

# gesis

Leibniz-Institut  
für Sozialwissenschaften

Working Papers

2012|24

## Eine Ein-Item-Skala zur Einschätzung von Attraktivität:

Das Attraktivitätsrating (AR1)

*Christoph J. Kemper, Johannes Lutz,  
Jutta Margraf-Stiksrud, Constanze Beierlein,  
Anastassiya Kovaleva & Beatrice Rammstedt*



GESIS-Working Papers 2012|24

## **Eine Ein-Item-Skala zur Einschätzung von Attraktivität:**

Das Attraktivitätsrating (AR1)

*Christoph J. Kemper, Johannes Lutz,  
Jutta Margraf-Stiksrud, Constanze Beierlein,  
Anastassiya Kovaleva & Beatrice Rammstedt*

## GESIS-Working Papers

GESIS – Leibniz-Institut für Sozialwissenschaften

Postfach 12 21 55

68072 Mannheim

Telefon: (0621) 1246 - 501

Telefax: (0621) 1246 - 500

E-Mail: [christoph.kemper@gesis.org](mailto:christoph.kemper@gesis.org)

[www.christoph-kemper.net](http://www.christoph-kemper.net)

ISSN: 1869-0491 (Online)

Herausgeber,

Druck und Vertrieb:

GESIS – Leibniz-Institut für Sozialwissenschaften  
Unter Sachsenhausen 6-8, 50667 Köln

# 1 Einleitung

---

Erhebungsinstrumente zur Erfassung psychologischer Merkmale wie beispielsweise Persönlichkeit, Risikobereitschaft, Werte, Lebenszufriedenheit, Attraktivität, Optimismus oder Intelligenz, werden immer häufiger in der Forschung eingesetzt. Neben der psychologischen Forschung werden psychologische Merkmale vermehrt auch in anderen Forschungsbereichen erhoben, zum Beispiel in den Sozialwissenschaften, den Wirtschaftswissenschaften, den Ingenieurwissenschaften und den Gesundheitswissenschaften. Die Forscherinnen und Forscher der verschiedenen Disziplinen erwarten von der Erfassung dieser Merkmale in Umfragen, Studien und Experimenten und deren Implementierung in ihre Modelle eine bessere Deskription und Prädiktion wissenschaftlich und gesellschaftlich relevanter Prozesse und Phänomene.

Der Zusammenhang zwischen psychologischen Merkmalen und sozioökonomischen Erfolgsgrößen, die bei Sozialwissenschaftlern und Ökonomen im Fokus des Interesses stehen, gilt als gut gesichert (Gottfredson, 1997; Gottfredson & Deary, 2004; Schmidt & Hunter, 1998; Strenze, 2007). Diese Studien belegen, dass die kognitiven Fähigkeiten einer Person, insbesondere die Intelligenz, der beste Prädiktor für ein erfolgreiches Leben sind. Personen mit hoher kognitiver Leistungsfähigkeit haben im Vergleich zu Personen mit niedriger kognitiver Leistungsfähigkeit mehr Erfolg in Schule, Studium, Ausbildung, Beruf und im Privatleben. Personen mit hoher kognitiver Leistungsfähigkeit haben zum Beispiel meist ein höheres Einkommen oder eine höhere Position im Beruf, lassen sich seltener scheiden und werden seltener delinquent oder arbeitslos. Neben der kognitiven Leistungsfähigkeit sind auch andere psychologische Merkmale dem sozioökonomischen Erfolg zuträglich, wie zum Beispiel Gewissenhaftigkeit und Optimismus (Barrick & Mount, 1991; Kemper, Beierlein, Kovaleva & Rammstedt, 2012).

Neben diversen sozioökonomischen Erfolgsgrößen beeinflussen psychologische Merkmale viele Prozesse und Phänomene, die mitunter weitreichende Implikationen für den Einzelnen, seine Mitmenschen oder die Gesellschaft als Ganzes haben. So kann beispielsweise auf Grundlage psychologischer Merkmale das Wahlverhalten prädiziert werden (Schumann & Schoen, 2005). Auch auf das Gesundheitsverhalten und dessen Folgen haben psychologische Merkmale einen Einfluss. Befunde aus der Fachliteratur zeigen, dass Merkmale wie Gewissenhaftigkeit und Optimismus die physische und die psychische Gesundheit, inklusive der Morbidität und der Mortalität beeinflussen können (Allison, Guichard, Fung & Gilain, 2003; Arthur & Graziano, 1996; Rasmussen, Scheier & Greenhouse, 2009). Schließlich stehen psychologische Merkmale im Zusammenhang mit der Entwicklung und Aufrechterhaltung psychischer Störungen, zum Beispiel Depression und Phobien, und mit Drogen- und Alkoholkonsum sowie delinquentem Verhalten (Block, Block & Keyes, 1988; Block, Gjerde & Block, 1991; Gottfredson, 1997).

Aufgrund dieser vielfältigen Beziehungen und der Nützlichkeit psychologischer Merkmale zur Verbesserung der Deskription und Prädiktion wissenschaftlich und gesellschaftlich relevanter Prozesse und Phänomene forderte kürzlich auch der Ökonomie-Nobelpreisträger James Heckman, dass zukünftige sozialwissenschaftliche Studien vermehrt validierte Persönlichkeitsskalen und Intelligenztests umfassen sollten (Borghans, Duckworth, Heckman & ter Weel, 2008). Dieser Standpunkt wird auch von anderen Forscherinnen und Forschern (Goldberg, 2005; Rammstedt, 2010a) und Institutionen (Rat für Sozial- und Wirtschaftsdaten, 2010) vertreten. Seit einigen Jahren sind diesbezüglich deutliche Bestrebungen erkennbar. Im Sozio-ökonomischen Panel (SOEP) wurden 2004 erstmals Risikoaversion und 2005 Persönlichkeit, Kontrollüberzeugung und Reziprozität erfasst; 2006 auch grundlegende Intelligenzmaße. Darüber hinaus wurden Skalen zur Erfassung psychologischer Merkmale in weitere Untersuchungen aufgenommen, zum Beispiel International Social Survey Programme (ISSP), Household, Income and Labour Dynamics in Australia (HILDA), UK Household Longitudinal Study (UKHLS) und DNB Household Survey (DHS). Der Bedarf an Verfahren zur Operationalisierung psychologischer Merkmale ist demnach gegeben und wird in den kommenden Jahren vermutlich weiter steigen.

Forscherinnen und Forscher, die entsprechende Merkmale in ihrer Untersuchung erfassen möchten, stehen allerdings zurzeit meist vor dem Problem, für ihre Zwecke geeignete Erhebungsinstrumente zu finden. Aus der psychologischen Forschung sind viele Erhebungsinstrumente bekannt (Leibniz-Zentrum für Psychologische Information und Dokumentation, 2011). Diese sind auch in den meisten Fällen sorgfältig im Hinblick auf ihre psychometrische Güte geprüft. Das heißt, durch empirische Untersuchungen wurde festgestellt, dass die Erhebungsinstrumente ein bestimmtes psychologisches Merkmal (Konstrukt) präzise messen (Gütekriterium der Reliabilität oder Messgenauigkeit), und dass sie das Merkmal, das sie zu messen intendieren, auch tatsächlich messen (Gütekriterium der Validität oder Gültigkeit). Allerdings beinhalten diese Erhebungsinstrumente oft viele Fragen oder Aufgaben (Items), um die Konstrukte zu erfassen. Während in der psychologischen Forschung, insbesondere aber in der psychologischen Einzelfalldiagnostik, eine möglichst breite Erfassung der jeweiligen Merkmale notwendig ist und die Dauer der Erhebung (Gütekriterium Ökonomie) eine untergeordnete Rolle spielt, ist die Erhebungsdauer in den meisten sozialwissenschaftlichen Untersuchungen ein kritischer Kostenfaktor. Hier spielt die Anzahl der Items bzw. die Länge von Erhebungsinstrumenten sehr wohl eine Rolle. Da die Erhebungsinstrumente für psychologische Merkmale zwar nicht im Zentrum des Interesses stehen, aber aufgrund ihrer Nützlichkeit dennoch erhoben werden sollen, werden oft auf der Grundlage etablierter psychologischer Skalen und Tests gekürzte Ad-Hoc-Instrumente erstellt. Der Einsatz solcher Ad-Hoc-Instrumente kann in manchen Fällen einen Mehrwert für eine bestimmte Untersuchung darstellen, bringt allerdings den Nachteil mit sich, dass Befunde zwischen Untersuchungen nicht mehr vergleichbar sind, wenn die eingesetzten Erhebungsinstrumente durch unterschiedliche Kürzungsstrategien zustande gekommen sind. Durch eine Kürzung sind genau genommen auch die Belege der psychometrischen Güte, die für das originäre Erhebungsinstrument empirisch ermittelt wurden, nicht mehr gültig. Die psychometrische Güte müsste erneut durch eine Kombination aus konzeptuellen Überlegungen und empirischen Prüfungen (Validierung) belegt werden (Stanton, Sinar, Balzer & Smith, 2002). Der Aufwand für den einzelnen Forscher wäre immens.

