
Der "Global Physical Activity Questionnaire" (GPAQ) im Licht des Omnibus 2011 und der Validierungsstudie 2015

Hanspeter Stamm, Rahel Bürgi und Markus Lamprecht
Lamprecht und Stamm Sozialforschung und Beratung AG
Forchstrasse 212 • 8032 Zürich • info@LSSFB.ch

Schlussbericht

Zürich, 6.11.2015

Inhalt

Zusammenfassung	1
Résumé	1
1. Ausgangslage und Fragestellung	2
2. Datenlage und Vorgehensweise	3
3. Resultate von Omnibus 2011 und Validierungsstudie im Vergleich	6
4. Diskussion und Folgerungen	10
Literaturhinweise	14
Anhang: Fragen des Global Physical Activity Questionnaire (GPAQ) im Omnibus 2011 des Bundesamts für Statistik	15

Zusammenfassung

Die vorliegende Studie ergänzt eine kürzlich vom Institut für Epidemiologie, Biostatistik und Prävention der Universität Zürich durchgeführte Validierungsstudie zum Global Physical Activity Questionnaire (GPAQ). Die Validierungsstudie, die auf einer schriftlichen Befragung und Akzelerometermessungen bei rund 350 Personen basierte, wird dabei mit der Omnibusbefragung 2011 des Bundesamts für Statistik verglichen, die telefonische Interviews mit über 5000 Personen umfasste.

Während die Befragungsresultate in der Validierungsstudie um knapp das Dreifache über den Akzelerometermessungen liegen, fallen die Befunde aus dem Omnibus 2011 noch einmal höher aus: Sie übertreffen die Resultate aus der Befragung der Validierungsstudie um knapp das Doppelte und liegen rund fünfmal höher als die Befunde aus der Akzelerometermessung. Mit Bezug zur ebenfalls erfassten Dauer des Sitzens liegen die Resultate aus dem Omnibus dagegen unter den Befunden aus der Validierungsstudie. Die Unterschiede zwischen den beiden Befragungen lassen sich nur schwer erklären, sie dürften jedoch aus einer Kombination verschiedener Faktoren entstanden sein (unterschiedliches Befragungsformat, Stichprobenunterschiede, Zeitpunkt der Befragung etc.).

Trotz des unterschiedlichen Gesamtniveaus der körperlichen Aktivität sind die Resultate der beiden Studien in dem Sinne konsistent, als ähnliche Unterschiede in der körperlichen Aktivität nach Alter, Geschlecht und Sprachregion nachgewiesen werden können. Entsprechend muss die in der Validierungsstudie nachgewiesene, akzeptable Validität des GPAQ zwar nicht grundsätzlich in Frage gestellt werden. Für die Verwendung in Studien, in denen nicht nur Differenzen zwischen verschiedenen sozialen Gruppen, sondern auch effektive Aktivitätsniveaus von Interesse sind, scheint der GPAQ im Licht der beträchtlichen Unterschiede in den Resultaten jedoch ungeeignet zu sein.

Résumé

La présente étude complète une étude de validation relative au *Global Physical Activity Questionnaire* (GPAQ) réalisée récemment par l’Institut d’épidémiologie, de biostatistique et de prévention de l’université de Zurich. Établie sur la base d’un questionnaire écrit et de mesures effectuées au moyen d’un accéléromètre auprès d’environ 350 personnes, l’étude de validation y est comparée à l’enquête Omnibus 2011 de l’Office fédéral de la statistique menée par téléphone auprès de plus de 5000 personnes.

Alors que les résultats d’enquête de l’étude de validation dépassent de trois fois les mesures de l’accéléromètre, ceux de l’Omnibus 2011 sont encore plus élevés. Ils dépassent de près du double les résultats de l’enquête de l’étude de validation et d’environ cinq fois les résultats des mesures de l’accéléromètre. Quant aux données également saisies sur la durée en position assise, les résultats de l’Omnibus sont par contre inférieurs à ceux de l’étude de validation. Les différences entre les deux enquêtes sont difficilement explicables mais devraient provenir d’une combinaison de différents facteurs (formats d’enquête différents, différences au niveau des échantillons, période de l’enquête, etc.).

