

Konstruktion und Validierung einer Skala zur relativen Messung von physischer Attraktivität mit einem Item

Das Attraktivitätsrating 1 (AR1)

Development and validation of a single-item scale for the relative assessment of physical attractiveness

The Attractiveness Rating 1 (AR1)

Johannes Lutz, Christoph J. Kemper, Constanze Beierlein, Jutta Margraf-Stiksrud und Beatrice Rammstedt

Zusammenfassung

Physische Attraktivität ist ein wichtiger Faktor sozialer Interaktion und Kognition, z.B. bei der Partnerwahl oder dem Eindruck, den sich Menschen von sich selbst und von anderen bilden. Auch im Rahmen von Umfragen könnte die physische Attraktivität sozial- und gesundheitswissenschaftliche Erklärungsmodelle sinnvoll ergänzen. So finden sich unter anderem Zusammenhänge mit Kriterien des Berufserfolgs, diversen Gesundheitsvariablen und der allgemeinen Lebenszufriedenheit. Die Messung physischer Attraktivität mittels absoluten Einschätzungen hat sich jedoch als problematisch erwiesen. Attraktivitätsurteile können z.B. von Charakteristika der Beurteiler beeinflusst werden. Vor allem die Interaktion zwischen Alter und Geschlecht von Beurteiler und Zielperson wird in der Literatur als Einflussfaktor berichtet. Ziel

Abstract

Physical attractiveness is an important factor of social interaction and cognition, such as in mate choice or in the impressions people form of themselves and of others. In the context of social and health sciences, physical attractiveness has also been shown to complement the prediction of various target variables. Among other relationships, physical attractiveness is associated with criteria of professional success, various health variables and overall life satisfaction. However, the measurement of physical attractiveness via absolute ratings has several shortcomings. For example, attractiveness judgments are influenced by characteristics of the rater. As pointed out in literature, especially the interaction between age and gender of rater and target person has a significant impact on rating scores. The aim of this study was to construct a scale for the



der vorliegenden Studie war daher die Konstruktion einer Skala zur ökonomischen, reliablen und validen Messung von physischer Attraktivität, die den in der Literatur berichteten Problemen Rechnung trägt. Um dies zu erreichen wird im Attraktivitätsrating 1 (AR1) ein relatives Messkonzept eingesetzt. Das AR1 wurde anhand von vier alters-, geschlechts- und bildungsheterogenen Stichproben konstruiert und validiert. Die berichteten Ergebnisse sprechen dafür, dass das AR1 das Merkmal physische Attraktivität adäquat abbildet.

economical, reliable and valid assessment of physical attractiveness, controlling for methodological problems reported in prior research. To achieve this, the Attractiveness Rating 1 (AR1) employs a relative measurement approach. The AR1 was constructed and validated using four diverse samples. The presented results suggest that the AR1 adequately measures physical attractiveness.

1 Theoretischer Hintergrund

Physische Attraktivität ist ein wichtiges Personenmerkmal zur Beschreibung und Erklärung menschlichen Erlebens und Verhaltens. Unter physischer Attraktivität wird die Attraktivität einer Person verstanden, die auf ihren körperlichen Merkmalen beruht (Asendorpf 2011, S.63). Die physische Attraktivität einer Person wird durch den Betrachter aus Merkmalen des Gesichts, wie der Größe der Nase, der Augen und des Kinns (Kindchenschema, vgl. Lorenz 1943) und aus der Figur, z.B. dem Taille-Hüfte-Verhältnis, erschlossen. Sie hat weitreichende Implikationen für die zwischenmenschliche Interaktion. Die physische Attraktivität ist wesentlicher Bestandteil des Selbstkonzepts und beeinflusst, wie sich eine Person selbst erlebt und somit auch wie sie sich anderen gegenüber verhält. Darüber hinaus beeinflusst die physische Attraktivität auch wie eine Person von ihren Interaktionspartnern gesehen und behandelt wird. Physisch attraktivere Menschen werden weniger attraktiven oder gar körperlich fehlgebildeten Menschen gegenüber positiver eingeschätzt und in der sozialen Interaktion, z.B. bei der Partnerwahl oder bei Bewerbungen, bevorzugt. Bestrebungen, dem gesellschaftlichen Schönheitsideal durch den Kauf modischer Kleidung und kosmetischer Produkte, durch Diäten, Sport, Sonnenbaden, Bodybuilding usw. näher zu kommen, erscheinen aufgrund der vorteilhaften Wirkung eines attraktiven Äußeren nur allzu verständlich. Aufgrund ihrer Rolle für das menschliche Erleben und Verhalten bei einer Vielzahl sozialer Phänomene und Prozesse ist die physische Attraktivität für die sozialwissenschaftliche Forschung und darüber hinaus ein hochgradig relevantes Konstrukt. Ein ökonomisches und reliables Verfahren für die Messung dieses Konstrukts zu entwickeln und der Forschung zur Verfügung zu stellen, ist Ziel der hier vorgestellten Arbeit.

Der Eindruck physischer Attraktivität entsteht in der sozialen Interaktion. Das Wahrnehmen von Proportionen, Haarfarbe, Frisur und mimischen Ausdrucks des Gesichts, löst auf Seiten des Betrachters einer Person zahlreiche Einschätzungsvorgänge aus, die in einer Zuschreibung von physischer Attraktivität resultieren. Dabei spielen insbesondere Gesichts- und Figurmerkmale eine Rolle (Gallup/Frederick 2010). Die Wahrnehmung physischer Attraktivität wird bspw. durch die Körperkomposition und die Beschaffenheit der Haut beeinflusst. So werden Frauen mit niedrigem Taille-Hüfte-Verhältnis und besonders ebenmäßiger Haut als attraktiver wahrgenommen (Fink et al. 2001; Singh et al. 2010). Bei Männern wirkt ein höheres Schulter-Hüfte-Verhältnis und moderat ausgeprägte Muskulatur attraktiv (Dijkstra/Buunk 2001; Frederick/Haselton 2007). Neben den genannten Merkmalen kommt dem Gesicht eine besondere Bedeutung für die physische Attraktivität zu. Dies wird bereits durch das heute nur noch selten verwendete Synonym „Antlitz“ impliziert. Es leitet sich aus „entgegenblicken“ ab. „Ein Gesicht wirkt dadurch, wie es dem Betrachter entgegenblickt“ (Margraf-Stiksrud 1991, S. 11). Die beiden Aspekte des Gesichts, die Mimik als dynamische Qualität und die Physiognomie als statische Qualität, sind wichtige Informationsquellen in der sozialen Interaktion (Margraf-Stiksrud 1991; Sergl 1991). So dient der mimische Ausdruck primär der Aussendung von Signalen über Stimmungen und innere Zustände und somit der zwischenmenschlichen Kommunikation. Die Physiognomie vermittelt Interaktionspartnern ebenfalls Informationen, auch z.B. über die Fruchtbarkeit einer Person (vgl. Rhodes 2006). Die Ausbildung bestimmter physiognomischer Merkmale, z.B. markantes Kinn, vorspringende Backenknochen, Gesichtsbehaarung bei Männern und volle Lippen, hohe Wangenknochen, schmale Wangen bei Frauen, wird durch Geschlechtshormone gesteuert, die auch als Indikatoren für ein starkes Immunsystem gelten (vgl. Rhodes 2006). Aus evolutionsbiologischer Perspektive lässt sich argumentieren, dass diese Merkmale biologisch verankerte Signale für Fruchtbarkeit darstellen, deren Erkennen daher einen Selektionsvorteil darstellt. Man vermutet, dass sie aus diesem Grund auch als attraktiv wahrgenommen werden (Bierhoff 2000; Rhodes 2006). Neben den oben genannten Merkmalen, die sich bei Männern und Frauen im Hinblick auf die Zuschreibung von physischer Attraktivität unterscheiden (Geschlechtsdimorphismus), gelten bei beiden Geschlechtern Symmetrie der Gesichtshälften und Durchschnittlichkeit der physiognomischen Merkmale als relevante Prädiktoren für physische Attraktivität (Rhodes 2006).