Die Bereitstellung von Beratungs- und Serviceleistungen für die Sozialwissenschaften gehört zu den Kernaufgaben von GESIS. Dies beinhaltet auch die Entwicklung und Verbreitung geprüfter, qualitativ hochwertiger und standardisierter Erhebungsinstrumente. Bisher liegen für die Erfassung von für die sozialwissenschaftliche Forschung relevanten psychologischen Merkmalen noch keine etablierten und disziplinübergreifend akzeptierten Erhebungsinstrumente vor, die eine ökonomische und effiziente Messung erlauben. Ziel der hier beschriebenen Skalenentwicklung war es daher, für das Konstrukt Attraktivität ein angemessenes Erhebungsinstrument zu entwickeln, es umfassend zu validieren und anschließend potentiellen Nutzern zur Verfügung zu stellen ([www.gesis.org/kurzskalen-psychologischer-merkmale](http://www.gesis.org/kurzskalen-psychologischer-merkmale)). Damit ist die Hoffnung verknüpft, dass durch den vermehrten Einsatz dieses standardisierten psychologischen Erhebungsinstruments eine erhöhte Anschlussfähigkeit und Vergleichbarkeit zwischen Untersuchungen und eine verbesserte Deskription und Prädiktion wissenschaftlich und gesellschaftlich relevanter Prozesse und Phänomene erzielt werden können.

## 2 Skalenkonzept

---

### 2.1 Theoretischer Hintergrund

Physische Attraktivität ist bereits seit dem 19. Jahrhundert (Galton, 1878) Betrachtungsgegenstand der psychologischen Forschung. Vor einem meist evolutionspsychologischen Hintergrund steht dabei oftmals die Frage im Mittelpunkt, welche physischen Merkmalsausprägungen als attraktiv wahrgenommen werden, inwiefern diese Ausprägungen tatsächliche Indikatoren interindividueller Unterschiede in reproduktiver Fitness der beurteilten Person darstellen und in welchen Beziehungen Attraktivitätseinschätzungen zu menschlichem Denken, Erleben und Verhalten stehen. Die physische Attraktivität einer Person erschließt sich für einen Betrachter aus einer Vielzahl körperlicher Attribute. Neben der Körperkomposition, dem Erscheinungsbild der Haut und der Stimme, sind es vor allem die Gesichtszüge, die einer Attraktivitätseinschätzung zugrunde liegen. Im Allgemeinen beurteilen Menschen symmetrische und durchschnittliche, prototypische Gesichter als attraktiv (Fink & Penton-Voak, 2002).

Die Wahrnehmung von Attraktivität scheint ein biologisch verwurzelter, genereller und kulturübergreifender Prozess zu sein. Frauen und Männer verschiedener Kulturen und sozialer Gruppen zeigen eine hohe Übereinstimmung in Attraktivitätseinschätzungen (Langlois, Kalakanis, Rubenstein, Larson, Hallam, & Smoot, 2000; Thornhill & Gangestad, 1999) und schon bei Neugeborenen zeigt sich eine Präferenz für attraktive Gesichter (Ramsey, Langlois, Hoss, Rubenstein, & Griffin, 2004). Neurowissenschaftliche Forschung identifizierte desweiteren spezifische Hirnstrukturen, die bei der Verarbeitung attraktivitätsbezogener Reize beteiligt sind (z.B. Winston, O'Doherty, Kilner, Perrett, & Dolan, 2007).

Diese Befunde werden im Rahmen evolutionstheoretischer Ansätze damit erklärt, dass Gesichtszüge valide Informationen über die reproduktive Eignung einer Person vermitteln (vgl. Gallup & Frederick, 2010). Tatsächlich finden sich vielfältige Korrelationen von Attraktivitätseinschätzungen mit verschiedensten Domänen menschlichen Erlebens, Denkens und Sozialverhaltens. Neben der offensichtlichen Rolle bei der Partnerwahl (Rhodes, 2006) steht physische Attraktivität in positivem Zusammenhang mit der allgemeinen Gesundheit (Thornhill & Gangestad, 1999), der Anzahl der Nachkommen (Jokela, 2009), Intelligenz (Kanazawa & Kovar, 2004), dem durchschnittlichen Einkommen (Frieze, Olson & Russell, 1991; Judge, Hurst & Simon, 2009) und verschiedenen anderen Kriterien des Berufserfolgs (Hosoda, Stone-Romero & Coats, 2003). Desweiteren beeinflusst die Attraktivitätseinschätzung eine Reihe sozialer Inferenzen wie die Zuschreibung von Kompetenz und Status, prosozialer Orientierung, Integrität und sozialer Anpassung (Eagly, Makhijani, Ashmore, & Longo, 1991; Kalick, 1987).

Physische Attraktivität wird standardmäßig durch die subjektiven Einschätzungen von Beurteilern anhand bipolarer Ratingskalen mit verbalen Ankern z.B. von „gar nicht attraktiv“ bis „sehr attraktiv“ erhoben (Langlois et al., 2000). Den Beurteilern wird dabei ein Foto der zu bewertenden Person vorgelegt und das subjektive Attraktivitätsrating auf der Skala festgehalten. Die Raterübereinstimmung (Konsensus) wird über die Berechnung statistischer Kennwerte wie Kappa (Cohen, 1960; Fleiss, 1971) oder dem Intraklassenkorrelationskoeffizienten (ICC; Shrout & Fleiss, 1979) ermittelt.

Die Nutzung solcher subjektiver Urteile hat sich jedoch als fehleranfällig bzw. sensibel für Charakteristika des Interviewers/Beurteilers erwiesen (Henss, 1991; Hönekopp, 2006; Nedelec & Beaver, 2011). So unterscheiden sich beispielsweise weibliche und männliche Beurteiler in der Attraktivitätseinschätzung von Männern (Nedelec & Beaver, 2011). Während Männer eher zu mittleren Urteilen neigen, beurteilen Frauen Männer im Vergleich öfter als sehr attraktiv oder sehr unattraktiv. Der generell negative Zusammenhang zwischen Alter und Attraktivität bei Attraktivitätseinschätzungen wird ebenfalls durch das Geschlecht und das Alter des Beurteilers beeinflusst (Henss, 1991). Bei Beurteilung von Personen des eigenen Geschlechts oder der eigenen Altersgruppe fällt der Zusammenhang zwischen Alter und Attraktivität signifikant niedriger aus. Kognitionspsychologische Untersuchungen unterstützen die

Rolle des Alters und Geschlechts des Beurteilers bei der Verarbeitung und Erinnerung von Gesichtszügen (Anastasi & Rhodes, 2006; Brigham & Barkowitz, 1978). Als *own-age bias* wird der Effekt bezeichnet, dass Gesichter von Personen der eigenen Altersgruppe besser wiedererkannt werden als Gesichter von älteren oder jüngeren Menschen. Dies wird mit der Häufigkeit erklärt, mit der Menschen mit bestimmten Gesichtern konfrontiert werden (Bartlett & Fulton, 1991). Gesichter, die dem eigenen ähnlich sind, z.B. der gleichen Altersgruppe, werden häufiger betrachtet und durch die resultierende Gewöhnung leichter wiedererkannt.

Diese Befunde verdeutlichen die Beeinflussbarkeit von Attraktivitätseinschätzungen durch individuelle Charakteristika der Beurteiler. Besonders das Geschlecht und das Alter der Rater können zu einer Verzerrung der Ergebnisse auf häufig genutzten Standardratingskalen führen. Im Rahmen der vorliegenden Arbeit wurde eine alternative Ratingskala zur Einschätzung von Attraktivität (Attraktivitätsrating-1, AR1) mit dem Ziel konstruiert, die oben beschriebenen Fehlerquellen zu minimieren. Zu diesem Zweck wird ein nach Alter und Geschlecht der zu beurteilenden Person gematchtes Bild als Anker des Attraktivitätsratings benutzt.

