle Einstellungsfragen wurden der 1996 durchgeführten Schwerpunktstudie „Ausländer" der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS) entnommen.

einer Befragungsteilnahme beziehen (für Einzelheiten der Skalenkonstruktion vgl. Stocké/Langfeldt 2004). Die Einstellungen werden mit siebenstufigen Antwortskalen erhoben und, nachdem die Polung der Einzelitems gleichgerichtet wurde, durch Mittelwertbildung zu einem Index der Umfrageeinstellung zusammengefasst.

– *Verankerung der Umfrageeinstellung (m_{VANK}):* Als Indikator für die Stärke der Umfrageeinstellung wird *erstens* die vor der aktuellen Befragung bereits absolvierte Anzahl an Interviews als Indikator der Umfrageerfahrung der Probanden herangezogen.[6] Dieser Operationalisierung liegt die Annahme zugrunde, dass die Kristallisation von Einstellungen mit der direkten Erfahrung mit dem Einstellungsgegenstand zunimmt (Wegener et al. 1995). Die *zweite* Operationalisierung ist die durchschnittliche Antwortgeschwindigkeit bei der Beantwortung der 16 Fragen nach der Einstellung zu Umfragen. Antwortlatenzen sind ein Indikator für die kognitive Zugänglichkeit und damit für die Stärke von Einstellungen (Bassili/Bors 1997). Die Prognosekraft der Antwortlatenzen für die Informationszugänglichkeit kann durch deren Standardisierung an der mittleren Antwortgeschwindigkeit der einzelnen Befragten verbessert werden (Fazio 1990). Daher wird für jeden Befragten die mittlere Antwortgeschwindigkeit bei 118 anderen Fragen berechnet und von der bei der Umfrageeinstellung beobachteten Antwortlatenz abgezogen. Die Ergebnisvariable nimmt negative Werte an, wenn die Zielfrage, relativ zur „normalen" Antwortgeschwindigkeit des betreffenden Befragten, überdurchschnittlich schnell beantwortet wird. Positive Werte bedeuten einen überdurchschnittlich langen Antwortgenerierungsprozess. Die standardisierten Antwortgeschwindigkeiten der einzelnen Einstellungsfragen werden durch Mittelwertbildung zu einem Index der kognitiven Verankerung der Umfrageeinstellung zusammengefasst.

– *Gesamtanreize durch soziale Erwünschtheit (SEU(SD)):* Die Stärke und Richtung des SD-Bias wird im Rahmen der RCT durch den Betrag und das Vorzeichen des Erwartungsnutzens sozial erwünschter Antworten prognostiziert. Diese Gesamtanreize werden durch eine multiplikative Verknüpfung ihrer Einzeldeterminanten berechnet ($SEU(SD) = U_{SD} \bullet w_P \; \Delta w_{TD}$). Die empirisch beobachteten Variablen werden hierzu auf die theoretisch angenommenen Wertebereiche standardisiert. Dieser Wertebereich liegt für die relative Erwünschtheit zwischen minus eins (negative Einstellung erwünschter) und plus eins (positive Einstellungen erwünschter), für das Anerkennungsbedürfnis zwischen null (kein Anerkennungsbedürfnis) und eins (starkes Anerkennungsbedürfnis) und für die Privatheitsbedingungen zwischen null (Angaben nicht wahrnehmbar) und eins (Angaben wahrnehmbar). Die Gesamtanreize variieren zwischen minus eins (starke Anreize für eine negative Einstellungsantwort) und plus eins (starke Anreize für positive Angaben).

5.3 Ergebnisse

5.3.1 Deskriptive Ergebnisse

Die Angaben der Befragten über ihre Einstellung zu Ausländern unterscheiden sich stark nach dem konkreten Inhalt der Items (vgl. Tab. 1). Die durchschnittlichen Antworten auf der Skala von 1 (negative Einstellung) bis 7 (positive Einstellung) liegen zwischen einem Wert von 3.3 bei Item 1 und einem Wert von 5.4 bei Item 3. Für alle Fragen der Gesamtskala lässt sich ein durchschnittlicher Skalenwert von 4.8 und damit ein eher positives Antwortverhalten feststellen. Die von den Befragten durchschnittlich wahrgenommenen Unterschiede in der Erwünschtheit positiver und negativer Einstellungsangaben liegen für die Gesamtskala bei einem Wert von 0.30, sodass positive Einstellungsangaben gesellschaftlich eher als akzeptierter angesehen werden. Dieser Kennwert variiert aber mit Werten zwischen –0.66 (Item 1) und +1.59 (Item 7) deutlich zwischen den Einzelfragen. Auf der Ebene der Einstellungsfragen besteht ein starker Zusammenhang zwischen den Antwortinhalten und den Erwünschtheitswahrnehmungen ($r = .82$; $p < 0.01$).[7]

Eine nach der Richtung der von den Befragten individuell wahrgenommenen Anreize differenzierte

[6] Das Ausmaß der Umfrageerfahrung der Probanden wurde zweistufig erfasst. Den Probanden, die im ersten Schritt angegeben hatten, vor der aktuellen Umfrage bereits an einer anderen Befragung teilgenommen zu haben, wurde die folgende Frage gestellt: „Abgesehen von der heutigen Befragung, wie häufig haben Sie bisher insgesamt an einer Umfrage teilgenommen?"

[7] Bei einer multivariaten Analyse der Erwünschtheitswahrnehmungen lassen sich Unterschiede nach der Bildung und dem Erwerbsstatus der Befragten feststellen. Demnach sehen Befragte mit einem Universitätsabschluss genauso wie Arbeiter eine positive Einstellung zu Ausländern als gesellschaftlich deutlich erwünschter an, als dies bei den jeweils anderen Sozialgruppen der Fall ist. Dagegen liegt hinsichtlich des Alters, des Einkommens und des Geschlechtes keine Differenzierung der Erwünschtheitswahrnehmungen vor (für diese Analyseergebnisse vgl. Stocké 2004).

Analyse zeigt, dass bezogen auf alle Items der Gesamtskala 44.1 % der Stichprobe eine *positive Einstellungsantwort* als erwünschter ansehen, während 36.0 % der Probanden eine *negative Haltung* zu Ausländern für erwünschter halten. Ein substantieller Anteil von 19.9 % der Informanten erwartet bei der Nennung positiver und negativer Einstellungen völlig identische Reaktionen der sozialen Umwelt.