Malgré la différence du niveau général d’activité corporelle, les résultats des deux études sont cohérents dans le sens qu’ils attestent des différences similaires au niveau de l’activité corporelle selon l’âge, le sexe et la région linguistique. Par conséquent, il n’est pas nécessaire de remettre en question la validité du GPAQ, définie comme acceptable selon l’étude de validation. Toutefois, pour une utilisation dans des études qui ne limitent pas leur champ de recherche aux différences entre divers groupes sociaux mais examinent aussi les niveaux d’activité effectifs, il semble que le GPAQ ne soit pas approprié en raison des importantes différences constatées quant aux résultats.

1. Ausgangslage und Fragestellung

Seit Jahrzehnten werden verschiedene Möglichkeiten zur Erhebung der körperlichen Bewegung mittels Fragebogen diskutiert. In Ermangelung von Geräten (Akzelero- und Inklinometer), mit denen die Aktivität direkt gemessen werden konnte, stellten Befragungen bis weit in die 2000er Jahre den "Königsweg" dar, um an Massendaten zum Bewegungsverhalten zu gelangen. Daran hat sich auch heute wenig geändert. Mittlerweile existiert zwar ein breites Angebot an Messgeräten, ihre Anwendung ist aber nach wie vor mit verschiedenen Problemen behaftet und vergleichsweise aufwendig. Entsprechend stellt sich gerade für grosse Bevölkerungsstudien weiterhin die Frage, wie die Bewegungsaktivität möglichst einfach und effizient erfragt werden könnte.

Ende der 1990er Jahre schien mit dem International Physical Activity Questionnaire (IPAQ) ein konsensfähiges Fragebogeninstrument zur Erfassung der körperlichen Aktivität gefunden worden zu sein. Ein Problem mit dem IPAQ bestand jedoch darin, dass er in zwei Versionen vorlag: Die kurze Version (IPAQ short) ist in Befragungen zwar gut einsetzbar, erfasst die Aktivität jedoch relativ grob und lässt keine Rückschlüsse auf den Bereich zu, in welchem Aktivitäten ausgeübt werden. Die lange Version (IPAQ long) erfasst das Bewegungsverhalten dagegen sehr differenziert, verlangt von den Befragten jedoch ein hohes Mass an Geduld und Konzentration. Vor diesem Hintergrund hat die WHO vor einigen Jahren den Global Physical Activity Questionnaire (GPAQ) entwickelt, der gleichsam eine mittlere Version des IPAQ darstellt und damit die Lösung des Widerspruchs zwischen einfacher Anwendbarkeit und differenzierter Datenerhebung verspricht.

Sowohl beim IPAQ als auch beim GPAQ und weiteren Fragenbatterien stellt sich jedoch die Frage, ob sie tatsächlich das messen, was sie zu messen vorgeben. In den vergangenen Jahren wurde eine Vielzahl von Validierungsstudien vorgelegt, die ein gemischtes Bild zeigen. Zwar scheinen die verschiedenen Instrumente einigermaßen valide und reliabel zu sein sowie verhältnismässig stabile Resultate zu generieren und die körperliche Aktivität tatsächlich in dem Sinne abzubilden, dass aktiver Personen über ein höheres Bewegungsniveau berichten als weniger aktive. Insgesamt scheinen die Fragebogeninstrumente das effektive Bewegungsverhalten im Vergleich zu direkten Messungen mittels Akzelerometer – die ihrerseits verschiedene Probleme beinhalten – jedoch deutlich zu überschätzen. Während sich dieser Befund bezüglich des IPAQ bereits mehrfach zeigte, gibt es erst wenige Validierungsstudien zum GPAQ.

Vor diesem Hintergrund hat eine Arbeitsgruppe des Instituts für Epidemiologie, Biostatistik und Prävention der Universität Zürich, finanziell unterstützt durch das Bundesamt für Gesundheit (BAG), kürzlich eine Validierungsstudie des GPAQ für die Schweiz durchgeführt (Wanner et al. 2015). Ausgehend von einer schriftlichen Datenerhebung im Rahmen des Ernährungspanels der ETH Zürich wurde eine Teilstichprobe von knapp 400 Personen für eine Woche mit Akzelerometern ausgerüstet, um das Bewegungsverhalten direkt zu erfassen.