## 2.2 Aufbau

Das AR1 zur Einschätzung der Attraktivität besteht aus einem Item (siehe Abschnitt 2.4 und Appendix B). Als Anker für die Attraktivitätseinschätzung wird zusammen mit der elfstufigen Antwortskala das Bild einer durchschnittlich attraktiven Person des gleichen Geschlechts und der gleichen Altersgruppe wie die zu beurteilende Person dargeboten. Die Einschätzung der Attraktivität erfolgt relativ zu diesem Vergleichsmaßstab, dem Ankerbild. 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. Insgesamt gibt es sechs Skalenvarianten, die sich nur im Ankerbild unterscheiden. Jede enthält ein Ankerbild, das dem Geschlecht (männlich oder weiblich) und dem Alter (junges, mittleres oder hohes Alter) der zu beurteilenden Person (annähernd) entspricht.

## 2.3 Auswertung

Da das AR1 aus nur einem Item besteht, stellt die Antwort des Beurteilers auf das Item gleichzeitig den Skalenwert dar. Der Wertebereich des Skalenwerts liegt zwischen -5 und 5 (für Referenzwerte siehe Abschnitt 5.4 und Appendix A).

## 2.4 Items

In Tabelle 1 sind Itemformulierung und deskriptive Statistiken des AR1 dargestellt. Sie basieren auf Daten einer umfangreichen, bevölkerungsrepräsentativen Zufallsstichprobe (Stichprobe 3, siehe Tabelle 2).

Tabelle 1: Itemformulierung und deskriptive Statistiken des AR1 aus Stichprobe 3.

	<i>M</i>	<i>SD</i>	Sch	Kurt
Bitte schätzen Sie die Attraktivität der Befragungsperson ein.	1.7	2.1	-.28	-.30
Geben Sie an, wie attraktiv oder unattraktiv die Befragungsperson im Vergleich zu der unten abgebildeten Person ist.				

Anmerkungen: *M* = Mittelwert, *SD* = Standardabweichung, Sch = Schiefe, Kurt = Kurtosis. *N* = 1134.

### 3 Durchführung

---

Das AR1 wurde als Forschungsinstrument für sozialwissenschaftliche Untersuchungen unterschiedlichster Art und Fragestellung entwickelt. Das AR1 kann entweder zur Fremdbeurteilung eingesetzt werden, d.h. ein Interviewer schätzt die Attraktivität einer Befragungsperson ein, oder zur Selbstbeurteilung, d.h. die Befragungsperson schätzt ihre eigene Attraktivität ein. Die Itemformulierung wurde mittels kognitiver Pretests optimiert, sodass Personen aus der deutschsprachigen Allgemeinbevölkerung ab 18 Jahren<sup>1</sup> damit umgehen können.

Das AR1 kann prinzipiell in unterschiedlichen Erhebungsmodi eingesetzt werden. Allerdings sind aufgrund zweier Anwendungsvoraussetzungen manche Modi weniger geeignet als andere. Zum Einen erfolgt die Einschätzung der Attraktivität abhängig vom Alter und Geschlecht der Befragungsperson. Diese Daten müssen demnach vor jeder Einschätzung erhoben werden. Zum Anderen sollte die Befragungsperson im Falle einer Fremdbeurteilung durch den Interviewer die Einschätzung ihrer Attraktivität nicht bemerken, da dies negative Effekte auf die Interviewsituation haben könnte. Beide Aspekte legen die Modi fest, in denen das AR1 sinnvoll einsetzbar ist. Alter und Geschlecht vor einer Einschätzung zu erheben und dem Beurteiler (Interviewer oder Befragungsperson) die passende Skalenvariante vorzugeben, ist prinzipiell bei allen Arten von computerunterstützten Erhebungen möglich, z.B. im CAPI-Modus (Computer Assisted Personal Interview), im CAWI-Modus (Computer Assisted Web Interview) und im CASI-Modus (Computer Assisted Self Interview). PAPI-Modus (Paper Assisted Personal Interview) und Selbstausfüller sind hingegen weniger geeignet, da der Interviewer nach der Erfragung von Alter und Geschlecht im Fragebogen blättern muss, um eine Einschätzung vorzunehmen und dies der Befragungsperson vermutlich auffällt. Bei einer Selbstbeurteilung durch die Befragungsperson muss diese selbst eine der sechs Skalenvarianten auswählen. Dieser Prozess scheint fehleranfällig und wenig praktikabel.

Für die Anwendung des AR1 in Erhebungen steht eine Vorlage zur Verfügung. Diese enthält die Instruktion, ein Ankerbild (als Beispiel) und die Antwortskala des AR1. Aufgrund des Copyrights können die sechs Ankerbilder, die als Grundlage der Skalenvarianten dienen, allerdings nicht im Anhang abgedruckt werden. Die Ankerbilder des AR1 stammen aus der FACES-Datenbank des Max-Planck-Instituts für Bildungsforschung in Berlin (<http://faces.mpdl.mpg.de>). Um die Ankerbilder zu Forschungszwecken einsetzen zu dürfen, ist eine Anmeldung beim Max-Planck-Institut für Bildungsforschung notwendig (für eine Anleitung siehe Appendix C). Die Ankerbilder der AR1 tragen die IDs 167, 054, 178, 080, 027 und 133 (aus Set A der Datenbank). Die Zuordnung der Ankerbilder zu den Skalenvarianten ist wie folgt: ID 167 für die Skalenvariante, die zur Einschätzung junger Männer im Alter von 18 - 35 Jahren verwendet wird, ID 054 für junge Frauen im Alter von 18 - 35 Jahren, ID 178 für Männer mittleren Alters (35 - 55 Jahre), ID 080 für Frauen mittleren Alters (35 - 55 Jahre), ID 027 für Männer hohen Alters (56 Jahre und älter), ID 133 für Frauen höheren Alters (56 Jahre und älter). Vor der Anwendung der AR1 in Erhebungen müssen die sechs Skalenvarianten zunächst erstellt werden, indem die oben genannten Bilder in die Vorlage aus Appendix B eingesetzt werden. Das dort aktuell enthaltene Ankerbild ist lediglich ein Beispielbild (ID 116), dessen Veröffentlichung vom Max-Planck-Institut für Bildungsforschung genehmigt wurde. Es gehört nicht zu einer der Skalenvarianten des AR1.

Die Durchführungsdauer des AR1 ist minimal. In 75% der CAPI-Interviews, in dem ein Interviewer eine Befragungsperson anhand des AR1 beurteilt, dauert die Durchführung 20 Sekunden oder weniger ( $P_{75} = 20$ ). Wenn die Skala zur Selbstbeschreibung der Befragungsperson verwendet werden soll, z.B. im

---

<sup>1</sup> Ausgenommen sind Personen, deren sprachliche oder kognitive Fähigkeiten oder deren Wahrnehmungsfähigkeiten, z. B. durch Seh- oder Hörschwäche, unzureichend sind, um die Items zu verstehen.

CASI, dauert die Durchführung in 75% der CASI-Interviews 40 Sekunden oder weniger ( $P_{75} = 40$ ). Ein weiterer wichtiger Aspekt bei der Verwendung des AR1 ist die Interviewerschulung. Die Reliabilität des AR1 hängt maßgeblich von der korrekten Anwendung durch den Interviewer ab (siehe Abschnitt 5.2). Dem Interviewer muss klar sein, dass er die relative und nicht die absolute Attraktivität von Befragungspersonen einschätzen soll.

## 4 Entwicklung und Validierung

---

Ziel der Skalenkonstruktion war es, eine Messung der physischen Attraktivität von Befragungspersonen zu ermöglichen, bei der Verzerrungen durch Beurteilermerkmale wie Alter und Geschlecht weitgehend kontrolliert werden können. Dies sollte über ein Messkonzept ermöglicht werden, bei dem Einschätzungen der Attraktivität einer Person (fortan auch Zielperson) relativ zu einer Person mit durchschnittlicher Attraktivität desselben Geschlechts und derselben Altersgruppe vorgenommen werden. Durch die Verwendung eines Ankerbilds mit Gesicht als Grundlage für die relative Einschätzung beziehen wir die physische Attraktivität einer Zielperson ausschließlich auf ihr Gesicht. Dies ist in gewisser Weise eine reduktionistische Sichtweise des Konstrukts, deckt sich aber mit Befunden aus der Attraktivitätsforschung, nach denen das Gesicht die wichtigste Quelle für eine Einschätzung der physischen Attraktivität ist (Mueser, Grau, Sussman & Rosen, 1984; Peters, Rhodes & Simmons, 2007).