Der bei der Analyse der aggregierten Unterschiede in den Erwünschtheitswahrnehmungen festgestellte niedrige Wert von 0.30 legt die Schlussfolgerung eines nur geringen Verzerrungspotentials durch soziale Erwünschtheit nahe. Dieser Eindruck wird bei einer nach dem Vorzeichen der individuellen Wahrnehmungen differenzierten Betrachtung der *Anreizstärke* stark relativiert. So nehmen Befragte mit einer Erwünschtheitswahrnehmung zugunsten einer positiven Haltung zu Ausländern durchschnittlich Anreize von 3.2 Skalenpunkten für eine solche Einstellungsantwort wahr (vgl. Tab. 1). Probanden mit einer entgegengesetzten Wahrnehmung sehen Anreize von 3.1 Skalenpunkten, sodass die beiden Teilstichproben ähnlich starke, von der Richtung her aber entgegengesetzte Anreize für Antwortverzerrungen annehmen.

Die Betrachtung der anderen Erklärungsfaktoren des SD-Bias zeigt *erstens*, dass die Befragten in unserer Stichprobe mit einem durchschnittlichen SD-Skalenwert von 5.7 ein Anerkennungsbedürfnis mittlerer Stärke haben (vgl. Tab. 2).[8] *Zweitens* stehen die Probanden Umfragen mit einem mittleren Einstellungswert von 4.9 recht positiv gegenüber, wobei dieser Wert auf der Ebene der Befragten zwischen einer relativ kritischen Bewertung von 3.3 und einem sehr positiven Wert von 6.6 variiert. Ein *drittes* Ergebnis ist, dass die Probanden über ein hohes Ausmaß an Umfrageerfahrung verfügen: 71 % haben zuvor bereits mindestens einmal an einem Interview teilgenommen. Die durchschnittlich absolvierte Zahl an Interviews liegt bei 4.5. Die Umfrageeinstellung ist *viertens* kognitiv sehr unterschiedlich stark verankert, sodass deren Antwortlatenz bis zu 5.3 Sekunden unter und 4.6 Sekunden über der „normalen" Antwortgeschwindigkeit der Befragten liegt. *Fünftens* variiert der im Rahmen der RCT prognostizierte Gesamtnutzen sozialer Erwünschtheit zwischen –0.90 und +0.75. Es liegen somit starke Gesamtanreize für die Nennung positiver wie auch negativer Einstellungen zu Ausländern vor.

5.3.2 Überprüfung der Prognosen der Rational-Choice Theorie

Nach den Vorhersagen der RCT wirkt sich die mit den Befragungsmodi verbundene Privatheit der Antwortsituation in Abhängigkeit vom Betrag und Vorzeichen der relativen Erwünschtheitswahrnehmungen in verschiedenem Ausmaß und verschiedener Richtung auf das Antwortverhalten aus. Diese nach der Privatheit des Befragungskontextes selektive Konformität sollte mit dem Anerkennungsbedürfnis der Probanden ansteigen. Die so prognostizierte Dreifach-Interaktion wird mit einem entsprechend berechneten multiplikativen Faktor auf ihre Erklärungskraft für die Einstellungsantworten überprüft.

Die in Tabelle 3 dargestellten Regressionsergebnisse belegen die Gefahr falscher Schlussfolgerungen über die Determinanten der Stärke und Richtung des SD-Bias, wenn nur die Haupteffekte der Erklärungsfaktoren berücksichtigt werden (vgl. Modell 1).[9] Wird das Erklärungsmodell in dieser Weise unvollständig spezifiziert, so lassen sich *erstens* keine Einflüsse der mit den Befragungsmodi jeweils verbundenen Privatheit auf die Antworten feststellen. Dieses Ergebnis deutet auf die scheinbare Abwesenheit eines SD-Bias hin. Der signifikante negative Haupteffekt der SD-Werte zeigt *zweitens*, dass mit zunehmendem Anerkennungsbedürfnis immer negativere Angaben über die Einstellungen zu Ausländern gemacht werden.

[8] Die Schätzung multiplikativer Interaktionsparameter führt zu hoher Multikolliniarität. Der damit verbundenen Inflationierung der für die Regressionsparameter geschätzten Standardfehler kann durch die Verwendung z-standardisierter Ausgangsvariablen begegnet werden (Cronbach 1987). Dadurch werden die Schätzergebnisse für Parameter niedrigerer Ordnung, nicht aber die für unser Analyseziel relevanten Parameter der höchsten Interaktionsstufe beeinflusst (Aiken/West 1991: 28ff.). In den folgenden Regressionsmodellen gehen daher die an Interaktionseffekten beteiligten Variablen z-standardisiert in die Analyse ein. In Tabelle 2 werden diese daher als Rohskalenwerte und in ihrer standardisierten Form dargestellt.

[9] Die abhängige Variable dieser und der folgenden Analysen besteht aus den kumulierten Antworten bei allen zehn Einstellungsfragen, sodass die Beobachtungen nicht als unabhängig angesehen werden können. Dies bewirkt eine Überschätzung der Zuverlässigkeit der Regressionsparameter. Dieses Problem wird durch Verwendung des „Huber-White Sandwich"-Schätzers für robuste Standardfehler mit den Befragten als Cluster-Variable korrigiert (STATA Corporation 1999: 165). In allen folgenden Regressionsanalysen werden neben den dargestellten Erklärungsfaktoren auch die Bildung, der Erwerbsstatus, das Einkommen, das Alter und das Geschlecht der Befragten sowie Item-Dummies als Kontrollvariablen in die Analyse einbezogen. Die resultierenden Parameter werden aus Platzgründen nicht berichtet.

Tabelle 1 Antwortverhalten der Befragten über ihre Einstellungen zu Ausländern und die Stärke sowie Richtung der wahrgenommenen relativen Erwünschtheit positiver und negativer Einstellungsangaben