Die Untersuchung kommt zum Schluss, dass der GPAQ "im schriftlichen Format [...] als valide eingestuft" werden könne, selbst wenn er das Aktivitätsniveau gegenüber den Akzelerometerdaten um das 2.8-fache überschätzt (Wanner et al. 2015: 3, 34). Aus diesem Grund müssten die "absoluten Werte, welche aus dem GPAQ resultieren, vorsichtig interpretiert werden". Der GPAQ scheint insbesondere intensive körperliche Aktivitäten relativ gut abzubilden, während er bei den moderaten Aktivitäten und den täglichen Wegstrecken weniger gut abschneidet. Letzteres dürfte nicht zuletzt damit zusammenhängen, dass die verwendeten Geräte nicht in der Lage sind, Wege, die mit dem Fahrrad zurückgelegt werden, zuverlässig zu erfassen. Ebenso räumen die Autor/innen ein, dass eine Unterscheidung von Settings (Arbeit, Freizeit, Verkehr) mit der Akzelerometermessung im Gegensatz zum Fragebogeninstrument nicht möglich sei (vgl. Wanner et al. 2015: 34).

Die Validierungsstudie von Wanner et al. (2015: 31ff.) enthält einen kurzen Vergleich mit anderen Validierungsstudien zum GPAQ und IPAQ sowie kurze Hinweise auf die Befunde aus zwei Validierungsstudien, welche vom Untersuchungsteam zum IPAQ und zum "Single Item Questionnaire" (SIQ) durchgeführt wurden. Der SIQ scheint bei vergleichenden Akzelerometermessungen dabei bessere

Validitätswerte zu erzielen als der GPAQ und der IPAQ (vgl. Abschnitt 4). Nicht verglichen wurden die Resultate aus der Validierungsstudie des GPAQ jedoch mit anderen schweizerischen Resultaten zum GPAQ, die 2011 im Rahmen des Omnibus des Bundesamts für Statistik (BFS) erhoben wurden. Diese Daten sind deshalb von Interesse, weil sie auf einer grossen repräsentativen Datenerhebung bei über 5'000 Personen basieren und damit detailliertere Analysen der Determinanten und Korrelate körperlicher Aktivität erlauben.

Vor diesem Hintergrund werden im vorliegenden Bericht die Resultate aus dem Omnibus 2011 mit denjenigen aus der Validierungsstudie von Wanner et al. (2015) verglichen. Dabei können nicht einfach die Resultate aus der bereits existierenden Studie zum Omnibus 2011 verwendet werden (vgl. Wiegand et al. 2012), weil die dort verwendete Aktivitätsklassifikation von derjenigen in der Validierungsstudie abweicht. Zudem müssen für den Vergleich verschiedene weitere Vorkehrungen getroffen werden, die im folgenden Abschnitt 2 diskutiert werden. Abschnitt 3 enthält die Resultate aus dem Vergleich der beiden Studien, die in Abschnitt 4 diskutiert und um einige Folgerungen für die zukünftige Verwendung des GPAQ ergänzt werden.

2. Datenlage und Vorgehensweise

Wie einleitend erwähnt, handelt es sich beim GPAQ um ein relativ neues Befragungsinstrument, welches von der WHO anfangs der 2000er Jahre als Reaktion auf verschiedene Probleme mit dem älteren IPAQ entwickelt wurde. Das Instrument wurde in der Schweiz im Jahr 2011 im Rahmen der ersten Omnibusstudie des Bundesamts für Statistik erstmals eingesetzt.

Mit dem Omnibus wird interessierten Bundesämtern die Gelegenheit gegeben, Themen abzufragen, welche nicht von den grossen, regelmässig stattfindenden Datenerhebungen (etwa der Schweizerischen Gesundheitsbefragung oder den Mikrozensusen) abgedeckt werden. Im Jahr 2011 umfasste der Omnibus neben dem GPAQ (im Auftrag des BAG, vgl. das Muster des Fragebogens im Anhang) Fragen zum Umweltbewusstsein (im Auftrag der Sektion Umwelt des BFS). Die *telefonische* Befragung wurde zwischen Mai und Juli 2011 in deutscher, französischer und italienischer Sprache durchgeführt. Insgesamt wurden 5'129 Personen im Alter von 15 bis 74 Jahren befragt.

Die Daten aus der Validierungsstudie basieren dagegen auf *schriftlichen* Interviews, die zunächst anlässlich des ETH-Ernährungspanels zwischen März und Mai 2014 erhoben wurden. Eine Teilstichprobe derjenigen Befragten, die sich zur Teilnahme an einer Zusatzstudie bereit erklärt hatten, wurden mit einem *schriftlichen* Fragebogen¹ und einem Akzelerometer versorgt. Der Fragebogen enthielt erneut die GPAQ-Fragen, während der Akzelerometer während sieben Tagen getragen werden sollte. Da das ETH-Ernährungspanel nur Personen aus der West- und Deutschschweiz umfasst, wurden zusätzlich Personen aus dem Tessin für die Validierungsstudie rekrutiert. Die Datenerhebung fand zwischen August 2014 und März 2015 statt. Insgesamt umfasste die Stichprobe 354 Personen im Alter ab 18 Jahren mit vollständigen Daten aus den Akzelerometermessungen und dem GPAQ-Fragebogen sowie zu den sozio-demografischen Merkmalen.