Die Auswirkungen der physischen Attraktivität auf Betrachter und Betrachtete sind in vielen Bereichen substantiell. Viele interpersonelle Prozesse werden durch physische Attraktivität beeinflusst. Zahlreiche Studien belegen, dass Attraktive im Vergleich zu weniger Attraktiven in der sozialen Interaktion oft bevorzugt

werden und Vorteile haben: Attraktivere bekommen häufiger als weniger Attraktive gut bezahlte berufliche Positionen mit hohem Prestige (z.B. Schuler/Berger 1979; Umberson/Hughes 1987). Attraktivere Kinder werden seltener von Lehrern diszipliniert (Clifford/Walster 1973). Attraktivere Kriminelle erhalten mildere Urteile als weniger Attraktive (Sigall/Ostrove 1975). Attraktiveren wird mehr Aufmerksamkeit geschenkt und ihnen wird häufiger geholfen (Langlois et al. 2000). Attraktive haben weiterhin diverse Vorteile in der sozialen Interaktion, die dem Aufbau intimer Beziehungen zuträglich sind (Tramitz 2000). Solche Befunde lassen sich zum Teil auf ein Attraktivitätsstereotyp zurückführen. Menschen neigen dazu, physisch Attraktiven sozial erwünschte Eigenschaften wie „interessant“, „warm“, „unabhängig“ oder „ehrlich“ zuzuschreiben und außerdem mehr Lebensglück, beruflichen Erfolg und soziale Kompetenz (Bierhoff 2000; Eagly et al. 1991). Neben diesen Zuschreibungen, gemäß der Heuristik „schön ist gut“, weisen Attraktive aber tatsächlich viele sozial erwünschte Eigenschaften auf. Sie sind z.B. geselliger und beliebter (Langlois et al. 2000). Außerdem steht physische Attraktivität in positivem Zusammenhang mit der allgemeinen Gesundheit (vgl. Thornhill/Gangestad 1999), der Anzahl der Nachkommen (Jokela 2009), der Intelligenz (Langlois et al. 2000; Kanazawa/Kovar 2004), dem durchschnittlichen Einkommen (Frieze et al. 1991; Judge et al. 2009) und verschiedenen anderen Kriterien des Berufserfolgs (Hamermesh/Biddle 1994; Hosoda et al. 2003).

Während Attraktive häufiger positive Reaktionen aus ihrem sozialen Umfeld erfahren und Vorteile in der sozialen Interaktion genießen, erfahren Menschen mit geringer Attraktivität zum Teil negative Reaktionen. Dies trifft besonders für Menschen zu, die stark von der Idealnorm abweichen, die z.B. Fehlbildungen im Gesicht aufweisen. Menschen mit normabweichenden physiognomischen Merkmalen werden oft mit Vorurteilen konfrontiert und in der sozialen Interaktion benachteiligt (siehe Heimes/Kemper 2001): Zu Menschen mit entstellenden Gesichtszügen wird eine größere soziale Distanz gehalten, z.B. beim Warten an der Bushaltestelle (Fleischer-Peters/Margraf-Stiksrud 1996). Houston und Bull (1994) beobachteten, dass Plätze in der Straßenbahn neben Menschen mit einem Feuermal (Nevus flammeus) im Gesicht häufiger gemieden werden. Neben einer größeren sozialen Distanz wurden auch Vorurteile gegenüber Personen mit Fehlbildungen berichtet, z.B. bei Personen mit schlecht operierter Lippen-Kiefer-Gaumenspalte (Bernstein 1982; Sergl 1991). Für die Vorurteile und das ablehnende Verhalten lassen sich verschiedene Gründe anführen. Unter anderem gehen die Normabweichungen in der Physiognomie mit einer Mimik einher, die dem Beobachter das Erkennen ausgesendeter Signale erschwert. Daraus resultiert eine Kommunikation, die von einem Inter-

aktionspartner nicht mehr eindeutig interpretiert werden kann (Fleischer-Peters/Margraf-Stiksrud 1996).

Die berichteten Befunde zeigen, dass dem Gesicht eine bedeutende Rolle bei der zwischenmenschlichen Begegnung zukommt. Dies scheint im Besonderen für das attraktive und „normkonforme“ Gesicht zu gelten, da es wichtige Voraussetzungen für eine reibungslose Kommunikation schafft. Das Gesicht daher bei der Messung des Konstrukts physische Attraktivität ins Zentrum zu stellen, liegt demnach nahe. Die Praxis der Attraktivitätsmessung sieht aktuell allerdings anders aus. Üblicherweise wird physische Attraktivität mittels mehrstufiger Ratingskalen mit verbalen Ankern, z.B. von „gar nicht attraktiv“ bis „sehr attraktiv“ gemessen (Langlois et al. 2000). Dabei wird von den Beurteilern eine globale Einschätzung gefordert, d.h. es wird nicht spezifiziert, welcher Aspekt von Attraktivität beurteilt werden soll. Welche äußerlichen Merkmale der Beurteiler letzten Endes berücksichtigt, bleibt dabei unklar. Diese Art der Messung, die z.B. im ALLBUS (Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften) eingesetzt wird, weist daher nicht nur eine Unterbetonung der Bedeutung des Gesichts auf, sondern auch eine mangelnde Standardisierung der Attraktivitätseinschätzung.

Weiterhin gehen mit der Messung von physischer Attraktivität Herausforderungen einher, die durch diese Art der Messung nicht adäquat adressiert werden. Attraktivitätseinschätzungen hängen von verschiedenen Merkmalen des Beurteilers ab, z.B. von seinem Vergleichsstandard, seinem Geschlecht und seinem Alter (Bierhoff 2000; Henss 1991; Hönekopp 2006; Nedelec/Beaver 2011). Männer, die kurz zuvor Filme mit attraktiven Frauen oder Aktfotographien angesehen haben, beurteilten durchschnittlich attraktive Frauen und ihre eigenen Frauen als weniger attraktiv im Vergleich zu Männern, bei denen der Vergleichsmaßstab nicht experimentell manipuliert wurde (Kontrasteffekt; Kenrick/Gutierrez 1980; Kenrick et al. 1989). Nedelec und Beaver (2011) berichten geschlechtsabhängige Unterschiede in der Attraktivitätseinschätzung von Männern. Während Männer bei der Einschätzung von Männern eher zu mittleren Urteilen neigen, beurteilen Frauen die Männer häufiger entweder als „sehr attraktiv“ oder „sehr unattraktiv“. Der generell negative Zusammenhang zwischen Alter und physischer Attraktivität wird ebenfalls durch das Geschlecht und auch das Alter des Beurteilers beeinflusst (Henss 1991). Bei der Beurteilung von Personen des eigenen Geschlechts oder der eigenen Altersgruppe fällt der Zusammenhang zwischen Alter und physischer Attraktivität signifikant niedriger aus. Der Effekt, dass Gesichter von Personen der eigenen Altersgruppe besser wiedererkannt (und somit auch anders eingeschätzt) werden als Gesichter von älteren oder jüngeren Menschen, wird als „Own-Age-Bias“ (OAB) bezeichnet (vgl. Bartlett/Fulton 1991; Rhodes/Anastasi 2012).

Diese Befunde verdeutlichen, welche Verzerrungen bei der üblichen Messung ohne Vergleichsmaßstab (fortan: „absolute“ Attraktivitätseinschätzung) auftreten können. Ziel der hier vorgestellten Arbeit war es, eine alternative Ratingskala (Attraktivitätsrating 1, AR1) zu konstruieren, die den oben beschriebenen Verzerrungen Rechnung trägt und diese minimiert. Zudem sollte diese optimal auf die Erfordernisse der sozialwissenschaftlichen Forschung ausgerichtet sein, d.h. eine möglichst ökonomische aber dennoch ausreichend reliable und valide Messung des Konstrukts erlauben. Dazu wird im AR1 ein „relatives Messkonzept“ verwendet, bei dem die Einschätzung der Attraktivität einer Person (fortan auch Zielperson) an einem Vergleichsmaßstab erfolgt. Als Vergleichsmaßstab dient ein „Ankerbild“ mit dem Gesicht einer durchschnittlich attraktiven Person (ungefähr) desselben Alters und desselben Geschlechts wie die Zielperson. Die Attraktivität der Zielperson soll vom Beurteiler relativ zum Ankerbild eingeschätzt werden. Durch dieses relative Messkonzept wird (1) die Grundlage von Attraktivitätseinschätzungen vereinheitlicht und die Vergleichbarkeit über Beurteiler erhöht, (2) der besonderen Rolle des Gesichts Rechnung getragen und (3) Verzerrungen der Attraktivitätsmessung durch Merkmale der Beurteiler, wie deren individuelle Präferenzen, deren Alter und deren Geschlecht, minimiert. Im Folgenden werden Konstruktion und empirische Überprüfung des AR1 dargestellt.

2 Methode

2.1 Stichproben

Das AR1 wurde anhand von vier alters-, geschlechts- und bildungsheterogenen Stichproben entwickelt und validiert. Details zur Zusammensetzung der Stichproben sind in Tabelle 1 dargestellt. Die Erhebung der Stichproben 1, 2 und 3 wurden von unabhängigen kommerziellen Anbietern durchgeführt.

Stichprobe 1 ist eine Quotenstichprobe, geschichtet nach den Merkmalen Geschlecht, Alter, Bildung und Bundesland ($N = 539$). Die Grundgesamtheit war definiert als „alle in der Bundesrepublik Deutschland in Privathaushalten lebenden deutschsprachigen Personen ab 18 Jahren.“ Die Daten wurden im Rahmen eines persönlich-mündlichen Interviews (CAPI, Computer Assisted Personal Interview) und durch die Vorgabe eines Papierfragebogens erhoben. Die Erhebung dauerte insgesamt im Mittel 53 Minuten ($SD = 12$).