	Antwortverhalten [a]	*Relative Erwünschtheit unterschiedlicher Einstellungen zu Ausländern:*			
		Positive Einstellung erwünschter [b]	Positive und negative Einstellung gleich erwünscht	Negative Einstellung erwünschter [c]	Relative Erwünschtheit insgesamt [d]
	Mittelwert	Anteil der Befragten Mittelwert	Anteil der Befragten	Anteil der Befragten Mittelwert	Anteil der Befragten Mittelwert
Item 1. „Die in Deutschland lebenden Ausländer sollten ihren Lebensstil ein bisschen besser an den der Deutschen anpassen“	3.3	30.0 % +2.1	24.0 %	46.0 % –2.8	100 % –0.7
Item 2. „Wenn Arbeitsplätze knapp werden, sollte man die in Deutschland lebenden Ausländer wieder in ihre Heimat zurückschicken“	5.2	45.4 % +3.2	21.3 %	33.3 % –3.1	100 % +0.4
Item 3. „Die in Deutschland lebenden Ausländer sollten sich ihre Ehepartner unter ihren eigenen Landsleuten auswählen“	5.4	47.3 % +3.5	20.7 %	32.0 % –2.8	100 % +0.8
Item 4. „Man sollte den in Deutschland lebenden Ausländern jede politische Betätigung in Deutschland untersagen“	5.1	50.0 % +3.2	19.3 %	30.7 % –3.1	100 % +0.6
Item 5. „Durch die vielen Ausländer in Deutschland fühlt man sich zunehmend als Fremder im eigenen Land“	5.0	46.7 % +3.1	14.0 %	39.3 % –2.7	100 % +0.4
Item 6. „Die in Deutschland lebenden Ausländer sind eine Belastung für das soziale Netz“	4.6	42.7 % +3.0	17.3 %	40.0 % –3.1	100 % +0.1
Item 7. [R] „Die in Deutschland lebenden Ausländer sind eine Bereicherung für die Kultur in Deutschland“	5.2	58.7 % +4.0	18.0 %	23.3 % –3.3	100 % +1.6
Item 8. „Die in Deutschland lebenden Ausländer begehen häufiger Straftaten als die Deutschen“	4.3	36.7 % +3.3	18.6 %	44.7 % –3.7	100 % –0.4
Item 9. „Die Anwesenheit von Ausländern in Deutschland führt zu Problemen auf dem Wohnungsmarkt“	5.0	41.4 % +2.8	27.3 %	31.3 % –2.6	100 % +0.4
Item 10. [R] „Die in Deutschland lebenden Ausländer sollten den gleichen Anspruch auf Sozialhilfe und andere Sozialleistungen haben wie die Deutschen“	4.7	42.0 % +3.2	18.7 %	39.3 % –3.7	100 % –0.1
Gesamt	4.8	44.1 % +3.2	19.9 %	36.0 % –3.1	100 % +0.3

[a] Wertebereich: 1 (sehr negative Einstellung) bis 7 (sehr positive Einstellung). [b] Wertebereich: > 0 (positive Einstellungen etwas erwünschter) bis +8 (positive Einstellungen stark erwünschter). [c] Wertebereich: < 0 (negative Einstellungen etwas erwünschter) bis –8 (negative Einstellungen stark erwünschter). [d] Wertebereich: –8 (negative Einstellungen erwünschter) bis +8 (positive Einstellungen erwünschter). [R] Antworten und relative Erwünschtheitswahrnehmungen bei den mit „R“ markierten Items wurden durch Rekodierung an die Richtung der Gesamtskala angepasst.

Tabelle 2 Deskriptive Statistiken der erklärenden Variablen

	Minimum/Maximum	Mittelwert	Standardabweichung
Bedürfnis nach sozialer Anerkennung			
– Rohwerte [a)]	0 / 10	5.71	2.21
– z-Werte	–2.58 / 1.95	0.0	1.0
Einstellung zu Umfragen			
– Rohwerte [b)]	3.3 / 6.6	4.9	0.65
– z-Werte	–2.6 / 2.6	0.0	1.0
Umfrageerfahrung in der Vergangenheit			
– Anzahl absolvierter Befragungen	0 / 30	4.48	5.69
– z-Werte	–0.79 / 4.49	0.0	1.0
Standardisierte Geschwindigkeit bei der Beantwortung der Umfrageeinstellungsitems			
– Rohwerte (hundertstel Sekunden)	–529.4 / 459.2	–36.7	140.4
– z-Werte	–3.5 / 3.5	0.0	1.0
Gesamtanreize durch soziale Erwünschtheit			
– Rohwerte [c)]	–0.90 / 0.75	0.001	0.19
– z-Werte	–4.77 / 3.95	0.0	1.0

[a)] Skalenwerte: 0 „kein Anerkennungsbedürfnis“, 10 „maximales Anerkennungsbedürfnis“; [b)] Skalenwerte: 1 „negative Umfrageeinstellung“, 7 „positive Umfrageeinstellung“; [c)] Skalenwerte: +1 „maximale Anreize für eine positive Einstellung“, –1 „maximale Anreize für eine negative Einstellung“.

Tabelle 3 Test der im Rahmen der RCT vorhergesagten Erklärungsfaktoren des Antwortverhaltens über die Einstellungen zu Ausländern (OLS-Regressionsparameter mit robuster Schätzung der Standardfehler)

	Modell 1 B (t-Wert)	Modell 2 B (t-Wert)
Rational-Choice Modell		
(1) RELATIVE ERWÜNSCHTHEIT	0.38 (5.2) **	0.22 (2.7) **
(2) ANERKENNUNGSBEDÜRFNIS	–0.26 (2.7) **	–0.16 (1.2)
(3) BEFRAGUNGSMODUS (CAPI) [a)]	0.06 (0.3)	0.09 (0.5)
(4) ERWÜNSCHTHEIT • MODUS	–	0.27 (2.2) **
(5) BEDÜRFNIS • MODUS	–	–0.16 (0.9)
(6) BEDÜRFNIS • ERWÜNSCHTHEIT	–	–0.05 (0.7)
(7) BEDÜRFNIS • ERWÜNSCHTHEIT • MODUS	–	0.25 (2.4) **
Konstante	1.83 (3.3) **	1.86 (3.3) **
R^2	0.26	0.27
N	1477	1477

Signifikanz: * $p < 0.1$; ** $p < 0.05$; Referenzkategorie: [a)] CASI

Dies legt die Schlussfolgerung eines globalen SD-Bias in Richtung auf negative Antworten nahe.