Das AR1 wurde in drei Schritten anhand dreier alters-, geschlechts- und bildungsheterogener Stichproben entwickelt und validiert.

(1) Als Ausgangspunkt für die Skalenkonstruktion wurde die FACES-Datenbank (Ebner, Riediger & Lindenberger, 2010) verwendet. Daraus wurden 171 durch eine Modelagentur vorselegierte Bilder von Personen mit durchschnittlicher Attraktivität (fortan Modelle) ohne emotionalen Ausdruck (Ebner, persönliche Mitteilung, 14.06.2010) entnommen. Ziel dieses Konstruktionsschritts war es, für jede mögliche Zielperson einen im Hinblick auf Alter und Geschlecht möglichst repräsentativen Vertreter mittlerer Attraktivität als Ankerbild auszuwählen, so dass beim zukünftigen Einsatz der Skala eine Verzerrung durch Alter und Geschlecht der Beurteiler minimiert wird. 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 vorgelegt. Diese schätzten die Attraktivität der Modelle anhand einer siebenstufigen Skala von „gar nicht attraktiv“ bis „sehr attraktiv“ ein. Diese Einschätzungen wurden aggregiert. Anschließend wurden aus jeder Gruppe 10 Bilder selektiert mit (a) mittleren Attraktivitätseinschätzungen und (b) einem wahrgenommenen Alter (vgl. Ebner et al., 2010) im Bereich von 25-30, 45-50, 65-70 Jahren.

(2) Im zweiten Schritt wurden aus den 10 Bildern pro Gruppe sechs Bilder ausgewählt. Die resultierenden 36 Bilder wurden mit Instruktionen und Antwortskala von „gar nicht attraktiv“ (1) bis „sehr attraktiv“ (7) in den Fragebogen von Stichprobe 1 integriert (siehe Tabelle 2). Diese Stichprobe 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 Erhebung erfolgte in zwei Wellen mit einem zeitlichen Abstand von 6 bis 10 Wochen. Aus Platzgründen konnten die Bilder nur in Welle 1 vorgegeben werden. Die Daten von Stichprobe 1 wurden im Rahmen eines persönlich-mündlichen Interviews (CAPI) oder durch die Vorgabe eines Papierfragebogens erhoben. Für die Einschätzung der 36 Bilder wurde ein Papierfragebogen vorgegeben. Die Erhebung dauerte insgesamt im Mittel 53 Minuten ( $SD = 12$ ). 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. Als Modelle mittlerer Attraktivität wurden die Bilder mit den IDs 167 als Anker für die Beurteilung von jungen Männern, 054 für die Beurteilung von jungen Frauen, 178 für die Beurteilung mittelalter Männer, 080 für die Beurteilung mittelalter Frauen, 027 für die Beurteilung älterer Männer und 133 für die Beurteilung älterer Frauen aus der FACES-Datenbank ausgewählt.

(3) Um die Validität des AR1 zu prüfen, wurde die so konstruierte Skala schließlich zwei weiteren Stichproben vorgegeben. Stichprobe 2 wurde im Internet erhoben (CAWI, Computer Assisted Web Interview). Bei Stichprobe 2 handelt es sich um eine Quotenstichprobe, geschichtet nach Geschlecht, Alter und Bildung ( $N = 741$ ). Grundgesamtheit waren die Teilnehmer eines Online-Access-Pools im Alter von 18 Jahren oder älter, die in Deutschland leben. Der Onlinefragebogen (Bearbeitungsdauer:  $M = 23$  Minuten,  $SD = 8$ ) enthielt soziodemographische Angaben, verschiedene Skalen zur Erfassung psychologischer Merkmale und das AR1. Um die Validität des AR1 zu prüfen, ließen wir von den Befragungspersonen in Stichprobe 2 Attraktivitätseinschätzungen von Bildern vornehmen, die von den Beurteilern in Stichprobe 1 bereits eingeschätzt wurden. Für jede Alter  $\times$  Geschlecht-Gruppe wurden drei Bilder von Modellen ausgewählt, die in Stichprobe 1 als gering, mittel oder hoch attraktiv eingeschätzt wurden. Die Attraktivität dieser Modelle wurde anhand des AR1 beurteilt. Wenn das AR1 eine valide Erfassung der Attraktivität erlaubt, dann sollten die Rangreihen der Einschätzungen von Stichprobe 1 und 2 annähernd übereinstimmen.

Tabelle 2: Charakteristika der drei Stichproben.

	Stichprobe 1 Welle 1	Stichprobe 1 Welle 2	Stichprobe 2	Stichprobe 3
<i>Stichprobe</i>				
Umfang [ $N$ ]	539	338	741	1134
Art	Quote	Quote	Quote	Zufall
Modus	CAPI, Papier	CAPI, Papier	CAWI	CASI, CASI
<i>Zusammensetzung</i>				
Geschlecht [% Frauen]	52.5%	52.1%	51.8%	55.6%
Alter [ $M(SD)$ ]	47.2 (15.2)	46.7 (15.1)	48.3 (13.0)	53.3 (18.4)
Bildung				
$\leq 9$ Jahre	44.7%	45.3%	40.1%	37.2%
10 Jahre	30.2%	27.9%	29.1%	37.0%
$\geq 11$ Jahre	23.7%	25.4%	30.8%	25.8%

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

Eine weitere Überprüfung der Konstruktvalidität des AR1 fand in Stichprobe 3 statt. Stichprobe 3 mit  $N = 1134$  Befragungspersonen ist eine Zufallsstichprobe, die repräsentativ für die Wohnbevölkerung in Deutschland über einem Alter von 18 Jahren ist. Sie wurde mithilfe des ADM-Stichprobensystem F2F (Random Route) der Arbeitsgemeinschaft deutscher Marktforschungsinstitute gezogen. Die Daten dieses Interview wurden im CAPI- und CASI-Modus erhoben (Bearbeitungsdauer:  $M = 43$ ,  $SD = 13$ ). Die Fragebogenbatterie beinhaltete neben dem AR1 umfangreiche soziodemographische Angaben, weitere psychologische Maße (z.B. BFI-10, Rammstedt & John, 2007, IE-4, Kovaleva, Beierlein, Kemper & Rammstedt, 2012; AKSU, Beierlein, Kovaleva, Kemper, & Rammstedt, 2012) und einige sozialwissenschaftliche Validierungsmaße, z.B. Einkommen, (durch den Interviewer) zugeschriebener sozialer Status der Befragungsperson, Gesundheitszustand etc. Mit diesen Maßen wurden Validitätskoeffizienten für das AR1 berechnet. Außerdem wurde das AR1 zweimal vorgegeben, um die Übereinstimmung zwischen Selbsteinschätzung und Fremdeinschätzung zu ermitteln. 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.

Alle drei Erhebungen wurden von unabhängigen kommerziellen Anbietern durchgeführt. Die Fragebögen der Erhebungen sind auf der Webseite [www.gesis.org/kurzskalen-psychologischer-merkmale](http://www.gesis.org/kurzskalen-psychologischer-merkmale) zu finden.

## 5 Gütekriterien

---

### 5.1 Objektivität

Unter Objektivität wird der Grad verstanden, in dem eine Messung unabhängig vom Untersucher ist (vgl. Lienert & Raatz, 1998). Diese bezieht sich auf verschiedene Phasen einer Untersuchung: Durchführung, Auswertung und Interpretation. Im Falle eines Face-to-face-Interviews hängt die Durchführungsobjektivität von dem Interviewer ab, der die Daten erhebt. Sie ist gegeben, wenn dieser sich bei der Vorgabe der Skala an die genauen Instruktionen und den Wortlaut der Items hält. Bei entsprechend geschulten Interviewern ist die Durchführungsobjektivität üblicherweise gewährleistet (Rammstedt, 2010b). Auswertungsobjektivität betrifft die numerische und kategoriale Auswertung des Antwortverhaltens der Befragten nach festgelegten Regeln (vgl. Lienert & Raatz, 1998). Diese ist beim AR1 vollständig gegeben, da jeder der elf Antwortkategorien eindeutig ein Zahlenwert zugewiesen wird (siehe Abschnitt 2.3). Interpretationsobjektivität ist gegeben, wenn die aus den Befragungsergebnissen gezogenen Schlüsse über verschiedene Forscher vergleichbar sind. Zur Maximierung der Interpretationsobjektivität sollte das Wissen der Forscher über die Messintention der Skala und über die Interpretation der quantitativen Messwerte vergleichbar sein (Rammstedt, 2010b). Durch die Standardisierung der Auswertung und die Zuweisung eines numerischen Messwerts, der die Attraktivität des Befragten beschreibt (siehe Befunde zur Validität in Abschnitt 5.3), kann auch die Interpretationsobjektivität des AR1 als gegeben angesehen werden.