Die Resultate aus der Validierungsstudie und zum GPAQ sind in verschiedenen Berichten dokumentiert (vgl. Wiegand et al. 2012, Wanner 2015). Überdies wurden die Daten zum Umfang des Sitzens in einer jüngeren Studie erneut genutzt (Stamm et al. 2015). Die Resultate aus den beiden Omnibus-Analysen können jedoch nicht direkt für den Vergleich mit der Validierungsstudie von Wanner et al. (2015) verwendet werden, weil der Fokus der Validierungsstudie auf dem Vergleich mit den Akzelerometerdaten und nicht auf einer möglichst detaillierten Analyse der GPAQ-Fragen lag. Als Folge wurden die Angaben aus dem GPAQ in der Validierungsstudie so zusammengefasst, dass sie den

¹ Die ursprüngliche, auch im Omnibus 2011 verwendete Version des GPAQ basiert auf einem persönlichen oder telefonischen Interview. Diese Version wurde für das ETH-Ernährungspanel auf eine schriftliche Befragung adaptiert.

direkten Akzelerometermessungen möglichst entsprechen. Konkret existieren im Bericht von Wanner et al. (2015) die folgenden Variablen, die mittels der Rohdaten aus dem Omnibus 2011 und den Analyserichtlinien der WHO (ohne Jahr) für den vorliegenden Bericht rekonstruiert wurden:²

- Gesamter Bewegungsumfang pro Woche (Minuten): Diese Angabe enthält die Summe aller fünf im GPAQ erfassten Bewegungskomponenten (Arbeit intensiv und moderat, Freizeit intensiv und moderat, Wegstrecken) pro Woche. Es handelt sich somit um die Angabe, wie viele Minuten pro Woche sich eine Person gemäss GPAQ bewegt.
- Gesamter Bewegungsumfang pro Woche (MET-Minuten): Hier werden die Angaben zu den einzelnen Bewegungskomponenten mit den metabolischen Äquivalenten (MET) gewichtet. Gemäss Analysevorgaben der WHO (ohne Jahr) werden die moderaten und körperlich aktiven Transportaktivitäten (Gehen zu Fuss oder Radfahren) dabei mit dem Faktor 4, die intensiven Aktivitäten dagegen mit dem Faktor 8 multipliziert.
- Gesamtumfang intensiver Aktivitäten pro Woche (Minuten): Diese Variable enthält die Summe der intensiven Aktivitäten in der Arbeit und der Freizeit pro Woche (ohne MET-Gewichtung).
- Gesamtumfang moderater Aktivitäten pro Woche (Minuten): Für diese Variable wurden neben den moderaten Aktivitäten in Arbeit und Freizeit auch die Transportaktivitäten mitberücksichtigt (ohne MET-Gewichtung).
- Gesamtdauer des Sitzens pro Woche (Std.): Diese Angabe bezieht sich schliesslich auf eine einfache Frage zur Dauer des Sitzens an einem "gewöhnlichen" Tag. Diese Zahl liegt im GPAQ auf Tagesbasis in Minuten vor, wurde für den Vergleich jedoch mit 7 multipliziert und durch 60 geteilt, um zu einer Schätzung für die wöchentliche Sitzdauer (in Stunden) zu gelangen.³

Im Bericht von Wanner et al. (2015: 64ff.) werden zu diesen Variablen arithmetische Mittelwerte und Standardabweichungen sowie Medianwerte und Interquartilsabstände nach Geschlecht, Sprachregion und Altersgruppe aufgeführt. Diese Analysen können mit dem Omnibus 2011 grundsätzlich repliziert werden.