Nach den Ergebnissen des vollständig spezifizierten, zweiten Regressionsmodells erweist sich der multiplikative Faktor aus dem Anerkennungsbedürfnis, den relativen Erwünschtheitswahrnehmungen und den Befragungsmodi als statistisch abgesicherter Prädiktor des Antwortverhaltens. Die Einstellungsangaben können signifikant besser vorhergesagt werden, wenn, ausgehend von einem reduzierten Regressionsmodell, die Determinanten des SD-Bias und insbesondere deren Zusammenspiel einbezogen werden ($F\ (6, 1247.9) = 16.5$; $p \leq 0.01$).[10]

[10] Das reduzierte Regressionsmodell enthält die in Fußnote 9 berichteten Kontrollvariablen und erzielt einen R^2-Wert von 0.21. Die Freiheitsgrade des F-Testes ergeben sich durch eine Gewichtung der Ausgangsfreiheitsgrade mit einem Greenhouse-Geissers Epsilon von 0.86. Dadurch wird berücksichtigt, in welchem Ausmaß die Daten

Tabelle 4 Angaben über die Einstellungen zu Ausländern nach dem Befragungsmodus sowie den relativen Erwünschtheitswahrnehmungen und dem Anerkennungsbedürfnis der Befragten (Vorhersagewerte aus Regressionsmodell 2)

	Anerkennungsbedürfnis schwach			*Anerkennungsbedürfnis stark*		
	Befragungsmodus			Befragungsmodus		
	Selbst-administriert	Interviewer-administriert	Differenz zwischen den Modi	Selbst-administriert	Interviewer-administriert	Differenz zwischen den Modi
Relative Erwünschtheit der Einstellung zu Ausländern						
positive Einstellung erwünschter	4.5	4.7	+ 0.2	4.0	4.5	+ 0.5
positive und negative Einstellung gleich erwünscht	4.2	4.4	+ 0.2	3.9	3.8	+ 0.1
negative Einstellung erwünschter	3.9	4.1	+ 0.2	3.7	3.1	– 0.6

Einstellungsangaben reichen von 1 (negative Einstellung) bis 7 (positive Einstellung). Die Prognosewerte gelten beim Anerkennungsbedürfnis für Befragte mit einer Standardabweichung über (starkes Bedürfnis) und unter (schwaches Bedürfnis) dem Stichprobenmittelwert. Bei den Erwünschtheitswahrnehmungen gehen Werte für Befragte mit einer Standardabweichung über (positive Einstellung erwünschter) und unter (negative Einstellung erwünschter) dem Stichprobenmittelwert in die Prognose ein. Für die Kategorie „gleich erwünscht" wurde der Stichprobenmittelwert eingesetzt. Kontrollvariablen sind bei kategorialen Variablen auf der Referenzkategorie und bei kontinuierlichen Faktoren auf dem Stichprobenmittelwert fixiert.

Der Inhalt der für unsere Fragestellung zentralen Dreifach-Interaktion lässt sich durch die auf der Grundlage von Regressionsmodell 2 prognostizierten Einstellungsantworten veranschaulichen. Es zeigen sich für Befragte mit einem schwachen Anerkennungsbedürfnis nur geringe Einflüsse der mit dem Befragungsmodus verbundenen Wahrnehmbarkeit der Antwortinhalte: Bei interviewer-administrierten Befragungen findet sich für alle Probanden eine nur um durchschnittlich 0.2 Skalenpunkten positivere Einstellung zu Ausländern. Auch wirken sich die Erwünschtheitswahrnehmungen nur als Haupteffekt, unabhängig vom Befragungsmodus und damit unbeeinflusst von der sozialen Realität im Interview auf die Antworten aus. Demnach lassen sich kaum Privatheitseffekte und insbesondere keine Differenzierung dieser Einflüsse nach der Richtung der wahrgenommenen Erwünschtheitsunterschiede feststellen.

Im Gegensatz zu Befragten mit schwachem Anerkennungsbedürfnis orientieren sich die Angaben der Probanden mit hohen SD-Werten stark an der in der Gesellschaft wahrgenommenen relativen Erwünschtheit von Einstellungen zu Ausländern. Es lassen sich für diese Gruppe substantielle Unterschiede im Antwortverhalten zwischen den Befragungsmodi feststellen, deren Richtung perfekt mit den wahrgenommenen Erwünschtheitsunterschieden übereinstimmt. So neigen bei Wahrnehmbarkeit der Antwortinhalte sowohl Befragte, die eine positive Haltung zu Ausländern als gesellschaftlich anerkannter ansehen, wie auch solche, die bei negativen Einstellungsangaben vorteilhaftere Reaktion der sozialen Umwelt erwarten, zu einer starken Konformität mit der jeweiligen Erwünschtheitswahrnehmung. Nach den vorliegenden Analyseergebnissen führt, je nach Richtung der Erwünschtheitswahrnehmungen, eine mangelnde Privatheit zu um 0.5 Skalenpunkten positiveren oder 0.6 Skalenpunkten negativeren Einstellungsangaben. Werden dagegen positive und negative Einstellungsantworten als identisch bewertet wahrgenommen, so lassen sich mit einem Effekt von 0.1 Skalenpunkten praktisch keine Einflüsse der Privatheit der Antwortsituation feststellen.

5.3.3 Überprüfung der Prognosen des Modells der Frame-Selektion

Die bisherige Analyse hat die Prognosen der „reinen" RCT über die Entstehungsbedingungen des SD-Bias und über das interaktive Zusammenspiel dieser Faktoren bestätigt. Diese Faktoren sind auch im MdFS relevant, deren Erklärungskraft wird aber bei Befragten mit einer zuverlässigen, kooperativen Rahmung der Befragungssituation ausgeschaltet. Eine solche Rahmung ist bei einer positiven *und* gleichzeitig kognitiv stark verankerten Umfrageein-

von der Annahme unabhängiger Beobachtungen abweichen (Stevens 1996: 459ff.).

Tabelle 5 Bedeutung des Inhalts und der Stärke der Umfrageeinstellung für den Einfluss der Anreize durch soziale Erwünschtheit auf die Angaben über die Einstellung zu Ausländern (OLS-Regressionskoeffizienten mit robuster Schätzung der Standardfehler)

	Modell 3 B (t-Wert)	Modell 4 B (t-Wert)	Modell 5 B (t-Wert)
MdFS: Verankerungsindikator „Antwortlatenz"			
(1) UMFRAGEEINSTELLUNG (Indexwerte)	0.41 (4.1) **	0.42 (4.1) **	0.51 (5.4) **
(2) ANREIZE SOZIALER ERWÜNSCHTHEIT (Indexwerte)	0.38 (5.5) **	0.36 (5.3) **	0.49 (6.5) **
(3) UMFRAGEERFAHRUNG (mehr als 4 Interviews) a)	–	–	0.37 (1.7)
(4) ANTWORTLATENZ (hundertstel Sekunden)	–	0.09 (1.2)	0.07 (0.9)
(5) ANREIZE • EINSTELLUNG	0.01 (0.2)	–0.05 (0.8)	0.07 (0.9)
(6) LATENZ • EINSTELLUNG	–	0.08 (1.2)	0.09 (1.0)
(7) LATENZ • ANREIZE	–	0.08 (1.4)	–0.06 (1.5)
(8) LATENZ • ANREIZE EINSTELLUNG	–	0.12 (2.9) **	0.03 (0.5)
MdFS: Verankerungsindikator „Umfrageerfahrung"			
(9) ANREIZE • ERFAHRUNG	–	–	–0.32 (2.3) **
(10) EINSTELLUNG • ERFAHRUNG	–	–	–0.33 (1.4)
(11) ANREIZE • EINSTELLUNG • ERFAHRUNG	–	–	–0.46 (2.7) **
Konstante	1.92 (3.9) **	1.86 (3.9) **	1.89 (3.9) **
R^2	0.29	0.30	0.32
N	1477	1477	1477