### 5.2 Reliabilität

Unter der Reliabilität oder Messgenauigkeit einer Skala versteht man den Grad der Genauigkeit, mit dem ein bestimmtes Merkmal erfasst wird (vgl. Lienert & Raatz, 1998). Die Reliabilität des AR1 wurde über die Testwiederholungsmethode (Retest) bestimmt. Dazu wurde eine nicht-studentische Gelegenheitsstichprobe erhoben (N = 28; Alter: M = 36.1, SD = 14.9; 60.7% Männer). Anhand des AR1 wurde die Attraktivität von sechs Modellen eingeschätzt, deren Bilder gemeinsam mit dem AR1 dargeboten wurden. In einem zeitlichen Abstand von ca. einer Woche (M = 7.5 Tage, SD = 1.5) wurden den Teilnehmern dieselben sechs Modelle erneut zur Beurteilung vorgelegt. Um die Retestreliaibilität zu ermitteln, wurden die Messwertreihen der beiden Testungen für jedes Modell miteinander korreliert. Die ermittelte Retest-Koeffizienten liegen in dieser Stichprobe zwischen  $r_{tt} = .46$  und  $.85$ . Für die Beurteilung der meisten Modelle in der Gelegenheitsstichprobe ergab sich demnach eine ausreichende bis gute Reliabilität für Gruppenuntersuchungen (vgl. Aiken & Groth-Marnat, 2006). Vereinzelt traten jedoch auch Reliabilitätskoeffizienten unterhalb der von Aiken und Groth-Marnat angegebenen Grenze von  $.60$  auf. Diese geringen Koeffizienten ergaben sich aufgrund eines mangelnden Instruktionsverständnisses einzelner Beurteiler. Anstatt relative Attraktivitätsurteile abzugeben, gaben sie vereinzelt absolute Urteile ab.

### 5.3 Validität

#### Inhaltliche Validität

Inhaltliche Validität ist gegeben, wenn ein Item das zu messende Konstrukt wirklich bzw. hinreichend präzise abbildet (Bühner, 2011). Im AR1 wird ein Beurteiler gebeten, die Attraktivität einer Zielperson einzuschätzen. Anhand des Ratings erfolgt also eine direkte Zuweisung eines numerischen Messwertes für das Konstrukt Attraktivität. Die inhaltliche Validität des AR1 ist daher gegeben.

## Konstruktvalidität

Konstruktvalidität ist gegeben, wenn sich der Messwert einer Skala als Indikator für das Merkmal eignet, das mit der Skala gemessen werden soll. Die Eignung wird daran beurteilt, wie gut sich die Skala im Hinblick auf Hypothesen, die sich aus dem Modell des zu erfassenden Merkmals ableiten lassen, empirisch bewährt. Bei der Konstruktvalidierung werden anhand theoretischer Überlegungen positive, negative und Nullbeziehungen einer zu validierenden Skala mit anderen empirischen Indikatoren postuliert (nomologisches Netzwerk) und getestet, ob sich das vorhergesagte Muster empirisch belegen lässt (Krohne & Hock, 2007).

Um die Konstruktvalidität des AR1 zu belegen, wurden für jede der sechs Skalenvarianten drei Bilder von Modellen ausgewählt, die von den Beurteilern in Stichprobe 1 als gering, mittel oder hoch attraktiv eingeschätzt wurden. Wenn das AR1 eine valide Einschätzung der Attraktivität erlaubt, dann sollten sich die Mittelwerte der drei unterschiedlich attraktiven Modelle 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). Die Ergebnisse sind in Tabelle 3 zu finden.

*Tabelle 3:* Mittelwerte und Standardabweichungen der AR1-Einschätzungen in Stichprobe 3 von Modellen mit geringer, mittlerer und hoher Attraktivität (eingeschätzt in Stichprobe 1) und Signifikanztests der Varianzanalyse.

Gruppe	Attraktivität			Effekt		Trend (linear)	
	gering	mittel	hoch	$F_{[2,1474]}$	$\eta_p^2$	$F_{[1,737]}$	$\eta_p^2$
Männlich, jung	-0.91 (2.00)	0.11 (1.94)	1.07 (2.11)	205.99***	.22	345.22***	.32
Weiblich, jung	-0.50 (2.07)	1.03 (2.00)	2.20 (2.21)	354.48***	.32	590.30***	.44
Männlich, mittelalt	-1.15 (1.93)	-0.69 (1.89)	2.01 (2.17)	573.49***	.44	798.92***	.52
Weiblich, mittelalt	-0.96 (1.98)	-0.43 (1.73)	1.12 (2.19)	266.68***	.27	385.16***	.34
Männlich, alt	-1.63 (2.22)	-0.16 (1.96)	0.33 (2.13)	225.03***	.23	337.18***	.31
Weiblich, alt	-0.81 (1.68)	1.46 (1.95)	0.64 (1.81)	313.27***	.30	260.76***	.26

*Anmerkung:*  $N = 741$ .  $F$  = Prüfstatistik des F-Tests.  $\eta_p^2$  = Effektstärke (partielles Eta-Quadrat).

Bei allen sechs Skalenvarianten unterschieden sich die Attraktivitätseinschätzungen der drei Modelle deutlich voneinander. Modelle, die in Stichprobe 1 als „gering attraktiv“ eingeschätzt wurden, zeigten in Stichprobe 2 deutlich geringere Mittelwerte im Vergleich zu Modellen, die als „mittel“ oder „hoch attraktiv“ eingeschätzt wurden. Ein linearer Trend der Mittelwerte über die drei „Attraktivitätsstufen“ konnte beobachtet werden. Für fünf der sechs Skalenvarianten steigen die Mittelwerte des AR1 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 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.

Um weitere Belege für die Konstruktvalidität des AR1 zu finden, wurde anhand von Stichprobe 3 versucht, aus der Fachliteratur bekannte typische Korrelate der Attraktivität zu replizieren. 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 des AR1 (Fremdeinschätzung der Attraktivität der Zielperson) findet sich erwartungsgemäß mit der Selbsteinschätzung der Attraktivität der Zielperson ( $r = .46, p < .01$ ). Desweiteren finden sich signifikante, wenn auch kleine Zusammenhänge der AR1-Werte mit soziodemographischen Variablen. So sind als attraktiver bewertete Befragungspersonen jünger, verfügen über eine bessere Bildung, ein größeres soziales Netzwerk und sind häufiger in einer festen Beziehung. Frauen werden generell als attraktiver beurteilt ( $r = .13, p < .01$ ).

Tabelle 4: Validitätskoeffizienten des AR1 (Fremdeinschätzung) für Stichprobe 3.

		Stichprobe 3
Attraktivität	<i>Selbsteinschätzung</i>	.46**
Soziodemographische Variablen	<i>Alter</i>	-.18**
	<i>Geschlecht</i>	.13**
	<i>Einkommen</i>	.12**
	<i>Bildung</i>	.18**
	<i>Größe soziales Netzwerk</i>	.11**
	<i>Partnerschaft ja/nein<sup>2</sup></i>	.17**
	<i>Zugeschriebener sozialer Status</i>	.22**
Status		
Intelligenz	<i>Kristalline Intelligenz</i>	.20**
Gesundheitsstatus	<i>Beeinträchtigung physisch</i>	-.18**
	<i>Beeinträchtigung psychisch</i>	-.12**
Selbstkonzept/-regulation	<i>Selbstwirksamkeit</i>	.21**
	<i>Internale Kontrollüberzeugung</i>	.20**
	<i>Externale Kontrollüberzeugung</i>	-.15**
Persönlichkeit	<i>Neurotizismus</i>	-.09**
	<i>Extraversion</i>	.21**
	<i>Offenheit</i>	.19**
	<i>Verträglichkeit</i>	.10**
	<i>Gewissenhaftigkeit</i>	.20**

Anmerkung:  $N = 1100 - 1134$ . Geschlecht = männlich (1), weiblich (2). Bildung = gering (1), mittel (2), hoch (3) (siehe Abschnitt 5.4). Partnerschaft = (1) nein, (2) ja <sup>1</sup>  $N = 647$ , <sup>2</sup>  $N = 53$ . \* =  $p < .05$ , \*\* =  $p < .01$ .