Stichprobe 2 wurde im Internet erhoben (CAWI, Computer Assisted Web Interview). Es handelt es sich um eine Quotenstichprobe, geschichtet nach Geschlecht, Alter und Bildung ($N = 741$). Grundgesamtheit waren die Teilnehmer eines Online-

Tabelle 1 Charakteristika der drei Stichproben

	Stichprobe 1	Stichprobe 2	Stichprobe 3	Stichprobe 4	
<i>Stichprobe</i>					
Umfang [N]	539	741	1.134	56	
Art	Quote	Quote	Zufall	Gelegenheit	
Modus	CAPI, Papier	CAWI	CAPI, CASI	Papier	
<i>Zusammensetzung</i>					
Geschlecht [% Frauen]	52,5%	51,8%	55,6%	82%	
Alter [M(SD)]	47.2 (15.2)	48.3 (13.0)	53.3 (18.4)	24.4 (5.9)	
Bildung	≤ 9 Jahre	44,7%	40,1%	37,2%	-
	10 Jahre	30,2%	29,1%	37,0%	-
	≥ 11 Jahre	23,7%	30,8%	25,8%	100%

CAPI = Computer Assisted Personal Interview, CAWI = Computer Assisted Web Interview, CASI = Computer Assisted Self Interview, Papier = Papierversion (Selbstaufüller).

Access-Panels im Alter von 18 Jahren oder älter, die in Deutschland leben. Die Bearbeitungsdauer des Fragebogens betrug im Mittel 23 Minuten ($SD = 8$).

Stichprobe 3 mit $N = 1.134$ Befragungspersonen ist eine Zufallsstichprobe, die repräsentativ für die Wohnbevölkerung in Deutschland über einem Alter von 18 Jahren ist. Sie wurde mithilfe des ADM-Stichprobensystem Face-to-Face (Random Route) der Arbeitsgemeinschaft deutscher Marktforschungsinstitute gezogen. Die Daten wurden im CAPI- und CASI-Modus (Computer Assisted Self Interview) erhoben. Die Erhebung dauerte durchschnittlich 43 Minuten ($SD = 13$).

Stichprobe 4 ($N = 56$) wurde im Rahmen von mehreren Lehrveranstaltungen des psychologischen Instituts der Universität Potsdam erhoben. Die Bearbeitungszeit des Papierfragebogens betrug maximal 2 Minuten.

2.2 Material

2.2.1 FACES-Datenbank

Als Ausgangspunkt für die Skalenkonstruktion wurde die FACES-Datenbank des Max-Planck-Instituts für Bildungsforschung verwendet (Ebner et al. 2010). FACES ist eine Sammlung von Portraitfotos von 171 Männern und Frauen dreier Altersklassen (junges, mittleres und hohes Alter). Von jeder der 171 Personen gibt es je zwei Fotos mit neutralem, traurigem, angewidertem, ängstlichem, wütendem und glücklichem Gesichtsausdruck. Durch die Vorselektion einer Modellagentur wur-

de sichergestellt, dass die in der Datenbank abgebildeten Personen (fortan auch: Modelle) möglichst durchschnittlich aussehen und keine auffälligen Merkmale wie Tätowierungen oder Piercings aufweisen.

2.2.2 AR1

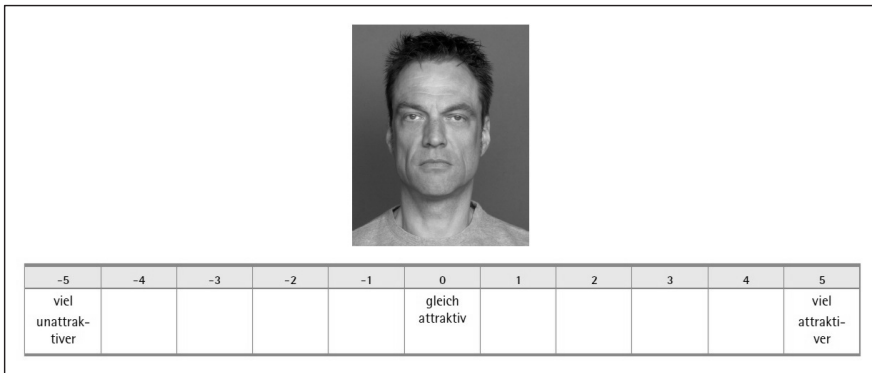
Das AR1 zur Einschätzung der Attraktivität besteht aus einem Item. Zusammen mit einer elfstufigen Antwortskala wird das Bild einer durchschnittlich attraktiven Person des gleichen Geschlechts und der gleichen Altersgruppe wie die zu beurteilende Person dargeboten (Ankerbild). Die Einschätzung der Attraktivität der Zielperson erfolgt relativ zu diesem Vergleichsmaßstab. Die elf Antwortkategorien des AR1 reichen von „viel unattraktiver“ (-5) bis „viel attraktiver“ (5) mit einer neutralen Mittelkategorie „gleich attraktiv“ (0), über der das Ankerbild dargeboten wird (siehe Abbildung 1). Die Instruktion an den Beurteiler lautet: „Bitte schätzen Sie die Attraktivität der Befragungsperson ein. Geben Sie an, wie attraktiv oder unattraktiv der Befragte im Vergleich zu der unten abgebildeten Person ist.“ (Instruktion für die Fremdeinschätzung; bei einer Selbsteinschätzung muss die Instruktion entsprechend angepasst werden).

Insgesamt gibt es sechs Skalenvarianten, die sich nur im Ankerbild unterscheiden. Jede enthält ein Ankerbild aus der FACES-Datenbank, das dem Geschlecht (männlich oder weiblich) und dem Alter (junges, mittleres oder hohes Alter) der zu beurteilenden Zielperson bestmöglich entspricht. Die Bilder mit den folgenden IDs wurden aus Set A der Datenbank ausgewählt: Nr. 167 als Anker für die Beurteilung von jungen Männern, Nr. 054 für die Beurteilung von jungen Frauen, Nr. 178 für die Beurteilung mittelalter Männer, Nr. 080 für die Beurteilung mittelalter Frauen, Nr. 027 für die Beurteilung älterer Männer und Nr. 133 für die Beurteilung älterer Frauen. Aufgrund des Copyrights können die sechs Ankerbilder, die als Grundlage der Skalenvarianten dienen, in dieser Publikation nicht abgedruckt werden. Sie können über die Onlinepräsenz des Max-Planck-Instituts einfach und kostenlos angefordert werden (<http://faces.mpdl.mpg.de>). Eine Anleitung dazu findet sich bei Kemper et al. (2012).

2.2.3 Fragebögen

Der Fragebogen von Stichprobe 1 enthielt soziodemographische Angaben, verschiedene psychologische und sozialwissenschaftliche Skalen und 36 Bilder von Modellen aus der FACES-Datenbank. Die Attraktivität der Modelle wurde auf einer Skala von „gar nicht attraktiv“ (1) bis „sehr attraktiv“ (7) beurteilt. Der Fragebogen

Abbildung 1 Attraktivitätsrating 1 mit Beispiel-Ankerbild



Anmerkung: Das Ankerbild in Abbildung 1 ist lediglich ein Beispiel. Die Skala darf in dieser Form nicht eingesetzt werden. Alle notwendigen Informationen zur Erstellung der sechs Skalenvarianten des AR1 sind in einem Arbeitsbericht von Kemper et al. (2012) zu finden.

in Stichprobe 2 enthielt soziodemographische Angaben, verschiedene Skalen zur Erfassung psychologischer Merkmale und 18 Bilder von Modellen, deren Attraktivität mit dem AR1 eingeschätzt werden sollte. Die Fragebogenbatterie in Stichprobe 3 beinhaltete neben dem AR1 umfangreiche soziodemographische Angaben und verschiedene psychologische und sozialwissenschaftliche Maße, z.B. zur Erfassung von Selbstwirksamkeit (AKSU; Beierlein et al. – dieses Heft), kristalliner Intelligenz (BEFKI GC-K; Schipolowski et al. – in diesem Heft), positiver und negativer Affektivität (Menold/Kemper 2013), physischer und psychischer Beeinträchtigung der Gesundheit (Andersen/Mühlbacher/Nübling 2007), Einkommen und Sozialstatus. Das AR1 war zweimal im Fragebogen enthalten. Am Anfang des Interviews (CAPI) nahm der Interviewer eine Einschätzung der Attraktivität der Befragungsperson vor. Am Ende des Interviews folgte ein kurzer Teil des Fragebogens im CASI-Modus. In diesem sollte die Befragungsperson selbst ihre Attraktivität anhand des AR1 einschätzen. Der Fragebogen von Stichprobe 4 enthielt neben dem AR1 zur Bewertung von sechs Modellen nur Fragen nach Alter und Geschlecht der Beurteiler.

2.3 Vorgehen und statistische Analysen

Die Entwicklung und Validierung des AR1 erfolgte in drei Schritten: (1) Vorauswahl durchschnittlich attraktiver Modelle aus der FACES-Datenbank, (2) Selektion von Ankerbildern aufgrund quantitativer Daten aus Stichprobe 1 und (3) empirische

Überprüfung der psychometrischen Güte des AR1 anhand der Stichproben 2, 3 und 4.