Signifikanz: * $p < 0.1$; ** $p < 0.05$; Referenzkategorie: a) vier oder weniger Interviews in der Vergangenheit

stellung zu erwarten. Sind beide Vorbedingungen erfüllt, so wirkt sich die Richtung und Stärke der Gesamtanreize durch soziale Erwünschtheit nicht auf das Antwortverhalten aus. Dies führt auf der Ebene der statistischen Analyse zur Vorhersage einer Dreifach-Interaktion zwischen dem Anreizindex sozialer Erwünschtheit sowie dem Inhalt und der Stärke der generalisierten Umfrageeinstellung.

Die in Tabelle 5 dargestellten Regressionsergebnisse zeigen *erstens*, dass der Inhalt der Umfrageeinstellung ohne Berücksichtigung der in der Stichprobe vorliegenden Heterogenität der Einstellungsstärke keinen statistisch signifikanten Einfluss auf die Beeinflussbarkeit der Befragten durch soziale Erwünschtheit hat: Der Interaktionsparameter zwischen dem Anreizindex und der Umfrageeinstellung ist kein signifikanter Prädiktor der Einstellungsangaben (vgl. Modell 3). Die Ergebnisse zeigen aber *zweitens*, dass die Erklärungskraft der Gesamtanreize durch soziale Erwünschtheit durch die Haltung der Befragten zu Umfragen vorhergesagt werden kann, wenn die durch die Antwortgeschwindigkeit erfasste unterschiedliche kognitive Zugänglichkeit der Einstellungsinhalte in die Analyse einbezogen wird: Die Dreifach-Interaktion zwischen dem Anreizindex, der Umfrageeinstellung und der zur Nennung dieser Einstellungsinhalte benötigten Zeit hat einen signifikanten Vorhersagewert für das Antwortverhalten (vgl. Modell 4). Die Analyse zeigt *drittens*, dass die Erfahrung mit Umfragen, der im Vergleich zu den Antwortlatenzen validere Indikator für die Stärke der Umfrageeinstellung und der bessere Prädiktor für deren Fähigkeit zur Rahmung der Interviewsituation ist: Bei gleichzeitiger Berücksichtigung der Dreifach-Interaktion zwischen dem Anreizindex, der Umfrageeinstellung und dem Ausmaß der Interviewerfahrung verlieren die Antwortlatenzen vollständig ihre Vorhersagekraft (vgl. Modell 5).[11] Wird das Ausmaß der Erfahrung mit Umfragen als Indikator der Einstellungsstärke herangezogen, so lässt sich die Erklärungsleistung der im RCT prognostizierten

[11] Zusätzliche Analysen (Ergebnisse nicht berichtet) haben gezeigt, dass die Moderatorwirkung der Umfrageerfahrung erst ab einem Schwellenwert von mehr als vier Interviewerfahrungen sprunghaft ansteigt. In der vorliegenden Analyse geht die Interviewerfahrung daher als Dummy-Variable mit den Ausprägungen null (4 oder weniger Interviewerfahrungen: 66 % Befragte) und eins (5 oder mehr Erfahrungen: 34 % Befragte) ein.

Tabelle 6 Sensibilität des Antwortverhaltens über die Einstellung zu Ausländern für Anreize durch soziale Erwünschtheit nach dem Inhalt und der Verankerung der Umfrageeinstellung (Vorhersagewerte aus Regressionsmodell 4 (Antwortlatenzen) und Modell 5 (Umfrageerfahrung))

	Verankerungsindikator Antwortlatenzen			Verankerungsindikator Umfrageerfahrung		
	Anreize sozialer Erwünschtheit für eine . . .			*Anreize sozialer Erwünschtheit für eine . . .*		
	negative Einstellung	positive Einstellung	Anreiz-sensibilität	negative Einstellung	positive Einstellung	Anreiz-sensibilität
Umfrageeinstellung: Inhalt / Verankerung						
positiv / hoch	3.8	4.0	+ 0.2	4.4	4.0	– 0.4
positiv / gering	3.7	4.8	+ 1.1	3.3	4.4	+ 1.1
negativ / hoch	2.8	3.7	+ 0.9	2.9	4.1	+ 1.2
negativ / gering	2.9	3.6	+ 0.7	2.1	3.0	+ 0.9

Einstellungsangaben reichen von 1 (negative Einstellung) bis 7 (positive Einstellung). Die prognostizierten Antworten gelten für Befragte mit Anreizwerten von einer Standardabweichung unter (Anreize für negative Angaben) und über (Anreize für positive Angaben) dem Stichprobenmittelwert. Die Gruppierungsdimension Inhalt/Verankerung der Umfrageeinstellung repräsentiert Befragte mit allen Ausprägungskombinationen von einer Standardabweichung über und unter dem Stichprobenmittelwert der Dimensionen (linker Tabellenteil). Für die Umfrageerfahrung wurden Werte von 0 (4 oder weniger Befragungen) und 1 (mehr als 4 Befragungen) in die Vorhersagegleichung eingesetzt (rechter Tabellenteil). Kontrollvariablen sind bei kategorialen Variablen auf der Referenzkategorie und bei kontinuierlichen Faktoren auf dem Stichprobenmittelwert fixiert.

Bestimmungsfaktoren substantiell verbessern, wenn die vom MdFS zusätzlich angenommenen Entstehungsbedingungen des SD-Bias berücksichtigt werden ($F(5.2,1247.9) = 21.0$; $p \leq 0.01$).[12]

Im linken Teil von Tabelle 6 wird der Interaktionseffekt zwischen den Anreizen durch soziale Erwünschtheit, dem Inhalt der Umfrageeinstellung und deren Antwortlatenz dargestellt. Es ist feststellbar, dass Befragte mit einer eher kritischen Haltung zu Umfragen, unabhängig von der kognitiven Zugänglichkeit ihrer Urteile, in starkem Ausmaß durch soziale Erwünschtheit beeinflusst werden: Die Anreizsensibilität beträgt bei langsamer Antwortlatenz 0.7 und bei schneller Antwortgeschwindigkeit 0.9 Skalenpunkte. Bei Befragten mit einer überdurchschnittlich positiven Umfrageeinstellung ist dagegen die Einstellungsstärke bedeutsam: Zeigt eine zögerliche Beantwortung der Fragen nach der Umfrageeinstellung eine geringe kognitive Zugänglichkeit dieser Bewertungen an, so lässt sich ebenfalls ein starker SD-Bias von 1.1 Skalenpunkten beobachten. Wenn die Befragten allerdings über eine positive *und* kognitiv leicht verfügbare Umfrageeinstellung verfügen, so lassen sich mit 0.2 Skalenpunkten nur sehr schwache soziale Erwünschtheitseinflüsse feststellen.