Im Rahmen des Fünf-Faktoren-Modell der Persönlichkeit finden sich geringe bis mittlere Zusammenhänge der AR1-Werte mit Extraversion, Offenheit und Gewissenhaftigkeit. Zusätzlich zeigt sich ein, wenn auch schwacher, negativer Zusammenhang mit Neurotizismus. Auch mit Aspekten des Selbstkonzepts und der Selbstregulation bestehen differentielle Korrelationen. Befragungspersonen mit höheren AR1-Werten weisen eine höhere Selbstwirksamkeit und internale Kontrollüberzeugungen auf.

Dagegen findet sich ein negativer Zusammenhang mit externalen Kontrollüberzeugungen. Übereinstimmend mit Ergebnissen der Attraktivitätsforschung korrelieren die AR1-Werte mit kristalliner Intelligenz ( $r = .20, p < .01$ ), physischen und psychischen Beeinträchtigungen der Gesundheit ( $r = -.18$  bzw.  $-.12, ps < .01$ ) und dem durch den Interviewer zugeschriebenen sozialen Status der Befragungsperson ( $r = .22, p < .01$ ).

#### 5.4 Referenzwerte

Im Appendix A sind Referenzwerte in Form von Gruppenmittelwerten und Standardabweichungen für das AR1 abgedruckt (Fremdeinschätzung der Attraktivität von Befragungspersonen durch den Interviewer). Diese wurden anhand der Zufallsstichprobe (Stichprobe 3) ermittelt und erlauben dem Anwender einen Vergleich der AR1-Werte aus seiner Untersuchung mit denen relevanter Subgruppen aus einer bevölkerungsrepräsentativen Zufallsstichprobe, zum Beispiel von Männern oder Frauen, von Personen mit unterschiedlicher Schulbildung oder unterschiedlichen Alters. Die Altersgruppen in der Tabelle von Appendix A wurden den Lebensphasen der bundesdeutschen Gesellschaft angepasst. Die Zeit von 18 bis 35 Jahren ist die der beruflichen Ausbildung und Familiengründung. Die Zeit der beruflichen Festigung, Karriere, Betreuung heranwachsender Kinder und Pflege älterer Angehöriger fällt in die Zeit zwischen 36 und 65 Jahren. Die dritte Lebensphase beginnt im Alter von 65 Jahren, wenn die berufliche Tätigkeit in den meisten Fällen abgeschlossen ist. Die Aufteilung der Bildungsstufen wurde nach der Dauer der schulischen Allgemeinbildung vorgenommen. Dabei gilt die Dauer der schulischen Bildung bis einschließlich 9 Jahren als geringes Bildungsniveau. Bei einer Schuldauer von 10 oder 11 Jahren handelt es sich um ein mittleres Bildungsniveau und bei mehr als 11 Jahren um ein hohes Bildungsniveau.

## 6 Kurzfassung

---

### Diagnostische Zielsetzung

Ziel der Skalenkonstruktion war es, eine ökonomische, reliable und valide Messung der physischen Attraktivität von Befragungspersonen zu ermöglichen, bei der die üblichen Verzerrungen durch Merkmale der Beurteiler wie Alter und Geschlecht weitgehend kontrolliert werden können. Dies wird über ein Messkonzept ermöglicht, bei dem Einschätzungen der Attraktivität einer Befragungsperson relativ zu einer Person mit durchschnittlicher Attraktivität desselben Geschlechts und derselben Altersgruppe vorgenommen werden.

### Aufbau

Das AR1 zur Einschätzung der Attraktivität besteht aus einem Item. Als Anker für die Attraktivitätseinschätzung wird zusammen mit einer elfstufigen Antwortskala das Bild einer durchschnittlich attraktiven Person des gleichen Geschlechts und der gleichen Altersgruppe (Modell) wie die zu beurteilende Person dargeboten. Die Einschätzung der Attraktivität erfolgt relativ zu diesem Vergleichsmaßstab, dem Ankerbild. 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 des durchschnittlich attraktiven Modells dargeboten wird. Insgesamt gibt es sechs Skalenvarianten, die sich nur im Ankerbild unterscheiden, das dem Geschlecht (männlich oder weiblich) und dem Alter (junges, mittleres oder hohes Alter) der zu beurteilenden Person entsprechen soll.

### Grundlagen und Konstruktion

Ziel der Skalenkonstruktion war die Messung der physischen Attraktivität von Befragungspersonen per Fremdeinschätzung. Durch die Verwendung eines Ankerbilds mit 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 das Gesicht die wichtigste Quelle für eine Einschätzung der physischen Attraktivität ist (Mueser, Grau, Sussman & Rosen, 1984; Peters, Rhodes & Simmons, 2007). Das AR1 wurde in mehreren Schritten anhand dreier alters-, geschlechts- und bildungsheterogener Stichproben entwickelt und validiert. Zunächst wurden die Ankerbilder ausgewählt, relativ zu denen die Einschätzung der Attraktivität erfolgen soll. Es wurde jeweils ein Ankerbild mit einem Modell durchschnittlicher Attraktivität für junge, mittelalte und ältere Männer und Frauen ausgewählt. Anschließend wurden die Reliabilität und verschiedene Aspekte der Konstruktvalidität überprüft.

### Gütekriterien

*Reliabilität* Die Reliabilität des AR1 wurde in einer nicht-studentischen Gelegenheitsstichprobe mittels Retest-Methode anhand von sechs zu beurteilenden Modellen bestimmt (Intervall:  $M = 7.5$  Tage,  $SD = 1.5$ ). Die ermittelte Retestreliabilität liegt in dieser Stichprobe zwischen  $r_{tt} = .46$  und  $.85$ . Für die Beurteilung der meisten Modelle in der Gelegenheitsstichprobe ergab sich demnach eine ausreichende bis gute Reliabilität für Gruppenuntersuchungen (vgl. Aiken & Groth-Marnat, 2006). Vereinzelt traten jedoch auch Reliabilitätskoeffizienten unterhalb der von Aiken und Groth-Marnat angegebenen Grenze von  $.60$  auf. Die geringen Koeffizienten ergaben sich aufgrund eines mangelnden Instruktionsverständnisses einzelner Beurteiler. Bei geschulten Beurteilern bzw. Interviewern kann von einer ausreichenden Reliabilität des AR1 für Forschungskontexte (Gruppenuntersuchungen) ausgegangen werden.

*Validität* Anhand dreier alters-, geschlechts- und bildungsheterogener Stichproben wurde das AR1 validiert. Dafür wurden Attraktivitätseinschätzungen der AR1 aus Stichprobe 2 mit Attraktivitätsein-

schätzungen aus Stichprobe 1 verglichen. Dabei zeigte sich, dass die Attraktivitätseinschätzungen aus den beiden Stichproben konvergierten. Die AR1-Werte von Personen, die in Stichprobe 1 als „gering“, „mittel“ oder „hoch attraktiv“ eingeschätzt wurden, stiegen linear über die „Attraktivitätsstufen“ an. Außerdem wurden anhand einer bevölkerungsrepräsentativen Zufallsstichprobe Validitätskoeffizienten ermittelt und versucht aus der Fachliteratur bekannte Korrelate der physischen Attraktivität anhand des AR1 zu replizieren. Die AR1-Werte korrelierten erwartungsgemäß mit der Selbsteinschätzung der Befragungsperson und dem Alter, mit Beeinträchtigungen der physischen und psychischen Gesundheit, dem sozioökonomischen Erfolg, sozialen Ressourcen und dem zugeschriebenen sozialen Status der Befragungsperson. Die Befunde sprechen für die Konstruktvalidität des AR1.

## 7 Bewertung

---

Das AR1 ist eine hochgradig ökonomische Skala zur Erfassung des Merkmals physische Attraktivität in sozialwissenschaftlichen Untersuchungen, die starken zeitlichen und monetären Restriktionen unterliegen und eine umfangreichere Erfassung nicht zulassen. Weiterhin erlaubt das AR1 eine Messung des Konstrukts, bei der die üblichen konfundierenden Einflüsse durch Merkmale der Beurteiler wie deren Alter und Geschlecht weitgehend kontrolliert werden. Die Skala kann in verschiedenen computerunterstützten Erhebungsmodi administriert werden, z.B. im CAPI-Modus (Computer Assisted Personal Interview), im CAWI-Modus (Computer Assisted Web Interview) und im CASI-Modus (Computer Assisted Self Interview). Im PAPI-Modus (Paper Assisted Personal Interview) und als Selbstausfüller ist das AR1 nicht sinnvoll einsetzbar. Validitätsbelege wurden für den CAPI- und den CAWI-Modus erbracht. Die Invarianz zwischen den Erhebungsmodi wurde bisher allerdings noch nicht geprüft. Ob Korrelationen und Mittelwerte des AR1 daher über Stichproben hinweg vergleichbar sind, die in unterschiedlichen Modi erhoben wurden, kann zurzeit nicht vorausgesetzt werden. Vor dem Einsatz des AR1 in Mixed-Mode-Designs sollte dieser Beleg erbracht werden (für aktualisierte Informationen bezüglich der psychometrischen Güte siehe [www.gesis.org/kurzskalen-psychologischer-merkmale](http://www.gesis.org/kurzskalen-psychologischer-merkmale)). Die Erhebung von Daten mit dem AR1 dauert deutlich weniger als eine Minute.