Bei der Aufbereitung und Interpretation der Resultate stellen die unterschiedlichen Stichproben jedoch eine erhebliche Herausforderung dar. Aus der Datenbeschreibung der Validierungsstudie geht hervor, dass Personen ab 18 Jahren untersucht werden, während im Omnibus 2011 15- bis 74-Jährige Personen befragt wurden. Während es möglich ist, die 15- bis 17-Jährigen im Omnibus aus den vergleichenden Analysen auszuschliessen, gibt es keine Möglichkeit, die über 74-Jährigen in der Validierungsstudie auszuschliessen, da uns nur die publizierten Resultate vorliegen. Wie ein Blick auf den oberen Teil von Tabelle 2.1 zeigt, dürfte der Anteil betagter Befragten in der Validierungsstudie jedoch vergleichsweise gering sein: Immerhin stimmt das Durchschnittsalter und auch der Anteil der 60-jährigen und älteren Personen in den beiden Untersuchungen recht gut überein.

Dasselbe gilt gemäss Tabelle 2.1 auch für die Geschlechterverteilung, nicht aber für die Sprachregionen, bei denen sich die Validierungsstudie durch eine stark disproportionale Stichprobe für den Tessin auszeichnet. Auch im Omnibus war die Stichprobe für die italienische Schweiz aus

² Grundsätzlich wäre es möglich, detailliertere Variablen zu bilden, welche auch die Aktivität während Arbeit, Freizeit und auf den täglichen Wegen abbildet. Dazu müssten jedoch die Daten aus der Validierungsstudie übernommen und neu codiert werden, was relativ aufwendig wäre. Für den vorliegenden Vergleich, welcher auf den im schriftlichen Schlussbericht von Wanner et al. (2015) publizierten Resultaten basiert, reichen jedoch die vereinfachten Variablen.

³ Soweit sich dies beurteilen lässt, sind die Angaben zur Sitzdauer die einzigen, in denen die Validierungsstudie von den Analyserichtlinien der WHO (ohne Jahr) abweicht. Die WHO weist die Sitzdauer pro Tag aus, während die Validierungsstudie den wöchentlichen Sitzumfang berechnet. Da sich Arbeitstage und Wochenende erheblich unterscheiden können, ist fraglich, ob die Berechnung eines wöchentlichen Sitzumfangs sinnvoll ist. Zwecks Vergleichs mit den Akzelerometermessungen wäre es allenfalls angemessener gewesen, diese auf einen täglichen Durchschnitt zurückzurechnen.

Fallzahlgründen zwar etwas zu gross, die gewichteten Daten deuten jedoch darauf hin, dass die Angaben in der Validierungsstudie die effektiven Verhältnisse um das Zehnfache übertreffen.⁴

Tabelle 2.1: Merkmale der Stichproben der Validierungsstudie und des Omnibus 2011

	Validierungs studie	Omnibus 2011 ungewichtet	Omnibus 2011 gewichtet
Alter			
18 bis 39 Jahre	37.3%	33.8%	39.3%
40 bis 59 Jahre	39.5%	42.5%	40.6%
60 Jahre und älter (Omnibus bis 74 Jahre)	23.2%	23.7%	20.0%
Altersdurchschnitt (Jahre)*	47.0	46.2	44.6
Geschlecht			
weiblich	50.6%	52.6%	50.4%
männlich	49.4%	47.4%	49.6%
Sprachregion			
Deutschschweiz	33.6%	69.5%	71.4%
Französische Schweiz	24.0%	21.2%	24.0%
Italienische Schweiz (Validierungsstudie: Tessin)	42.4%	9.3%	4.6%
Fallzahl*	354	4853	

* Werden die 15- bis 17-Jährigen im Omnibus 2011 mitberücksichtigt, so beträgt der Altersdurchschnitt 44.6 (ungewichtet) bzw. 43.0 Jahre (gewichtet). Die Fallzahl beträgt 5'129.

Diese Befunde haben für die folgenden Analysen einige Konsequenzen: Für eine möglichst zuverlässige Analyse müssten im Omnibus eigentlich die gewichteten Werte verwendet werden, weil diese eine Reihe von kleineren Stichprobenverzerrungen (z.B. die vergrösserte Stichprobe in der italienischen Schweiz) korrigieren. Geht man jedoch von systematischen Unterschieden zwischen den Sprachregionen aus, so müsste auf die Verwendung der Gewichtungsfaktoren verzichtet werden. Im folgenden werden daher für den Omnibus jeweils sowohl gewichtete als auch ungewichtete Resultate berichtet.