Die im rechten Teil von Tabelle 6 dargestellten Ergebnisse bei Verwendung der Umfrageerfahrung als einem von den Antwortlatenzen konzeptionell komplett unabhängigem Indikator der Einstellungsstärke repliziert die bei den Antwortlatenzen beobachteten Zusammenhänge annähernd perfekt. Auch hier erweisen sich alle Befragten mit einer überdurchschnittlich negativen und solche mit einer positiven, aber auf wenig direkte Erfahrung beruhenden Einstellung, als stark durch soziale Erwünschtheit beeinflussbar: Das Antwortverhalten unterscheidet sich nach den Anreizen um 0.9 (negative Umfrageeinstellung, wenig Erfahrung), um 1.1 Skalenpunkte (negative Einstellung, viel Erfahrung) und um 1.2 Skalenpunkte (positive Einstellung, wenig Erfahrung). Nur wenn die Probanden Umfragen positiv gegenüberstehen und dieses Urteil zugleich auf häufige direkte Erfahrung beruht, widerstehen die Befragten der Versuchung zu Antwortverzerrungen in die Richtung der als sozial erwünscht wahrgenommenen Haltung zu Ausländern. Für diese Gruppe von Befragten findet sich sogar eine 0.4 Skalenpunkte starke Tendenz zur Umkehrung der Effekte. Diese unerwartete Umkehrung der Anreizwirkung lässt sich so interpretieren, dass die stark

[12] Das reduzierte Regressionsmodell enthält die Item-Dummies, die sozialstrukturellen Kontrollvariablen sowie den im Rahmen der RCT vorhergesagten Anreizindex. Der hierbei erzielte R^2-Wert von 0.25 steigt auf 0.31 an, wenn die Interaktion zwischen dem Anreizindex, der Umfrageeinstellung sowie der Umfrageerfahrung in das Regressionsmodell eingeführt werden (Ergebnisse nicht berichtet). Die Freiheitsgrade des F-Testes sind mit einem Greenhouse-Geissers Epsilon von 0.86 gewichtet.

zu korrekten Angaben motivierten Befragten dieser Teilstichprobe auf jeden Fall eine Beeinflussung durch soziale Erwünschtheit vermeiden wollen: Als Ergebnis resultiert eine Überkorrektur der in der sozialen Situation im Interview wahrgenommenen Ansprüche an das eigene Antwortverhalten. Es kann festgehalten werden, dass sich die Vorhersagen des MdFS darüber, wann sich die Anreize durch soziale Erwünschtheit auf die Einstellungsangaben auswirken und wann nicht, insgesamt empirisch gut bestätigen lassen.

6. Zusammenfassung und Diskussion

Nach den Ergebnissen des *ersten* Teils unserer empirischen Untersuchung kann die Vorhersage der RCT bestätigt werden, dass mit einem SD-Bias nur dann gerechnet werden muss, wenn die Befragten ein starkes Anerkennungsbedürfnis haben, wegen der Wahrnehmbarkeit der Antwortinhalte Sanktionen befürchten und deutliche Bewertungsunterschiede zwischen den Antwortmöglichkeiten wahrnehmen. In unserer Studie hat sich auch gezeigt, dass die Erklärungskraft jeder einzelnen Entstehungsbedingung des SD-Bias davon abhängt, ob alle anderen Vorbedingungen des SD-Bias erfüllt sind. So wirken sich etwa Unterschiede in der Privatheit der Antwortsituation trotz hohem Anerkennungsbedürfnis nicht auf die Einstellungsangaben aus, wenn die Antwortoptionen als gleich erwünscht wahrgenommen werden. Ein nur schwach ausgeprägtes Anerkennungsbedürfnis und eine gesicherte Privatheit der Antwortsituation haben den gleichen negativen Effekt auf die Erklärungskraft der jeweils anderen Faktoren für die Antwortinhalte.

Die vorliegenden Ergebnisse machen klar, dass bei einer angemessenen Analyse der Entstehungsbedingungen des SD-Bias alle Einzeldeterminanten gleichzeitig einbezogen und deren Interaktion berücksichtigt werden muss. Es hat sich gezeigt, dass eine in dieser Hinsicht unvollständige Untersuchungsanlage zu irreführenden Schlussfolgerungen über die Existenz und Richtung eines SD-Bias führen kann. So wirken sich in unserer Studie, im Unterschied zu anderen Ergebnissen, die verschiedenen Erhebungsmodi als Einzelfaktor *nicht* auf die Angaben über die Einstellungen zu Ausländern aus. Dieser Widerspruch lässt sich bei Berücksichtigung der Erwünschtheitswahrnehmungen auflösen. So konnten in unserer Stichprobe zwei, etwa gleich große normative Milieus mit Erwünschtheitswahrnehmungen und Antwortverzerrungen in unterschiedlichen Richtungen identifiziert werden, die sich auf der Aggregatebene gegenseitig aufheben. Bei der isolierten Analyse der Erklärungsfaktoren hat sich außerdem gezeigt, dass Befragte mit stärkerem Anerkennungsbedürfnis *negativere* Einstellungen zu Ausländern berichten. Die naheliegende Schlussfolgerung globaler Antwortverzerrungen in Richtung auf negative Einstellungsantworten erweist sich bei Berücksichtigung aller Erklärungsfaktoren als Fehlschluss: Ein zunehmendes Anerkennungsbedürfnis verstärkt vielmehr den SD-Bias in Übereinstimmung mit den unterschiedlichen Erwünschtheitswahrnehmungen der Befragten.