Die empirischen Belege der Validierungsstudien sprechen für die psychometrische Güte des AR1. Im Hinblick auf die Reliabilität sei allerdings einschränkend erwähnt, dass diese nicht durchgehend als ausreichend belegt werden konnte. Bei der Beurteilung weniger Modelle traten auch Reliabilitätskoeffizienten unter der für Gruppenuntersuchungen angegebenen Grenze von .60 auf. Neben dem geringen Stichprobenumfang kann hierfür insbesondere die mangelnde Übung der Beurteiler im Umgang mit sozialwissenschaftlichen Fragebögen im Allgemeinen und mit dem AR1 im Besonderen angeführt. Manche Beurteiler gaben trotz anderslautender Instruktion absolute statt relative Attraktivitätsurteile ab. Dies unterstreicht die besondere Bedeutung von Schulungen der Beurteiler bzw. der Interviewer in der Anwendung des AR1. Mit entsprechender Schulung sollte mit dem AR1 eine für Gruppenuntersuchungen mindestens ausreichende Reliabilität erzielt werden können.

Auch die Konstruktvalidität des AR1 konnte belegt werden. Zunächst wurden Attraktivitätseinschätzungen auf Grundlage des AR1 aus Stichprobe 2 mit Attraktivitätseinschätzungen aus Stichprobe 1 verglichen. Die Bilder der beurteilten Modelle wurden in Stichprobe 1 als „gering“, „mittel“ oder „hoch attraktiv“ eingeschätzt. Bei dem Vergleich zeigte sich, dass die AR1-Werte von Modellen, die in Stichprobe 1 als „gering“, „mittel“ oder „hoch attraktiv“ eingeschätzt wurden, linear über die „Attraktivitätsstufen“ ansteigen, d.h. Modelle, die in Stichprobe 1 als „gering attraktiv“ eingeschätzt wurden, zeigten in Stichprobe 2 deutlich geringere Mittelwerte im Vergleich zu Modellen, die als „mittel“ oder „hoch attraktiv“ eingeschätzt wurden. Als „hoch attraktiv“ eingeschätzte Personen zeigten im Mittel die höchsten AR1-Werte im Vergleich zu Personen, die als „gering“ oder „mittel“ attraktiv eingeschätzt wurden. Mit Ausnahme einer Skalenvariante konnte dieser lineare Anstieg der AR1-Werte bei allen der übrigen fünf Skalenvarianten und über alle Attraktivitätsstufen gezeigt werden. Der Effekt ist daher vermutlich der spezifischen Auswahl des Modells mit hoher Attraktivität (für Bedingung „weiblich/alt“) in der Validierungsstudie geschuldet.

Weitere Belege für die Konstruktvalidität des AR1 wurden korrelativ ermittelt. Die empirischen Validitätskoeffizienten spiegeln die aus der psychologischen Literatur bekannten Beziehungen des Konstrukts „Physische Attraktivität“ angemessen wider: Die höchste Korrelation des AR1, also der fremd eingeschätzten Attraktivität der Befragungsperson, zeigte sich mit der Selbsteinschätzung der Befragungsperson. Weitere mit der Fachliteratur konsistente Korrelate von geringer Stärke ( $.10 < r < .30$ ) wurden mit dem Alter (negativer Zusammenhang), mit Beeinträchtigungen der physischen und psychischen Gesundheit (negativer Zusammenhang), sozioökonomischen Erfolgsvariablen wie Einkommen

und Bildung (positiver Zusammenhang), sozialen Ressourcen wie Partnerschaft (positiver Zusammenhang) und dem zugeschriebenen sozialen Status der Befragungsperson (positiver Zusammenhang) gefunden.

Die empirischen Belege der Validierungsstudien sprechen dafür, dass das AR1 nicht nur eine ökonomische, sondern auch eine für Forschungskontexte ausreichend reliable und eine valide Erfassung des Merkmals physische Attraktivität erlaubt. In mehreren umfangreichen alters-, geschlechts- und bildungsheterogenen Stichproben, darunter auch eine repräsentative Zufallsstichprobe, konnten die Reliabilität des AR1 und unterschiedliche Aspekte der Konstruktvalidität belegt werden. Die AR1 ermöglicht somit die Fremdeinschätzung der physischen Attraktivität von Befragungspersonen durch Interviewer, ohne die bei einer Fremdeinschätzung üblicherweise auftretenden konfundierenden Einflüsse durch Merkmale der Beurteiler, wie deren Alter und Geschlecht. Ein direkter empirischer Vergleich des hier vorgeschlagenen Messkonzeptes einer relativen Attraktivitätsmessung mit der traditionellen „absoluten“ Messung steht allerdings noch aus. Eine entsprechende Studie ist aktuell in Planung (für Befunde siehe demnächst oben genannte Internetseite).

## 8 Literatur

---

- Allison, P. J., Guichard, C., Fung, K. & Gilain, L. (2003). Dispositional optimism predicts survival status 1 year after diagnosis in head and neck cancer patients. *Journal of Clinical Oncology*, *21*, 543-548.
- Anastasi, J. S. & Rhodes, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology*, *8*, 237-253.
- Arthur, W., Jr. & Graziano, W. G. (1996). The five-factor model, conscientiousness, and driving accident involvement. *Journal of Personality*, *64*, 593-618.
- Barrick, M. R. & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, *44*, 1-26.
- Bartlett, J. C. & Fulton, A. (1991). Familiarity and recognition of faces in old age. *Memory & Cognition*, *19*, 229-238.
- Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2012). Ein Messinstrument zur Erfassung subjektiver Kompetenzerwartungen: Allgemeine Selbstwirksamkeit Kurzskala (ASKU) (GESIS Working Papers 2012|17). Köln: GESIS.
- Block, J., Block, J. H. & Keyes, S. (1988). Longitudinally foretelling drug usage in adolescence: Early childhood personality and environmental precursors. *Child Development*, *59*, 336-355.
- Block, J., Gjerde, P. F. & Block, J. H. (1991). Personality antecedents of depressive tendencies in 18-year-olds: A prospective study. *Journal of Personality and Social Psychology*, *60*, 726-738.
- Borghans, L., Duckworth, A. L., Heckman, J. J. & ter Weel, B. (2008). *The Economics and Psychology of Personality Traits*. IZA Discussion Paper No. 3333. Bonn: IZA.
- Brigham, J. C. & Barkowitz, P. (1978). Do 'They all look alike'? The effect of race, sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Social Psychology*, *8*, 306-318.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. PS Psychologie. München: Pearson Studium.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.
- Eagly, A. H., Ashmore, R. D., Makhijani, M. G. & Longo, L. C. (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., Riediger, M., & Lindenberger, U. (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.
- Fink, B. & Penton-Voak, I. S. (2002) Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, *11*, 154-158.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*, 378-382.
- Frieze, I. H., Olson, J. E. & Russell, J. (1991). Attractiveness and income for men and women in management. *Journal of Applied Social Psychology*, *21*, 1039-1057.
- Gallup, G. R. & Frederick, D. A. (2010). The science of sex appeal: An evolutionary perspective. *Review of General Psychology*, *14*, 240-250.