Für die Konstruktion des AR1 wurden 171 Bilder von Personen mit neutralem Gesichtsausdruck aus der FACES-Datenbank entnommen und deren Attraktivität ermittelt. Ziel war es, für jede mögliche Zielperson einen im Hinblick auf Alter und Geschlecht möglichst repräsentativen Vertreter mittlerer Attraktivität als Ankerbild auszuwählen. Die 171 Bilder der Modelle wurden daher in sechs Gruppen eingeteilt, in junge, mittelalte und ältere Männer und Frauen (jeweils 27-29 Bilder) und vier Beurteilern aus dem GESIS-Projekt Standardskalen vorgelegt. Diese schätzten die Attraktivität der Modelle anhand einer siebenstufigen Skala von „gar nicht attraktiv“ bis „sehr attraktiv“ ein. Ziel hierbei war eine erste grobe Einschätzung, um Modelle aus dem Bilderpool zu entfernen, die von einer durchschnittlichen Attraktivität abweichen. Die Einschätzungen der Beurteiler wurden aggregiert, um stabilere Schätzungen zu erhalten. Anschließend wurden aus jeder Alter×Geschlecht-Gruppe 10 Bilder selektiert mit ungefähr mittleren Attraktivitätseinschätzungen (nach Standardisierung wurden pro Gruppe diejenigen Modelle mit z-Werten um 0 selektiert) und einem wahrgenommenen Alter im Bereich von 25-30, 45-50, 65-70 Jahren (Alterseinschätzungen berichtet bei Ebner et al. 2010).

Im zweiten Schritt wurden aus den 10 Bildern pro Gruppe weitere vier aufgrund markanter Merkmale nach Expertendiskussion (Projekt Standardskalen) entfernt. Pro Gruppe verblieben sechs Bilder. Die resultierenden 36 Bilder wurden mit Instruktionen und Antwortskala in den Fragebogen von Stichprobe 1 integriert. Für die Auswahl von Ankerbildern durchschnittlich attraktiver Modelle wurde die Stichprobe der Befragungspersonen beziehungsweise Beurteiler nach Alter und Geschlecht in sechs Gruppen unterteilt. Die Attraktivitätseinschätzungen wurden anschließend innerhalb jeder Gruppe gemittelt und aus jeder Gruppe wurde ein Bild mit mittlerer Attraktivitätseinschätzung gewählt. Das heißt, dass zum Beispiel das Ankerbild für die Einschätzung der Attraktivität von älteren Männern anhand der gemittelten Einschätzungen der älteren männlichen Beurteiler aus Stichprobe 1 ausgesucht wurde. Auf diese Weise wurden für alle sechs Alter×Geschlecht-Gruppen Ankerbilder ausgewählt.

Im dritten Schritt wurden zwei unterschiedliche Aspekte der Konstruktvalidität überprüft und die Reliabilität des AR1 geschätzt. Die Konstruktvalidität wurde über Unterschiede der AR1-Mittelwerte zwischen Modellen mit unterschiedlichen Attraktivitätsniveaus und Zusammenhänge des AR1 mit aus der Literatur bekannten Korrelaten von physischer Attraktivität ermittelt. Für jede der sechs Skalenvarianten wurden drei Bilder von Modellen ausgewählt, die von den Beurteilern in Stichprobe 1 als gering (z-Wert ca. -3), mittel (z-Wert ca. 0)

oder hoch attraktiv (z-Wert ca. 3) eingeschätzt wurden. Wenn das AR1 eine valide Einschätzung der Attraktivität erlaubt, dann sollten sich die AR1-Mittelwerte der drei unterschiedlich attraktiven Modelle in einer separaten Stichprobe deutlich voneinander unterscheiden. Die Einschätzung der Modelle aus Stichprobe 1 in „gering“, „mittel“ oder „hoch attraktiv“ sollte sich in den Mittelwerten des AR1 aus Stichprobe 2 widerspiegeln. Wenngleich die Attraktivität in beiden Stichproben mit unterschiedlichen Messkonzepten (absolute vs. relative Einschätzung) erfasst wurde, sollten die Rangreihen der Messwerte dennoch miteinander korrespondieren. Um dies zu prüfen, wurden sechs Varianzanalysen mit Messwiederholung durchgeführt. Messwiederholungsfaktor war die Attraktivität der Modelle (gering, mittel, hoch, klassifiziert nach Stichprobe 1). Ein linearer Trendtest wurde verwendet, um die A-priori-Hypothese steigender AR1-Werte über die drei Attraktivitätsstufen hinweg zu prüfen. Eine weitere Überprüfung der Konstruktvalidität des AR1 fand in Stichprobe 3 statt, indem die Attraktivitätseinschätzungen mit verschiedenen Variablen korreliert wurden, die mit physischer Attraktivität in Zusammenhang stehen. Ziel war hierbei prinzipiell eine Replikation von in der Literatur berichteten Korrelationen. Gleichzeitig sollten aber auch mögliche Unterschiede zwischen Kriteriumskorrelationen von absoluten und relativen Attraktivitätseinschätzungen identifiziert werden. Weiterhin wurde das AR1 in Stichprobe 3 sowohl dem Interviewer als auch der Befragungsperson vorgelegt, um, als weiteren Validitätsbeleg, die Übereinstimmung zwischen Selbst- und Fremdeinschätzung physischer Attraktivität zu ermitteln.

Ein Schätzwert für die Reliabilität der Attraktivitätseinschätzungen mit dem AR1 wurde in Form der Beurteilerübereinstimmung (Interraterreliabilität) für jede der sechs Skalenvarianten in Stichprobe 4 bestimmt. Bei der Bewertung einer Zielperson durch mehrere Beurteiler wird üblicherweise die Konsistenz und/oder die Übereinstimmung der Urteile über die Berechnung statistischer Kennwerte wie z.B. Cronbach α (Cronbach 1951) oder dem Intraklassenkorrelationskoeffizienten (ICC; McGraw/Wong 1996; Shrout/Fleiss 1979) ermittelt, um sicher zu stellen, dass den Einschätzungen ein gemeinsamer Attraktivitätsstandard zugrunde liegt. Je stärker dies der Fall ist, desto ähnlicher fallen die Attraktivitätsurteile auch aus. Im Gegensatz zu Konsistenzmaßen wie Cronbach α berücksichtigen Maße der Übereinstimmung auch systematische Niveauunterschiede zwischen den Beurteilern. Im Vergleich zur hohen Konsistenz von Attraktivitätsurteilen ($r = .85 - .94$, Rubenstein et al. 2002), liegen Kennwerte der Übereinstimmung häufig auf einem niedrigen bis mittleren Niveau (ICCs = $.26 - .34$, Hassebrauck 1993; $.30 - .50$, Thornhill/Gangestad 1999). Die offensichtliche Diskrepanz dieser beiden Ansätze wird bei Hönekopp (2006) und Hassebrauck (1993) kritisch diskutiert. Um die absolute

Übereinstimmung der Beurteiler zu bestimmen, wurde nach den Empfehlungen von McGraw und Wong (1996) der Intraklassenkorrelationskoeffizient (ICC[A, 1]) berechnet. Dazu beurteilten alle Probanden in Stichprobe 4 die Attraktivität von sechs Modellen anhand des AR1. Die Modelle wurden als Portraitfotos oberhalb des AR1 mit der Instruktion präsentiert, die Attraktivität der Modelle anhand des AR1 einzuschätzen.

3 Ergebnisse

3.1 Deskriptive Statistiken

Die mittlere Attraktivitätseinschätzung der Zielpersonen in der bevölkerungsrepräsentativen Stichprobe 3 liegt bei 1.66 (SD = 2.14). Die Interviewer schätzten die Zielpersonen somit im Mittel als mehr als eine Skalenstufe attraktiver ein als die AR1-Ankerfotos. In Tabelle 2 sind deskriptive Statistiken für das AR1 nach Alter, Geschlecht und Bildung aufgeführt, um den Anwendern einen Vergleich der AR1-Werte aus ihrer Untersuchung mit denen relevanter Subgruppen aus einer bevölkerungsrepräsentativen Zufallsstichprobe, zum Beispiel von Männern oder Frauen, von Personen mit unterschiedlicher Schulbildung oder unterschiedlichen Alters, zu ermöglichen.

3.2 Reliabilität

Die Beurteilungen der sechs Modelle in Stichprobe 4 fielen sehr konsistent aus (Cronbach $\alpha = .99$). Die Reliabilität des AR1 wurde anhand der absoluten Beurteilerübereinstimmung geschätzt. Diese wurde über die Berechnung des Intraklassenkorrelationskoeffizienten (ICC[A, 1], McGraw/Wong 1996) bestimmt. Sie fällt mit ICC = .65 in den Bereich angemessener Reliabilität (vgl. Shrout 1998) und liegt somit über den üblicherweise in der Literatur berichteten Kennwerten (siehe Diskussion).