Obwohl die von der RCT vorhergesagten Entstehungsbedingungen des SD-Bias gut bestätigt werden konnten, erweisen sich diese in unserem *zweiten* Analyseteil als unvollständig. Demnach wirken sich die in der RCT definierten, instrumentell rationalen Anreize durch soziale Erwünschtheit genau unter der vom MdFS vorhergesagten Randbedingung *nicht* auf das Antwortverhalten aus. Diese Randbedingung ist eine positive *und* kognitiv stark verankerte Umfrageeinstellung der Befragten. Dieses Ergebnis findet sich gleichermaßen bei Verwendung von Antwortlatenzen und der Umfrageerfahrung als Indikatoren der kognitiven Zugänglichkeit der Umfrageeinstellung. Dagegen gelten die von der RCT vorhergesagten Determinanten des SD-Bias uneingeschränkt, wenn die Befragten eine eher kritische Haltung zu Umfragen haben oder wenn es sich bei der Umfrageeinstellung um „Non Attitudes“ handelt. Unsere Ergebnisse bestätigen somit die bereits bei der Erklärung der Entstehungsbedingungen normativen Handelns (Stocké 2002) und bei der Vorhersage ehelicher (In-) Stabilität (Esser 2002) festgestellte empirische Angemessenheit des MdFS.

Nach unseren Ergebnissen hat sich die Wahrnehmbarkeit der Antwortinhalte in der Interviewsituation als zentrale Entstehungsbedingung des SD-Bias erwiesen. Demnach sollte *erstens* die Privatheit der Interviewsituation durch die Verwendung selbst-administrierter Befragungsformen sichergestellt werden. Die hierbei fehlende motivierende und die Befragten unterstützende Funktion der Interviewer ist jedoch negativ für die resultierende Datenqualität (Lyberg/Kasprzyk 1991). Es ist daher ratsam, gezielt nur solche Umfragen oder Befragungsteile selbst-administriert durchzuführen, bei denen ein SD-Bias erwartet werden muss. Das zentrale Kriterium bei dieser Entscheidung sind die Erwünschtheitswahrnehmungen der Befragten. Ob und in welchem Umfang bei einem bestimmten Thema Erwünschtheitsunterschiede zwischen den Antwort-

möglichkeiten wahrgenommen werden und ob es hierbei Gruppenunterschiede gibt, wurde bisher nur wenig erforscht. Es wäre daher notwendig, so die *zweite* Schlussfolgerung unserer Untersuchung, zumindest für häufig eingesetzte Fragebogeninstrumente der angewandten Umfrageforschung empirisch ermittelte Kennwerte der Erwünschtheitswahrnehmungen zur Verfügung zu stellen. Obwohl unsere Analyse der Allbus-Ausländeritems hierzu ein erster Schritt ist, können die festgestellten Ergebnisse wegen der relativ geringen Fallzahl und der regional begrenzten Stichprobe nicht einfach auf bundesdeutsche Verhältnisse generalisiert werden.

Die vorliegende Untersuchung hat *drittens* deutlich gemacht, dass die Einstellung zu Umfragen und die Umfrageerfahrung wichtige Entstehungsbedingungen des SD-Bias sind. In anderen Studien hat sich gezeigt, dass die qualitative Bewertung vorangegangener Befragungen ein wichtiger Bestimmungsfaktor der Umfrageeinstellung ist (Stocké/Langfeldt 2004). Demnach prägt die Umfrageteilnahme den Inhalt der Umfrageeinstellung und vermittelt über die Zunahme der Einstellungsstärke, so das Ergebnis unserer Untersuchung, gleichzeitig auch die Bedeutung dieser Einstellung für die Entstehung des SD-Bias. Demnach sollten Umfrageforscher bei der Gestaltung der Interviews, etwa in Hinblick auf deren Länge und den Schwierigkeitsgrad der Fragen, beachten, dass sich die Belastung der Befragten auf deren zukünftige Empfänglichkeit für Anreize durch soziale Erwünschtheit und damit auf die realisierbare Datenqualität auswirkt.

Literatur

Aiken, L.S. / West, S.G., 1991: Multiple Regression: Testing and Interpreting Interactions. Newbury Park: Sage.

Aquilino, W.S., 1994: Interview Mode Effects in Surveys of Drug and Alcohol Use. Public Opinion Quarterly 58: 210–240.

Bassili, J.N. / Bors, D.A., 1997: Using Response-Latency to Increase Lead Time in Election Forecasting. Canadian Journal of Behavioural Science 29: 231–238.

Brunet, A. / Boucher, C. / Boyer, R., 1996: Social Desirability in the Assessment of Trauma. Psychological Reports 79: 511–514.

Chen, P.Y. / Dai, T. / Spector, P.E. / Jex, S.M., 1997: Relation Between Negative Affectivity and Positive Affectivity: Effects of Judged Desirability of Scale Items and Respondents' Social Desirability. Journal of Personality Assessment 69: 183–198.

Cronbach, L.J., 1987: Statistical Tests for Moderator Variables: Flaws in Analyses Recently Proposed. Psychological Bulletin 102: 414–417.

Crowne, D.P., 1979: The Experimental Study of Personality. New York: Wiley.

Davies, M.F. / French, C.C. / Keogh, E., 1998: Self-Deceptive Enhancement and Impression Management Correlates of EPQ-R Dimensions. Journal of Psychology. Interdisciplinary & Applied 132: 402–408.

DeMaio, T.J., 1984: Social Desirability and Survey Measurement: A Review. S. 257–282 in: C.F. Turner / E. Martin (Hrsg.), Surveying Subjective Phenomena. New York: Russell Sage Foundation.

Diekmann, A., 1996: Homo ÖKOnomicus. Anwendungen und Probleme der Theorie rationalen Handelns im Umweltbereich. Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 36: 89–118.

Duck, R.J. / Hunsberger, B., 1999: Religious Orientation and Prejudice: The Role of Religious Proscription, Right-Wing Authoritarianism. The International Journal for the Psychology of Religion 9: 157–179.

Edwards, A.L., 1957: The Social Desirability Variable in Personality Assessment and Research. New York: Dryden Press.

Erbslöh, B. / Koch, A., 1988: Die Non-Response-Studie zum Allbus 1986: Problemstellung, Design, erste Ergebnisse. ZUMA-Nachrichten 22: 29–44.

Esser, H., 1991: Die Erklärung systematischer Fehler in Interviews: Befragtenverhalten als „rational choice". S. 59–78 in: R. Wittenberg (Hrsg.), Person – Situation – Institution – Kultur. Günter Büschges zum 65. Geburtstag. Berlin: Duncker & Humblot.

Esser, H., 2001: Soziologie. Spezielle Grundlagen. Band 6: Sinn und Kultur. Frankfurt a. M.: Campus.

Esser, H., 2002: Ehekrisen: Das (Re-)Framing der Ehe und der Anstieg der Scheidungsraten. Zeitschrift für Soziologie 31: 472–496.