Ein weiteres Problem betrifft die Durchführung von statistischen Tests zwecks Klärung der Frage, ob gefundene Unterschiede zwischen dem Omnibus und der Validierungsstudie signifikant sind. Dabei gilt es festzuhalten, dass in der Validierungsstudie nur punktuell Vertrauensintervalle und andere Teststatistiken angegeben wurden. Für die arithmetischen Mittelwerte konnten die 95%-Vertrauensintervalle jedoch nachträglich geschätzt werden⁵, während sie für den Omnibus 2011 direkt mit dem verwendeten Statistikprogramm (SPSS, Prozedur "Complex Samples") berechnet wurden.

⁴ Die Detailtabelle 1 im Anhang der Validierungsstudie (Wanner et al. 2015: 62) deutet überdies darauf hin, dass sich die Alters- und Geschlechterverteilungen zwischen den drei Sprachregionen unterscheiden. So zeichnen sich die Deutsch- und Westschweiz durch einen höheren Frauenanteil aus, während die Gruppe der 60-Jährigen und Älteren in der italienischen Schweiz sehr klein (8%) und diejenige der unter 40-Jährigen überdurchschnittlich gross ist (53%). Diese letztere Gruppe ist dagegen in der Westschweiz zugunsten der mittleren Altersgruppe (48%) verhältnismässig klein (17%).

⁵ Für die Schätzung des Vertrauensintervalls (VI) wurde die folgende Formel verwendet:

$$VI = +/- (s/\sqrt{n}) * t \quad \text{mit } s: \text{Standardabweichung}$$

n: ungewichtete Stichprobengrösse
t: t-Schätzwert

Der Schätzwert für t variiert nach Stichprobengrösse. Bei Teilstichproben zwischen n=80 und n=354 liegt er im Bereich von 1.99 bis 1.97. Für grosse Stichproben mit n > 500 liegt der Wert bei 1.96.

Schliesslich muss darauf hingewiesen werden, dass sich zusätzlich Unterschiede zwischen den beiden Befragungen ergeben haben können, weil unterschiedlich und zu verschiedenen Zeitpunkten befragt wurde (telefonisch vs. schriftlich) und die Teilnehmer/innen der Validierungsstudie wegen der bereits einmal absolvierten GPAQ-Fragenbatterie (im Rahmen des Ernährungspanels) und anstehenden Akzelerometermessung möglicherweise sensibilisierter waren für die Frage nach ihrem Bewegungsverhalten. Auf diese Problematik wird im Schlussabschnitt zurückzukommen sein (Abschnitt 4).

3. Resultate von Omnibus 2011 und Validierungsstudie im Vergleich

Die Tabellen 3.1 bis 3.3 zeigen verschiedene Gegenüberstellungen von Validierungsstudie und Omnibus 2011. In Tabelle 3.1 sind die arithmetischen Mittelwerte und Vertrauensintervalle für die Akzelerometermessung und den GPAQ-Fragebogen aus der Validierungsstudie sowie die ungewichteten und gewichteten Resultate aus dem GPAQ des Omnibus 2011 enthalten, während Tabelle 3.2 die entsprechenden Medianwerte und Interquartilsabstände enthält. Beide Tabellen enthalten neben den Gesamtwerten für die verschiedenen Aktivitätsvariablen auch die Detailresultate nach Geschlecht, Alter und Sprachregion und sind daher etwas unübersichtlich. Aus diesem Grund enthält Tabelle 3.3 eine vereinfachte Darstellung, aus der hervorgeht, um welches Vielfache die Resultate aus dem Omnibus 2011 (nur gewichtete Daten) die Akzelerometermessung und die GPAQ-Befunde aus der Validierungsstudie sowie diese letzteren Angaben die Akzelerometermessungen über- oder unterschätzen.

Ein Blick auf die drei Tabellen zeigt vor allem ein wichtiges Resultat: Die Resultate des Omnibus 2011 liegen deutlich und signifikant höher als diejenigen aus dem GPAQ der Validierungsstudie, welche ihrerseits höher liegen als die Akzelerometermessungen. In aller Regel überschneiden sich die Vertrauensintervalle (Tabelle 3.1) nicht. Bei den Interquartilsabständen gibt es zwar verschiedene Überlagerungen, doch erstaunt dies angesichts des grossen Streubereichs der Angaben nicht.