3.3 Validität

Die Ergebnisse der AR1-Einschätzungen in Stichprobe 2 von Modellen mit geringer, mittlerer und hoher Attraktivität (eingeschätzt in Stichprobe 1) und Signifikanztests der Varianzanalysen sind in Tabelle 3 zu finden. Bei allen sechs Skalenvarianten unterschieden sich die Attraktivitätseinschätzungen der drei Modelle signifikant voneinander (alle $F_{S[2,1480]} \geq 205.99$, $ps < .001$, $\eta_p^2 \geq .22$). Modelle, die

Tabelle 2 Deskriptive Statistiken (Stichprobe 3)

Geschlecht	Bildung	AR1-Werte (Fremdeinschätzung)							
		18-35 Jahre		36-65 Jahre		>65 Jahre		Gesamt	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Männlich	gering	1.23	1.80	1.05	2.15	0.83	1.95	0.97	2.01
	mittel	0.75	2.32	1.57	1.99	0.80	1.53	1.21	1.96
	hoch	1.92	1.95	2.51	1.71	1.20	1.79	2.05	1.86
	Gesamt	1.38	2.07	1.67	2.05	0.89	1.82	1.34	2.00
Weiblich	gering	2.11	2.49	1.59	2.47	1.31	1.98	1.49	2.23
	mittel	2.16	2.12	2.14	2.30	1.38	1.89	2.02	2.20
	hoch	2.44	2.07	2.43	2.26	1.77	1.82	2.34	2.13
	Gesamt	2.26	2.14	2.06	2.35	1.39	1.94	1.92	2.21
Gesamt	gering	1.60	2.14	1.33	2.33	1.09	1.98	1.23	2.14
	mittel	1.76	2.26	1.92	2.20	1.08	1.72	1.71	2.15
	hoch	2.23	2.03	2.46	2.01	1.44	1.81	2.21	2.01
	Gesamt	1.92	2.15	1.89	2.23	1.14	1.89	1.66	2.14

N = 1.134

in Stichprobe 1 als „gering attraktiv“ eingeschätzt wurden, zeigten in Stichprobe 2 geringere Mittelwerte im Vergleich zu Modellen, die als „mittel“ oder „hoch attraktiv“ eingeschätzt wurden. Signifikante lineare Trends der Mittelwerte über die drei Attraktivitätsstufen konnten für alle Skalenvarianten beobachtet werden (alle $F_{S[1,740]} \geq 260.76$, $ps < .001$, $\eta_p^2 \geq .31$). Ein quadratischer Trend konnte nur für die Skala „weiblich, alt“ mehr Varianz aufklären als der lineare Trend ($F_{[1,740]} = 360.83$, $p < .001$, $\eta_p^2 = .33$). Für fünf der sechs Skalenvarianten steigen die Mittelwerte des AR1 also mit den Attraktivitätsstufen von „gering“ über „mittel“ bis „hoch attraktiv“ an. Einzige Ausnahme war die Skalenvariante zur Einschätzung von älteren Frauen. Hier korrespondierte die Höhe der Mittelwerte in der AR1 nicht mit den Stufen „mittel“ und „hoch attraktiv“. Die Stufe „hoch attraktiv“ wies einen niedrigeren Mittelwert als die Stufe „mittel“ auf. Da die übrigen Befunde klare Übereinstimmungen der Messwertreihen zeigen, könnte dieser Effekt durch die spezifische Auswahl der Modelle für die Gruppe „weiblich, alt“ erklärt werden. Die praktische Bedeutsamkeit der Effekte der beobachteten Mittelwertsunterschiede und des linearen Anstiegs der Mittelwerte ist als groß einzustufen ($\eta_p^2 > .14$, nach Cohen, 1988). Demnach erlaubt das AR1 eine valide Erfassung unterschiedlicher Grade der Attraktivität von Befragungspersonen.

Tabelle 3 Mittelwerte und Standardabweichungen der AR1-Einschätzungen in Stichprobe 2

Gruppe	Attraktivität			Trend (linear)		Trend (quadratisch)	
	gering	mittel	hoch	$F_{[1,740]}$	η_p^2	$F_{[1,740]}$	η_p^2
Männlich jung	-0.91 (2.00)	0.11 (1.94)	1.07 (2.11)	349.96***	.32	0.24	.00
Weiblich, jung	-0.50 (2.07)	1.03 (2.00)	2.20 (2.21)	590.30***	.44	15.79*	.01
Männlich, mittelalt	-1.15 (1.93)	-0.69 (1.89)	2.01 (2.17)	798.92***	.52	214.79***	.23
Weiblich, mittelalt	-0.96 (1.98)	-0.43 (1.73)	1.12 (2.19)	385.16***	.34	56.27***	.07
Männlich, alt	-1.63 (2.22)	-0.16 (1.96)	0.33 (2.13)	337.18***	.31	45.83***	.06
Weiblich, alt	-0.81 (1.68)	1.46 (1.95)	0.64 (1.81)	260.76***	.26	360.83***	.33

$N = 741$. $F =$ Prüfstatistik des F -Tests. $\eta_p^2 =$ Effektstärke (partielles Eta-Quadrat). * = $p < .05$, *** = $p < .001$.

Um weitere Belege für die Konstruktvalidität des AR1 zu finden, wurde anhand von Stichprobe 3 versucht, aus der Fachliteratur bekannte typische Korrelationen von Attraktivität zu replizieren und einen Vergleich zwischen den anhand des AR1 berechneten Validitätskoeffizienten mit auf relativen Attraktivitätsmaßen beruhenden Korrelationen vorzunehmen. Die Validitätskoeffizienten sind in Tabelle 4 zu finden. Die praktische Bedeutsamkeit der im Folgenden berichteten Validitätskoeffizienten wird nach den Richtlinien von Cohen (1992) vorgenommen: kleiner Effekt ($r = .10$), mittlerer Effekt ($r = .30$), starker Effekt ($r = .50$). Die höchste Korrelation findet sich erwartungsgemäß zwischen der Selbst- und Fremdeinschätzung der Befragungspersonen anhand des AR1 ($r = .46$, $p < .01$). Diese Korrelation liegt im oberen Bereich der üblicherweise in der Literatur berichteten Übereinstimmungen zwischen Fremd- und Selbsteinschätzungen physischer Attraktivität. Feingold (1992) berichtet z.B. in einer metaanalytischen Untersuchung eine mittlere Übereinstimmungen von $r = .24$. Des Weiteren finden sich signifikante, wenn auch nominell kleine Zusammenhänge der AR1-Werte mit soziodemographischen Variablen. So sind als attraktiver bewertete Befragungspersonen jünger, verfügen über ein höheres Einkommen, eine bessere Bildung, ein größeres soziales Netzwerk und sind häufiger in einer festen Beziehung. Eine negative Korrelation physischer Attraktivität mit dem Alter wird häufig berichtet (z.B. Henss 1991; Jones/Hill, 1993; Teuscher/Teuscher 2007). Die mit dem AR1 erhobenen Urteile zeigen dabei einen schwächeren Zusammenhang mit dem Alter als häufig in der Literatur angegeben ($r = -.18$, $p < .01$). Henss (1991) fand z.B. in zwei Experimenten Korrelationen zwischen physischer Attraktivität und Alter im Bereich von $r = -.33$ – $-.83$ bei Frauen

und $r = .00 - -.69$ bei Männern. Der hier gefundene Zusammenhang mit dem Einkommen ($r = .11, p < .05$) deckt sich mit einer Reihe in der Literatur berichteter Korrelationen physischer Attraktivität mit diversen Kriterien des beruflichen und akademischen Erfolgs. Pfeifer (2011) sowie Roszell et al. (2001) fanden dabei ähnliche Effektstärken für den Zusammenhang von Attraktivität und Einkommen ($r = .03 - .08$). Die Korrelation mit dem Bildungsniveau ($r = .18, p < .01$) bestätigt die Ergebnisse von Umberson und Hughes (1987), die ebenfalls einen linearen Zusammenhang dieses Merkmals mit physischer Attraktivität berichten. Vor dem Hintergrund, dass als attraktiver wahrgenommene Personen stärker ausgeprägte soziale Fertigkeiten berichten und in sozialen Situationen weniger befangen sind (Feingold 1992), erscheint es auch plausibel, dass sie über ein größeres soziales Netzwerk verfügen ($r = .11, p < .01$). Weiterhin sind attraktive Menschen eher in einer Partnerschaft als weniger attraktive ($r = .17, p < .01$). Die Höhe dieser Korrelation ist vergleichbar mit den Ergebnissen, die bei Townsend und Levy (1990) für verschiedene Beziehungstypen angegeben werden ($r = .15 - .27$). Frauen wurden generell als attraktiver beurteilt ($r = .13, p < .01$). Übereinstimmend mit Ergebnissen der Attraktivitätsforschung korrelieren die AR1-Werte mit kristalliner Intelligenz ($r = .20, p < .01$). Zebrowitz et al. (2002) berichten analog, abhängig vom Alter der beurteilten Personen, Korrelationen von $r = .11 - .26$ zwischen Attraktivität und Intelligenz. Die Interviewer schrieben attraktiveren Personen außerdem einen höheren sozialen Status zu ($r = .22, p < .01$). Tatsächliche Unterschiede bezüglich sozialer Kompetenz und Dominanz zwischen unterschiedlich attraktiven Menschen werden in diversen Studien berichtet (z.B. Eagly et al. 1991; Feingold 1992; Langlois et al. 2000). Erwartungsgemäße Zusammenhänge finden sich auch bezüglich der Variablen im Bereich Gesundheit und Lebensqualität. Physische Attraktivität korreliert positiv mit der allgemeinen Lebenszufriedenheit ($r = .19, p < .01$) und positiver Affektivität ($r = .29, p < .01$). Vergleichbare Zusammenhänge werden z.B. bei Umberson und Hughes (1987) für positiven Affekt ($\beta = .10$) und Lebenszufriedenheit ($\beta = .06$) berichtet. Negative Zusammenhänge finden sich entsprechend mit selbsteingeschätzten physischen und psychischen Beeinträchtigungen der Gesundheit ($r = -.18$ bzw. $-.12, ps < .01$) sowie negativer Befinden ($r = -.11, p < .01$). Dies deckt sich mit den Ergebnissen von Shackelford und Larsen (1999), die u.a. Zusammenhänge von physischer Attraktivität mit besserer kardiovaskulärer Gesundheit ($r = .15$) und seltenerem Auftreten von Kopfschmerzen ($r = -.17$) berichten. Weiterhin besteht ein positiver Zusammenhang mit der allgemeinen Selbstwirksamkeit ($r = .21, p < .01$), ähnlich den in der Literatur berichteten stärker ausgeprägten internalen Kontrollüberzeugungen attraktiver Menschen (Feingold 1992).