Fazio, R.H., 1990: A Practical Guide to the Use of Response Latency in Social Psychological Research. S. 74–97 in: C. Hendrick / M.S. Clark (Hrsg.), Review of Personality and Social Psychology: Vol. 11. Research Methods in Personality and Social Psychology. Newbury Park, CA: Sage.

Fisher, R.J. / Katz, J.E., 2000: Social-Desirability Bias and the Validity of Self-Reported Values. Psychology & Marketing 17: 105–120.

Huang, C. / Liao, H. / Chang, S.-H., 1998: Social Desirability and the Clinical Self-Report Inventory: Methodological Reconsideration. Journal of Clinical Psychology 54: 517–528.

Jobe, J.B. / Pratt, W.F. / Tourangeau, R. / Baldwin, A.K. / Rasinski, K.A., 1997: Effects of Interview Mode on Sensitive Questions in a Fertility Survey. S. 311–329 in: L. Lyberg / P. Biemer / M. Collins / E. De Leeuw / C. Dippo / N. Schwarz / D. Trewin (Hrsg.), Survey Measurement and Process Quality. New York: Wiley.

Krysan, M., 1998: Privacy and the Expression of White Racial Attitudes. Public Opinion Quarterly 62: 506–544.

Krysan, M. / Schuman, H. / Scott, L.J. / Beatty, P., 1994: Response Rates and Response Content in Mail Versus Face-to-Face Surveys. Public Opinion Quarterly 58: 381–399.

Lyberg, L. / Kasprzyk, D., 1991: Data Collection Methods and Measurement Error: An Overview. S. 237–257 in:

P.P. Biemer / R.M. Groves / L.E. Lyberg / N.A. Mathiowetz / S. Sudman (Hrsg.), Measurement Errors in Surveys. New York: Wiley.

Mangione, T.W. / Hingson, R. / Barrett, J., 1982: Collecting Sensitive Data. A Comparison of Three Survey Strategies. Sociological Methods & Research 10: 337–346.

Mielke, R., 1995: Der Bildungsgrad und die Einstellung gegenüber Ausländern. Bielefelder Arbeiten zur Sozialpsychologie 173. Bielefeld: Universität Bielefeld.

Nicotera, A.M., 1996: An Assessment of the Argumentativeness Scale For Social Desirability Bias. Communication Reports 9: 23–36.

Paulhus, D.L., 1984: Two-Component Models of Socially Desirable Responding. Journal of Personality and Social Psychology 46: 598–609.

Phillips, D.L. / Clancy, K.J., 1970: Response Bias in Field Studies of Mental Illness. American Sociological Review 35: 503–514.

Phillips, D.L. / Clancy, K.J., 1972: Some Effects of „Social Desirability“ in Survey Studies. American Journal of Sociology 77: 921–940.

Reinecke, J., 1991: Interviewer- und Befragtenverhalten. Theoretische Ansätze und methodische Konzepte. Opladen: Westdeutscher Verlag.

Reuband, K.-H. / Blasius, J., 1996: Face-to-Face-, telephonische und postalische Befragungen. Ausschöpfungsquoten und Antwortmuster in einer Großstadt-Studie. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48: 296–318.

Rogelberg, S.G. / Fisher, G.G. / Maynard, D.C. / Hakel, M.D. / Horvath, M., 2001: Attitudes Toward Surveys: Development of a Measure and its Relationship to Respondent Behavior. Organizational Research Methods 4: 3–25.

Schuette, R.A. / Fazio, R.H., 1995: Attitude Accessibility and Motivation as Determinant of Biased Processing: A Test of the MODE Model. Personality and Social Psychology Bulletin 21: 704–710.

STATA Corporation, 1999: STATA. Reference Manual. Release 6. Vol. 3. College Station: Stata Press.

Stevens, J., 1996: Applied Multivariate Statistics for the Social Sciences, (Third Edition). Mahwah, N.J.: Erlbaum.

Stocké, V., 2001: Socially Desirable Response Behavior as Rational Choice: The Case of Attitudes Towards Foreigners. Arbeitspapier No. 01–22: Sonderforschungsbereich 504, Universität Mannheim.

Stocké, V., 2002: Framing und Rationalität. Die Bedeutung der Informationsdarstellung für das Entscheidungsverhalten. München: Oldenbourg.

Stocké, V., 2003: Deutsche Kurzskala zur Erfassung des Bedürfnisses nach sozialer Anerkennung. in: A. Glöckner-Rist (Hrsg.), ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente. Mannheim: Zentrum für Umfragen, Methoden und Analysen.

Stocké, V., 2004: The Interdependence of Determinants for the Strength and Direction of Social Desirability Bias in Racial Attitude Surveys. Arbeitspapier No. 04-14: Sonderforschungsbereich 504, Universität Mannheim.

Stocké, V. / Langfeldt, B., 2004: Effects of Survey Experience on Respondents' Attitudes towards Surveys. Bulletin de Methodologie Sociologique 81: 5–32.

Tourangeau, R. / Rips, L.J. / Rasinski, K., 2000: The Psychology of Survey Response. Cambridge: Cambridge University Press.

Vella-Broderick, D.A. / White, V., 1997: Response Set of Social Desirability in Relation to the Mental, Physical and Spiritual Well-being Scale. Psychological Reports 81: 127–130.

Wegener, D.T. / Downing, J. / Krosnick, J.A. / Petty, R.E., 1995: Measures and Manipulations of Strength-Related Properties of Attitudes: Current Practice and Future Directions. S. 455–487 in: R.E. Petty / J.A. Krosnick (Hrsg.), Attitude Strength. Antecedents and Consequences. Mahwah, N.J.: Erlbaum.

Wyner, G.A., 1980: Response Errors in Self-Reported Number of Arrests. Sociological Methods & Research 9: 161–177.

Summary: Rational choice theory and its generalization in the model of frame-selection both claim to be exhaustive explanations of socially desirable response behavior in surveys, which allows researchers to take the interaction of explanatory factors into account. The determinants predicted in these approaches have been tested only incompletely and, in part, with inconsistent results. The following article analyzes some of the main hypotheses of both approaches, using responses about attitudes towards ethnic outgroups. In the first step the analysis confirms the predictions from straightforward rational choice theory. Accordingly, socially desirable response behavior only is to be expected when the motivational, cognitive and social preconditions are fulfilled simultaneously. This already very differentiated explanation proved to be incomplete when the additional cultural factors assumed in the model of frame-selection were analyzed in a second step. These determinants for the respondents' tendency toward a "cooperative" framing of the interview permits an identification of those conditions when instrumentally rational incentives are completely eliminated by the factor of social desirability. According to our results the validity of rational choice theory is restricted to important but specific conditions which are predicted in the model of frame-selection.