- Galton, F. (1878). Composite portraits. *Journal of the Anthropological Institute of Great Britain and Ireland*, 8, 132-142.
- Goldberg, L. R. (2005). *Why Personality Measures Should Be Included in Epidemiological Surveys: A brief Commentary and a Reading List*. Oregon Research Institute.
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24, 79-132.
- Gottfredson, L. S. & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, 13, 1-4.
- 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., Stone-Romero, E. F. & Coats, G. (2003). The effects of physical attractiveness on job-related outcomes: A meta-analysis of experimental studies. *Personnel Psychology*, 56, 431-462.
- International Test Commission (2010). International Test Commission Guidelines for Translating and Adapting Tests. Gefunden am 27.06.2012 unter <http://www.intestcom.org>.
- 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.
- Judge, T. A., Hurst, C. & Simon, L. S. (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. & Kovar, J. L. (2004). Why beautiful people are more intelligent. *Intelligence*, 32, 227-243.
- Kemper, C. J., Beierlein, C., Kovaleva, A. & Rammstedt, B. (2012). Entwicklung und Validierung einer ultrakurzen Operationalisierung des Konstrukts Optimismus-Pessimismus - Die Skala Optimismus-Pessimismus-2 (SOP2). Manuskript eingereicht zur Publikation.
- Kovaleva, A., Beierlein, C., Kemper, C. J., & Rammstedt, B. (2012). Eine Vier-Item-Skala zur Erfassung von internaler und externaler Kontrollüberzeugung (IE-4) (GESIS Working Papers 2012|19). Köln: GESIS.
- Krohne, H. W. & Hock, M. (2007). *Psychologische Diagnostik: Grundlagen und Anwendungsfelder*. Stuttgart: Kohlhammer.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M. & Smoot, M. (2000). Maxims of myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, 126, 390-423.
- Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID). (Hrsg.). (2011). *Verzeichnis Testverfahren. Kurznamen. Langnamen. Autoren. Testrezensionen* (18., aktualisierte Auflage). Trier: ZPID [Online verfügbar unter: [http://www.zpid.de/pub/tests/verz\\_teil1.pdf](http://www.zpid.de/pub/tests/verz_teil1.pdf); Stand: 30.6.2011].
- Lienert, G. A. & Raatz, U. (1998). Testaufbau und Testanalyse (6. Aufl.). Grundlagen Psychologie. Weinheim: Julius Beltz.
- Mueser, K. T., Grau, B. W., Sussman, S. & Rosen, A. J. (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. & Beaver, K. M. (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., Rhodes, G. & Simmons, L. W. (2007). Contributions of the face and body to overall attractiveness. *Animal Behaviour, 73*, 937-942.
- Rammstedt, B. (2010a). Subjective indicators. In *Building on progress. Expanding the research infrastructure for the social, economic, and behavioral sciences* (S. 813-824). Opladen: Budrich UniPress.
- Rammstedt, B. (2010b). Reliabilität, Validität, Objektivität. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 239-258). Wiesbaden: VS.
- Rammstedt, B. & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*, 203-212.
- Ramsey, J. L., Langlois, J. H., Hoss, R. A., Rubenstein, A. J. & Griffin, A. (2004). Origins of a Stereotype: Categorization of Facial Attractiveness by 6-Month-Old Infants. *Developmental Science, 7*, 201-211.
- Rasmussen, H. N., Scheier, M. F. & Greenhouse, J. B. (2009). Optimism and physical health: A meta-analytic review. *Annals of Behavioral Medicine, 37*, 239-256.
- Rat für Sozial- und Wirtschaftsdaten (2010). *Building on progress: Expanding the research infrastructure for the social, economic, and behavioral sciences*. Opladen: Budrich UniPress.
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review Of Psychology, 57*, 199-226.
- Schmidt, F. L. & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.
- Schumann, S. & Schoen, H. (2005). *Persönlichkeit: Eine vergessene Größe der empirischen Sozialforschung*. Wiesbaden: VS Verlag.
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.
- Stanton, J. M., Sinar, E. F., Balzer, W. K. & Smith, P. C. (2002). Issues and strategies for reducing the length of self-report scales. *Personnel Psychology, 55*, 167-194.
- Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*, 401-426.
- Thornhill, R. & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences, 3*, 452-460.
- Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I. & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia, 45*, 195-206.

## Appendix A

---

### Referenzwerte

Tabelle A: Referenzwerte für das AR1

Geschlecht	Bildung	Altersgruppen						Gesamt (Alter)	
		18-35		36-65		>65		M	SD
		M	SD	M	SD	M	SD		
Männlich	gering	1.23	1.80	1.05	2.15	.83	1.95	.97	2.01
	mittel	.75	2.32	1.57	1.99	.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	.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 (Geschlecht)	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

Anmerkungen: N = 1134.

## Appendix B

### AR1

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.



-5	-4	-3	-2	-1	0	1	2	3	4	5
viel unattrak- tiver					gleich attraktiv					viel attrakti- ver

---

## Appendix C

---

### Anleitung zum Download der Ankerbilder aus der Datenbank FACES

Um die AR1 nutzen zu können, benötigen Sie die sechs Ankerbilder. Aus rechtlichen Gründen können wir diese nicht direkt an Sie weitergeben. Sie können sich aber ganz einfach kostenlos bei der Datenbank FACES anmelden und dort die entsprechenden Bilder herunterladen.

#### Wichtige Informationen zur Nutzung:

- Die FACES Datenbank darf nur für wissenschaftliche Zwecke und nur unter strikter Einhaltung des FACES Database Release Agreement genutzt werden. Das heißt unter anderem:
  - Das Instrument darf nicht an Dritte weitergegeben werden (Ausnahme: Personen/Institutionen, die mit der Datenerhebung beauftragt sind)
  - Sie dürfen die Bilder nur mit dem AR1 verwenden
  - Sie dürfen die Bilder nicht verändern
  - Jede Nutzung des AR1/der Bilder muss beantragt werden
- Bei der Anmeldung müssen Sie eine Affiliation an einer anerkannten wissenschaftlichen Einrichtung nachweisen
- Sollten Sie die AR1 einsetzen zitieren Sie bitte neben dem AR1 [Kemper, C. J., Lutz, J., Margraf-Stiksrud, J., Beierlein, C., Kovaleva, A., & Rammstedt, B. (2012). Eine Ein-Item-Skala zur Einschätzung von Attraktivität: Das Attraktivitätsrating (AR1) (GESIS Working Papers 2012|24). Köln: GESIS] auch die FACES Datenbank, aus der das Bildmaterial stammt [Ebner, N. C., Riediger, M., & Lindenberger, U. (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. doi:10.3758/BRM.42.1.351]

#### Was ist zu tun, wenn Ihre Erhebung von einem Erhebungsinstitut durchgeführt wird?

- In diesem Fall ist die Weitergabe des AR1/der Bilder an das Erhebungsinstitut erlaubt
- Geben Sie dies aber unbedingt im FACES-Anmeldeformular an
- Sie müssen dafür Sorge tragen (optimaler Weise durch vertragliche Vereinbarung), dass das Erhebungsinstitut das AR1/die Bilder nicht weitergibt und sich auch ansonsten an das FACES Database Release Agreement hält

#### Wie können Sie die Bilder downloaden?

##### Schritt 1: Account erstellen

1. Besuchen Sie die Seite <http://faces.mpib-berlin.mpg.de>
2. Klicken Sie oben rechts auf „Registration“
3. Laden Sie das Anmeldeformular (PDF) herunter, indem Sie im Text auf „application form“ klicken.
4. Füllen Sie das Formular aus und senden Sie es an den FACES Technical Agent ([faces-database@mpib-berlin.mpg.de](mailto:faces-database@mpib-berlin.mpg.de)). (Falls die Erhebung von einem Erhebungsinstitut durchgeführt wird, vergessen Sie bitte nicht, dies im Formular anzugeben!)

5. Sobald Ihr Zugang erstellt ist, erhalten Sie eine E-Mail mit den Zugangsdaten

*Schritt 2: Bilder finden*

1. Sobald Ihr Zugang erstellt ist, gehen Sie wieder auf <http://faces.mpib-berlin.mpg.de>. Klicken Sie oben rechts auf „Login“ und loggen Sie sich mit den zugesendeten Zugangsdaten ein.
2. Klicken Sie im oberen rechten Bereich auf „advanced search“
3. Geben Sie folgende Suchkriterien an:
  - *Emotion*: neutrality
  - *Picture Set*: A
  - *ID*: Hier geben Sie eine ID an, um die richtige Person auszuwählen (167 für jung/männlich, 054 für jung/weiblich, 178 für mittleres Alter/männlich, 080 für mittleres Alter/weiblich, 027 für hohes Alter/männlich, 133 für hohes Alter/weiblich)  
Anmerkung: Vergessen Sie nicht, bei zweistelligen IDs eine 0 voranzustellen.  
...und klicken Sie unten rechts auf „search“. Als Suchergebnis sollen Sie genau ein Bild erhalten
4. Klicken Sie auf das Bild
5. Klicken Sie erneut auf das Bild, um es zu vergrößern
6. Speichern Sie das Bild ab, z.B. mit Rechtsklick -> „Speichern unter...“