Tabelle 4 Validitätskoeffizienten des AR1 in Stichprobe 3

	AR1 (Fremdeinschätzung)
<i>Attraktivität</i>	
Selbsteinschätzung	.46**
<i>Soziodemographische Variablen</i>	
Alter	-.18**
Geschlecht	.13**
Einkommen ¹	.11*
Bildung	.18**
Größe des sozialen Netzwerks (Anzahl Personen)	.11**
Partnerschaft ja/nein ²	.17**
Länge der Beziehung in Jahren ²	-.14**
<i>Status und Erfolg</i>	
Anerkennung durch Vorgesetzten ³	.08
Zugeschriebener sozialer Status	.22**
<i>Kognitive Fähigkeiten</i>	
Kristalline Intelligenz	.20**
<i>Gesundheit und Lebensqualität</i>	
Positive Affektivität	.29**
Negative Affektivität	-.11**
Beeinträchtigung physisch	-.18**
Beeinträchtigung psychisch	-.12**
Lebenszufriedenheit	.19**
<i>Selbstwirksamkeit</i>	
Allgemeine Selbstwirksamkeit	.21**

N = 1097-1134. Geschlecht = männlich (1), weiblich (2). Bildung = gering (1), mittel (2), hoch (3). Partnerschaft = (1) nein, (2) ja. ¹*N* = 541, ²*N* = 534, ³*N* = 462. * = *p* < .05, ** = *p* < .01.

4 Diskussion

Durch ihre besondere Rolle in Prozessen der sozialen Kognition, z.B. der Eindrucksbildung, und ihrer Implikationen für menschliches Erleben und Verhalten ist die physische Attraktivität für eine Vielzahl sozialer Phänomene relevant. Vermehrt wird diese daher auch in Umfragen erfasst. Die weit verbreitete Messung über absolute Globalurteile auf Ratingskalen weist allerdings methodische Schwächen auf. In der hier beschriebenen Skalenentwicklung wird erstmals ein relatives Messkon-

zept umgesetzt, um diesen methodischen Schwächen Rechnung zu tragen und die Standardisierung der Attraktivitätseinschätzungen zu verbessern.

Das AR1 wurde in mehreren Schritten anhand von vier Stichproben entwickelt und validiert. In einem mehrstufigen Verfahren wurden zunächst Skalenvarianten für die Attraktivitätsbeurteilung von jungen, mittelalten und älteren Männern und Frauen erstellt und anschließend auf ihre psychometrische Güte hin überprüft. Dafür wurden Attraktivitätseinschätzungen mit dem AR1 mit Attraktivitätseinschätzungen, die mit einem üblicherweise eingesetzten absoluten Messverfahren erhoben wurden, verglichen. Weiterhin wurden anhand einer bevölkerungsrepräsentativen Zufallsstichprobe Validitätskoeffizienten ermittelt und aus der Fachliteratur bekannte Korrelate der physischen Attraktivität anhand des AR1 repliziert. In einer weiteren Stichprobe wurde die Reliabilität des AR1 anhand der Beurteilerübereinstimmung bestimmt.

Im Folgenden wird die psychometrische Güte des AR1 diskutiert: Objektivität, Reliabilität und Konstruktvalidität. Die Durchführungsobjektivität hängt bei einem Face-to-face-Interview von dem Interviewer ab, der die Daten erhebt. Sie ist gegeben, wenn dieser sich bei der Messung an die Instruktionen hält. Bei entsprechend geschulten Interviewern ist die Durchführungsobjektivität also üblicherweise gewährleistet (Rammstedt 2010). Auswertungs- und Interpretationsobjektivität sind durch das Messkonzept des AR1 ebenfalls gewährleistet.

Die von verschiedenen Beurteilern mit dem AR1 erhobenen Attraktivitätsurteile der gleichen Personen zeigten ein hohes Maß an Übereinstimmung. Die anhand von Stichprobe 4 berechnete Interraterreliabilität fällt dabei höher aus, als dies in der Literatur für Attraktivitätsbeurteilungen üblicherweise berichtet wird (vgl. Thornhill/Gangestad 1999). Nach den Kriterien von Shrout (1998) zu Anforderungen an Instrumente der klinischen Diagnostik liegen die in dieser Studie ermittelten Werte im Bereich angemessener Interraterreliabilität. Dieses Ergebnis könnte einen ersten Hinweis darauf darstellen, dass den mit dem AR1 erhobenen Attraktivitätsurteilen in stärkerem Maß ein geteilter Bewertungsmaßstab zugrunde liegt, als dies bei herkömmlich verwendeten, absoluten Messinstrumenten der Fall ist.

Im Rahmen der Validitätsanalysen zeigte sich, dass die AR1-Werte von Personen, die in Stichprobe 1 als „gering“, „mittel“ oder „hoch attraktiv“ eingeschätzt wurden, linear über die „Attraktivitätsstufen“ ansteigen, d.h. Personen, die in Stichprobe 1 als „gering attraktiv“ eingeschätzt wurden, zeigten in Stichprobe 2 signifikant geringere Mittelwerte im Vergleich zu Personen, die als „mittel“ oder „hoch attraktiv“ eingeschätzt wurden, und als „hoch attraktiv“ eingeschätzte Personen zeigten im Mittel die höchsten AR1-Werte im Vergleich zu Personen, die als „gering“ oder „mittel“ attraktiv eingeschätzt wurden. Das bedeutet, dass die mit

dem AR1 erhobenen, relativen Attraktivitätswerte grundsätzlich mit standardmäßig eingesetzten absoluten Ratingverfahren vergleichbar sind.

Die empirischen Validitätskoeffizienten des AR1 spiegeln die erwarteten und aus der Literatur bekannten Beziehungen des Konstrukts physische Attraktivität angemessen wider. Die Korrelation zwischen der Fremdeinschätzung der Interviewer und der Selbsteinschätzung der Befragungsperson fiel dabei höher aus, als dies in der Attraktivitätsforschung üblicherweise berichtet wird. Weitere mit der Fachliteratur konsistente Korrelationen von geringer bis mittlerer Stärke wurden mit dem Alter, mit Indikatoren der Gesundheit und der Lebenszufriedenheit, sozioökonomischen Erfolgsvariablen, sozialen Ressourcen und dem zugeschriebenen sozialen Status der Befragungsperson gefunden. Die Ergebnisse der Korrelationsanalysen lassen sich, wie im Folgenden dargestellt, gut in die bestehende Attraktivitätsforschung integrieren.

Wie bei Attraktivitätsbeurteilungen häufig berichtet (z.B. Henss 1991), zeigen die mit dem AR1 erhobenen Urteile einen negativen Zusammenhang mit dem Alter der zu bewertenden Person. Ältere Menschen werden als weniger attraktiv eingeschätzt. Im Vergleich zu in der Literatur berichteten Korrelationen mit absoluten Messverfahren fällt der Einfluss des Alters der Zielperson auf die AR1-Werte allerdings geringer aus. Dieser vergleichsweise geringe Zusammenhang der AR1-Werte mit dem Alter der Zielpersonen lässt den vorläufigen Schluss zu, dass das hier verwendete, relative Messkonzept tatsächlich eine weniger durch Beurteilermerkmale verzerrte Messung physischer Attraktivität erlaubt als die üblicherweise verwendeten absoluten Urteile.

Die Korrelationen von physischer Attraktivität mit dem Einkommen, dem Bildungsniveau und der Zuschreibung von sozialem Status fielen in Richtung und Stärke erwartungsgemäß aus und zeichnen ein Bild von Attraktivität als einem generellen Prädiktor sozioökonomischen Erfolgs. Feingold (1992; siehe auch Kalick 1988) bezeichnet physische Attraktivität als Statuscharakteristikum. Ein entsprechender Effekt auf die Eindrucksbildung findet sich auch im Rahmen der vorliegenden Untersuchung. Die Interviewer ordneten attraktivere Personen signifikant häufiger in eine statushöhere Gesellschaftsschicht ein. Entgegen den Erwartungen fand sich in den hier präsentierten Daten allerdings kein Zusammenhang mit der empfundenen Anerkennung durch Vorgesetzte. Eine positive Diskriminierung aufgrund der Attraktivität (vgl. Morrow et al. 1990) kann dadurch allerdings nicht ausgeschlossen werden, da sie sich evtl. primär in Einstellungsentscheidungen ausdrückt.