Beschränken wir die Betrachtung auf die Gesamtaktivität, so fällt auf, dass die Befragten des Omnibus 2011 über körperliche Aktivitäten im Umfang von rund 26 Std. pro Woche (über 1500 Minuten) berichten, während sich diese Werte in der Validierungsstudie auf rund die Hälfte (rund 850 Minuten) und in der Akzelerometermessung sogar nur auf rund einen Fünftel (rund 300 Minuten) belaufen. Wie dem unteren Teil der Tabellen zu entnehmen ist, sind die Abweichungen jedoch je nach Aktivitätskategorie sehr unterschiedlich. Der Omnibus scheint vor allem intensive Aktivitäten stark zu überschätzen (Faktor 2.8 gegenüber der Validierungsstudie und Faktor 15.1 gegenüber der Akzelerometermessung), während die Angaben zu den moderaten Aktivitäten und zur Sitzdauer besser übereinstimmen. Bei der Sitzdauer fällt überdies auf, dass diese im Omnibus geringer ausfällt als in der Validierungsstudie und der Akzelerometermessung. Insgesamt sind die Angaben zur körperlichen Aktivität im Omnibus also deutlich höher, zum Sitzen dagegen deutlich geringer als in der Validierungsstudie und der Akzelerometermessung.

Lenkt man den Blick auf die Detailresultate zu den verschiedenen Komponenten der körperlichen Aktivität, so zeigt sich jedoch gesamthaft betrachtet eine recht hohe Konsistenz der Angaben zwischen Omnibus 2011 und Validierungsstudie in dem Sinne, dass die beiden Studien in der Regel bei denselben Bevölkerungsgruppen höhere oder geringere Werte ausweisen. So sind die Geschlechterunterschiede relativ gering, wobei sich die Frauen durch ein geringeres Mass an intensiven und ein höheres Mass an moderaten Aktivitäten auszeichnen. Beim Alter erweisen sich die über 60-Jährigen im Omnibus 2011 als eher etwas aktiver als in der Validierungsstudie, was vor allem auf die höheren Werte für die intensiven Aktivitäten zurückzuführen ist. Und bei den Sprachregionen gibt es zwischen der Validierungsstudie und dem Omnibus 2011 verschiedene deutliche Abweichungen bei der italienischen Schweiz, was wohl auf die spezielle regionale Stichprobe der Validierungsstudie zurückzuführen ist (vgl. Abschnitt 4). Insgesamt gelangt man bezüglich sozialer Unterschiede im Aktivitäts- und Sitzverhalten jedoch zu vergleichbaren Resultaten unabhängig davon, ob man sich auf

die Akzelerometermessungen, die Validierungsstudie oder den Omnibus 2011 bezieht. Dies ist ein erster Hinweis darauf, dass die Messung mittels GPAQ zwar nicht sehr zuverlässig, jedoch einigermassen valide sein dürfte.

Tabelle 3.1: Vergleich der Resultate aus der Validierungsstudie und dem Omnibus 2011, Mittelwerte und 95%-Vertrauensintervalle