Die gefundene Korrelation mit Intelligenz lässt sich gut in die Befunde von Zebrowitz et al. (2002) einordnen, die ähnlich starke Zusammenhänge fanden.

Anzumerken ist hier, dass Zebrowitz et al. auf vollständige IQ-Werte (Stanford-Binet Test und Wechsler Adult Intelligence Scales – Revised) zurückgreifen konnten, während in der vorliegenden Studie aus ökonomischen Gründen nur eine Intelligenzkomponente, die kristalline Intelligenz, betrachtet wurde. Dadurch ist nicht auszuschließen, dass die Varianz der Kriteriumsvariable Intelligenz in dieser Studie eingeschränkt war. Untersuchungen mit einer elaborierteren Intelligenzdiagnostik könnten möglicherweise noch stärkere Korrelationen mit dem AR1 finden. Dass Attraktivität offenbar ein generelles Kriterium der Partnerwahl ist, zeigt sich auch in der hier betrachteten Stichprobe. Attraktive Menschen sind häufiger in einer Partnerschaft als weniger Attraktive. Während attraktivere Menschen häufiger in Beziehungen sind als weniger attraktive, gaben sie allerdings eine geringere Beziehungsdauer an.

Die mit dem AR1 gemessene physische Attraktivität steht weiterhin erwartungsgemäß in Zusammenhang mit mehreren Affekt- und Gesundheitsvariablen. Dabei berichten als attraktiver eingeschätzte Menschen positivere Affektivität, eine höhere Lebenszufriedenheit und weniger psychische und physische Beeinträchtigungen. Diese Befunde decken sich mit Ergebnissen anderer Studien, die nahelegen, dass Attraktivität mit positiver Affektivität und subjektiv wahrgenommener Gesundheit korreliert, aber auch mit physischer Gesundheit (Umberson/Hughes 1987; Feingold 1992; Shackelford/Larsen 1999).

In der metaanalytischen Studie von Feingold (1992) zeigt sich auch, dass attraktivere Menschen stärkere Kontrollüberzeugungen, also den Glauben, dass sie ihr Leben selbst kontrollieren, sowie stärker ausgeprägtes Selbstwertgefühl berichten. In der vorliegenden Studie konnte eine positive Korrelation mit dem Konstrukt allgemeine Selbstwirksamkeit festgestellt werden. Das Konzept Selbstwirksamkeit ist in seiner allgemeinen, situationsunspezifischen Form vergleichbar mit dem Konstrukt der internalen Kontrollüberzeugungen. Attraktive Menschen sind im Vergleich zu weniger attraktiven eher der Überzeugung, dass ihre eigenen Handlungen einen starken Einfluss auf wichtige Aspekte ihres Lebens haben.

Zusammenfassend lässt sich feststellen, dass die AR1-Werte in erwarteter Weise mit verschiedenen aus der Attraktivitätsforschung bekannten Variablen korrelieren. Die Höhe der Validitätskoeffizienten lag in den meisten hier betrachteten Fällen in den in der Literatur berichteten Bereichen. Eine Ausnahme stellt die hohe Korrelation von Fremd- und Selbsturteilen dar. Des Weiteren scheinen die AR1-Werte weniger abhängig vom Geschlecht der Zielpersonen zu sein, als dies an anderer Stelle für Attraktivitätsurteile berichtet wird. Wir werten diese Befunde als erste Hinweise auf die Nützlichkeit des im AR1 verwendeten relativen Messkonzepts im Vergleich zu absoluten Attraktivitätsmaßen.

Die berichteten Befunde dieser Studie legen nahe, dass sich anhand des AR1 das Merkmal physische Attraktivität adäquat messen lässt. Allerdings unterliegt die vorliegende Untersuchung und das Messkonzept des AR1 gewissen Einschränkungen. Durch die Verwendung eines Ankerbilds mit dem Gesicht als Grundlage für die relative Einschätzung beziehen wir die physische Attraktivität einer Befragungsperson ausschließlich auf ihr Gesicht. Dies ist in gewisser Weise eine reduktionistische Sichtweise des Konstrukts, deckt sich aber mit Befunden aus der Attraktivitätsforschung, nach denen die Attraktivität des Gesichts der stärkste Prädiktor einer globalen Einschätzung der physischen Attraktivität ist (Mueser et al. 1984; Peters et al. 2007). Außerdem kann aufgrund der aktuellen Daten nicht abschließend beantwortet werden, ob das relative Messkonzept der AR1 tatsächlich klassischen absoluten Maßen überlegen ist. Dazu müssten systematische Vergleiche der beiden Messkonzepte durchgeführt werden. Zukünftige Studien könnten z.B. die beiden Faktoren Messkonzept und Merkmale der Beurteiler (quasi-)experimentell variieren. Da in der vorliegenden Untersuchung eine vergleichsweise niedrige Korrelation von physischer Attraktivität und Alter gefunden wurde, könnte sich das AR1 auch in einer solchen experimentellen Studie als weniger anfällig für Beurteilermerkmale erweisen. Ein möglicher Einwand gegen die grundsätzliche Relevanz des hier betrachteten Merkmals für die sozialwissenschaftliche Umfrageforschung könnte sich auf die Größe der gefundenen Validierungskoeffizienten beziehen. Rein nominal sind diese mit aufgeklärten Varianzanteilen von maximal acht Prozent als klein zu bewerten (Cohen 1988). Auf einer theoretischen Ebene betrachtet, erscheinen die gefundenen Effekte allerdings realistisch und sind vergleichbar mit Ergebnissen aus metaanalytischen Untersuchungen (Eagly et al. 1991; Feingold 1992; Langlois et al. 2000). Kriterien wie Erfolg im Beruf, sozialer Status und die psychische und physische Gesundheit können nur schwerlich vollständig über einige wenige Variablen erklärt werden. Trotzdem bedeutet dies nicht, dass kleine Einflussgrößen ignoriert werden sollten, da auch sie praktisch relevant sein können (vgl. Abelson 1985; Rosenthal 1990). Frieze et al. (1991) berichten z.B. eine Einkommensdifferenz von \$2.200 pro Jahr zwischen hoch und niedrig attraktiven Männern unter Kontrolle des Bildungsniveaus.

Die Ergebnisse der Überprüfung der psychometrischen Qualität des AR1 sprechen dafür, dass das AR1 physische Attraktivität reliabel, valide und ökonomisch abbildet. Damit ist das AR1 optimal für den Einsatz in Erhebungen mit beschränkten zeitlichen oder monetären Ressourcen, z.B. in sozialwissenschaftlichen oder gesundheitswissenschaftlichen Umfragen, geeignet.

Literatur

- Abelson, R. P., 1985: A variance explanation paradox: When a little is a lot. *Psychological Bulletin* 97: 129–133.
- Andersen, H. H., A. Mühlbacher u. M. Nübling, 2007: Die SOEP-Version des SF 12 als Instrument gesundheitsökonomischer Analysen. SOEP Papers on Multidisciplinary Panel Data Research, 6. Berlin: DIW.
- Asendorpf, J., 2011: *Persönlichkeitspsychologie*. Heidelberg: Springer.
- Bartlett, J. C. und A. Fulton, 1991: Familiarity and recognition of faces in old age. *Memory & Cognition* 19: 229–238.
- Bernstein, N., 1982: Psychological results in burns: The damaged self-esteem. *Clinics in Plastic Surgery* 9: 337–346.
- Bierhoff, H. W., 2000: *Sozialpsychologie*. Ein Lehrbuch. Stuttgart: Kohlhammer.
- Clifford, M. M. und E. Walster, 1973: The effect of physical attractiveness on teacher expectations. *Sociology of Education* 46: 248–258.
- Cohen, J., 1992: A power primer. *Psychological Bulletin* 112: 155–159.
- Cohen, J., 1988: *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum Associates.
- Cronbach, L. J., 1951: Coefficient alpha and the internal structure of tests. *Psychometrika* 16: 297–334.
- Dijkstra, P. und B. P. Buunk, 2001: Sex differences in the jealousy-evoking nature of a rival's body build. *Evolution and Human Behavior* 22: 335–341.
- Eagly, A. H., R. D. Ashmore, M. G. Makhijani und L. C. Longo, 1991: What is beautiful is good, but...: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin* 110: 109–128.
- Ebner, N. C., M. Riediger und U. Lindenberger, 2010: FACES: A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods* 42: 351–362.
- Feingold, A., 1992: Good-looking people are not what we think. *Psychological Bulletin* 111: 304–341.
- Fink, B., K. Grammer und R. Thornhill, 2001: Human (*Homo sapiens*) facial attractiveness in relation to skin texture and color. *Journal of Comparative Psychology* 115: 92–99.
- Fleischer-Peters, A. und J. Margraf-Stiksrud, 1996: Auswirkungen von Normabweichungen auf die Psyche. S. 131–152 in: H.G. Sergl (Hrsg.): *Psychologie und Psychosomatik in der Zahnheilkunde*. München: Urban & Schwarzenberg.
- Frederick, D. A. und M. G. Haselton, 2007: Why is muscularity sexy? Tests of the fitness indicator hypothesis. *Personality and Social Psychology Bulletin* 33: 1167–1183.
- Frieze, I. H., J. E. Olson und J. Russell, 1991: Attractiveness and income for men and women in management. *Journal of Applied Social Psychology* 21: 1039–1057.
- Gallup, G. R. und D. A. Frederick, 2010: The science of sex appeal: An evolutionary perspective. *Review of General Psychology* 14: 240–250.
- Hamermesh, D. S. und J. E. Biddle, 1994: Beauty and the labor market. *American Economic Review* 84: 1174–1194.
- Hassebrauck, M., 1993: Die Beurteilung der physischen Attraktivität. S. 29–59 in: M. Hassebrauck und R. Niketta (Hrsg.): *Physische Attraktivität*. Göttingen: Hogrefe.
- Heimes, H. und C. J. Kemper, 2001: Beeinflusst die Persönlichkeit unsere Reaktion auf physische Attraktivität. Unveröffentlichte Semesterarbeit. Philipps-Universität Marburg.
- Henss, R., 1991: Perceiving age and attractiveness in facial photographs. *Journal of Applied Social Psychology* 21: 933–946.

- Hönekopp, J., 2006: Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology – Human Perception and Performance* 32: 199-209.
- Hosoda, M., E. F. Stone-Romero und G. Coats, 2003: The effects of physical attractiveness on job-related outcomes: A meta-analytical analysis of experimental studies. *Personnel Psychology* 56: 431-462.
- Houston, V. und R. Bull, 1994: Do people avoid sitting next to someone who is facially disfigured? *European Journal of Social Psychology* 24: 279-284.
- Jokela, M., 2009: Physical attractiveness and reproductive success in humans: Evidence from the late 20th century United States. *Evolution and Human Behavior* 30: 342-350.
- Jones, D. und K. Hill, 1993: Criteria of facial attractiveness in five populations. *Human Nature* 4: 271-296.
- Judge, T. A., C. Hurst und L. S. Simon, 2009: Does it pay to be smart, attractive, or confident (or all three)? Relationships among general mental ability, physical attractiveness, core self-evaluations, and income. *Journal of Applied Psychology* 94: 742-755.
- Kalick, S. M., 1988: Physical attractiveness as a status cue. *Journal of Experimental Social Psychology* 24: 469-489.
- Kanazawa, S. und J. L. Kovar, 2004: Why beautiful people are more intelligent. *Intelligence* 32: 227-243.
- Kemper, C. J., J. Lutz, J. Margraf-Stiksrud, C. Beierlein, A. Kovaleva und B. Rammstedt, 2012: Eine Ein-Item-Skala zur Einschätzung von Attraktivität: Das Attraktivitätsrating (AR1). *GESIS Working Papers*2012/24. GESIS – Leibniz Institut für Sozialwissenschaften. Mannheim. http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_arbeitsberichte/WorkingPapers_2012-24.pdf (27.05.2013).
- Kenrick, D. T., und S. E. Gutierrez, 1980: Contrast effects and judgments of physical attractiveness: When beauty becomes a social problem. *Journal of Personality and Social Psychology* 38: 131-140.
- Kenrick, D. T., S. E. Gutierrez und L. L. Goldberg, 1989: Influence of popular erotica on judgments of strangers and mates. *Journal of Experimental Social Psychology* 25: 159-167.
- Langlois, J. H., L. Kalakanis, A. J. Rubenstein, A. Larson, M. Hallam und M. Smoot, 2000: Maxims of myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin* 126: 390-423.
- Lorenz, K., 1943: Die angeborenen Formen möglicher Erfahrung. *Zeitschrift für Tierpsychologie* 5: 235-409.
- Margraf-Stiksrud, J., 1991: Die Funktion des Gesichts für die psychische Entwicklung des Menschen. S. 11-18 in: H.G. Sergl und H. Müller-Fahlbusch (Hrsg.): *Jahrbuch der Psychologie und der Psychosomatik in der Zahlheilkunde*, Bd. 2. Berlin: Quintessenz.
- McGraw, K. O. und S. P. Wong, 1996: Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1: 30-46.
- Menold, N. und C. Kemper, 2013. Verbal and numerical labels? And how many categories? The impact of rating scale formats on quality metrics of multi-item sets. Manuscript in preparation.
- Morrow, P. C., J. C. McElroy, B. G. Stamper und M. A. Wilson, 1990: The effects of physical attractiveness and other demographic characteristics on promotion decisions. *Journal of Management* 16: 723-736.
- Mueser, K. T., B. W. Grau, S. Sussman und A. J. Rosen, 1984: You're only as pretty as you feel: Facial expression as a determinant of physical attractiveness. *Journal of Personality and Social Psychology* 46: 469-478.

- Nedelec, J. L. und K. M. Beaver, 2011: Beauty is in the sex of the beholder: An examination of the effects of interviewer characteristics on assessments of respondent attractiveness. *Personality and Individual Differences* 51: 930-934.
- Peters, M., G. Rhodes und L. W. Simmons, 2007: Contributions of the face and body to overall attractiveness. *Animal Behaviour* 73: 937-942.
- Pfeifer, C., 2011: Physical attractiveness, employment, and earnings. Discussion Paper No. 5664. Leuphana University Lüneburg.
- Rammstedt, B., 2010: Reliabilität, Validität, Objektivität. S. 239258 in: C. Wolf und H. Best (Hrsg.): *Handbuch der sozialwissenschaftlichen Datenanalyse*. Wiesbaden: VS.
- Rhodes, G., 2006: The evolutionary psychology of facial beauty. *Annual Review of Psychology* 57: 199-226.
- Rhodes, M. G. und J. S. Anastasi, 2012: The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin* 138: 146-174.
- Rosenthal, R., 1990: How are we doing in soft psychology? *American Psychologist* 45: 775-777.
- Roszell, P., D. Kennedy und E. Grabb, 2001: Physical attractiveness and income attainment among Canadians. *The Journal of Psychology* 123: 547-559.
- Rubenstein, A. J., J. H. Langlois und L. A. Roggman, 2002: What makes a face attractive and why: The role of averageness in defining facial beauty. S. 1-33 in: G. Rhodes und L.A. Zebrowitz (Hrsg.): *Facial attractiveness: Evolutionary, cognitive, and social perspectives*. Ablex: Westport, CT.
- Schuler, H. und W. Berger, 1979: Physische Attraktivität als Determinante von Beurteilung und Einstellungsempfehlung. *Psychologie und Praxis* 23: 59-70.
- Sergl, H. G., 1991: Psychosoziale Auswirkungen einer Entstellung im Mund- und Gesichtsbereich. S. 19-27 in: H.G. Sergl und H. Müller-Fahlbusch (Hrsg.): *Jahrbuch der Psychologie und der Psychosomatik in der Zahnheilkunde*, Bd. 2. Berlin: Quintessenz.
- Shackelford, T. K. und R. J. Larsen, 1999: Facial attractiveness and physical health. *Evolution and Human Behavior* 20: 71-76.
- Shrout, P. E., 1998: Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research* 7: 301-317.
- Shrout, P. E. und J. L. Fleiss, 1979: Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin* 86: 420-428.
- Sigall, H. und Ostrove, N., 1975: Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology* 31: 410-414.
- Singh, D., B. J. Dixson, T. S. Jessop, B. B. Morgan und A. F. Dixson, 2010: Cross-cultural consensus for waist-hip ratio and women's attractiveness. *Evolution and Human Behavior* 31: 176-181.
- Teuscher, U. und C. Teuscher, 2007: Reconsidering the double standard of aging: Effects of gender and sexual orientation on facial attractiveness ratings. *Personality and Individual Differences* 42: 631-639.
- Thornhill, R. und S. W. Gangestad, 1999: Facial attractiveness. *Trends in Cognitive Sciences* 3: 452-460.
- Townsend, J. M. und G. D. Levy, 1990: Effects of potential partners' physical attractiveness and socioeconomic status on sexuality and partner selection. *Archives of Sexual Behavior* 19: 149-164.
- Tramitz, C., 2000: Die Annäherung – der Erstkontakt zwischen Mann und Frau. S. 33-52 in: P. Kaiser (Hrsg.): *Partnerschaft und Paartherapie*. Göttingen: Hogrefe.

- Umberson, D. und M. Hughes, 1987: The impact of physical attractiveness on achievement and psychological well-being. *Social Psychology Quarterly* 50: 227-236.
- Zebrowitz, L. A., J. A. Hall, N. A. Murphy und G. Rhodes, 2002: Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin* 28: 238-249.

Anschrift des Autors	Johannes Lutz Universität Potsdam Department Psychologie / Sozialpsychologie Karl-Liebknecht-Str. 24-25 14476 Potsdam jlutz@uni-potsdam.de
Ko-Autor/-innen	Christoph J. Kemper Institut für medizinische und pharmazeutische Prüfungsfragen (IMPP), Mainz Constanze Beierlein GESIS – Leibniz-Institut für Sozialwissenschaften Mannheim Jutta Margraf-Stiksrud Philipps-Universität Marburg Beatrice Rammstedt GESIS – Leibniz-Institut für Sozialwissenschaften Mannheim