	Validierungsstudie		Omnibus 2011	
	Akzelerometer	GPAQ	GPAQ ungewich.	GPAQ gewichtet
Gesamtaktivität (Min.)	301 (284-318)	853 (753-953)	1551 (1511-1591)	1534 (1487-1581)
weiblich	290 (265-315)	849 (713-984)	1572 (1518-1625)	1535 (1475-1595)
männlich	311 (288-334)	857 (709-1004)	1529 (1470-1588)	1533 (1461-1604)
18-39 Jahre	293 (267-319)	719 (566-872)	1541 (1473-1609)	1498 (1420-1576)
40-59 Jahre	315 (288-342)	779 (639-920)	1588 (1524-1653)	1596 (1520-1674)
60 Jahre und älter	287 (246-328)	1194 (940-1447)	1499 (1426-1573)	1478 (1397-1560)
Deutschschweiz	296 (266-326)	892 (711-1073)	1679 (1629-1728)	1647 (1590-1704)
französische Schweiz	279 (246-312)	767 (563-970)	1232 (1156-1307)	1228 (1144-1312)
italienische Schweiz	315 (288-342)	871 (720-1021)	1308 (1195-1422)	1343 (1117-1569)
Gesamtaktivität (MET-Min.)	-	4197 (3717-4677)	8352 (8118-8587)	8305 (8028-8583)
weiblich	-	4019 (3321-4717)	8119 (7821-8418)	7957 (7620-8293)
männlich	-	4380 (3721-5039)	8606 (8239-8972)	8656 (8214-9097)
18-39 Jahre	-	3899 (3107-4691)	8430 (8028-8832)	8219 (7752-8686)
40-59 Jahre	-	3753 (3121-4385)	8552 (8174-8930)	8635 (8183-9087)
60 Jahre und älter	-	5434.4 (4212-6656)	7878 (7440-8316)	7811 (7325-8297)
Deutschschweiz	-	4239 (3359-5120)	9039 (8748-9330)	8891 (8558-9224)
französische Schweiz	-	3450 (2553-4348)	6665 (6214-7116)	6695 (6175-7215)
italienische Schweiz	-	4587 (3848-5327)	6963 (6280-7646)	7453 (5914-8992)
Intensive Aktivitäten (Min.)	36 (29-43)	197 (164-230)	537 (513-560)	542 (514-570)
weiblich	31 (22-40)	156 (110-202)	458 (430-486)	454 (422-486)
männlich	42 (32-52)	238 (192-284)	623 (585-660)	631 (585-677)
18-39 Jahre	44 (33-55)	256 (195-318)	566 (526-607)	556 (509-604)
40-59 Jahre	42 (30-54)	159 (123-195)	550 (513-587)	562 (516-608)
60 Jahre und älter	13 (3-23)	165 (86-244)	470 (426-514)	474 (426-523)
Deutschschweiz	25 (14-36)	168 (115-222)	581 (552-610)	576 (543-609)
französische Schweiz	20 (11-30)	96 (52-140)	435 (390-480)	445 (390-501)
italienische Schweiz	54 (42-66)	276 (219-333)	432 (361-503)	520 (338-702)
Moderate Aktivitäten (Min.)	264 (249-279)	656 (568-744)	1014 (986-1043)	991 (957-1025)
weiblich	259 (236-282)	693 (565-821)	1114 (1072-1155)	1081 (1035-1127)
männlich	269 (249-289)	619 (498-740)	906 (868-945)	902 (853-951)
18-39 Jahre	248 (226-270)	462 (338-586)	975 (926-1023)	942 (887-998)
40-59 Jahre	273 (250-296)	621 (490-752)	1039 (992-1085)	1035 (978-1092)
60 Jahre und älter	275 (235-315)	1029 (810-1248)	1029 (978-1080)	1004 (946-1062)
Deutschschweiz	271 (244-298)	724 (573-875)	1097 (1062-1133)	1071 (1029-1113)
französische Schweiz	259 (229-289)	671 (480-862)	797 (745-849)	783 (725-841)
italienische Schweiz	262 (238-286)	595 (462-728)	876 (793-959)	823 (679-967)
Sitzen (Std.)	61.8 (60.9-62.7)	51.5 (48.9-52.0)	35.3 (34.6-36.0)	36.3 (35.5-37.1)
weiblich	59.8 (58.6-61.0)	48.5 (45.0-52.0)	32.5 (31.6-33.3)	34.1 (33.1-35.2)
männlich	63.7 (62.4-65.0)	54.5 (50.8-58.2)	38.3 (37.3-39.4)	38.5 (37.3-39.7)
18-39 Jahre	63.1 (61.6-64.6)	52.9 (48.5-57.3)	38.7 (37.5-39.9)	39.3 (37.9-40.7)
40-59 Jahre	61.0 (59.6-62.4)	52.9 (49.0-56.8)	36.0 (35.0-37.1)	36.6 (35.4-37.9)
60 Jahre und älter	60.8 (59.0-62.6)	46.8 (41.7-51.9)	29.0 (28.0-30.0)	29.6 (28.3-30.8)
Deutschschweiz	61.4 (59.7-63.1)	50.7 (46.4-55.0)	36.6 (35.9-37.4)	37.6 (36.7-38.5)
französische Schweiz	61.1 (59.5-62.7)	53.7 (48.1-59.3)	33.0 (31.6-34.4)	33.8 (32.1-35.5)
italienische Schweiz	62.3 (61.0-63.6)	50.9 (47.0-54.8)	30.2 (28.2-32.3)	29.7 (25.2-34.2)

Tabelle 3.2: Vergleich der Resultate aus der Validierungsstudie und dem Omnibus 2011, Medianwerte und Interquartilsabstände

	Akzelerometer	GPAQ	GPAQ ungewich.	Omnibus 2011
				GPAQ gewichtet
Gesamtaktivität (Min.)	287 (183-407)	512 (270-960)	1080 (535-2280)	1050 (520-2220)
weiblich	286 (161-385)	480 (270-945)	1185 (570-2265)	1110 (540-2220)
männlich	297 (197-419)	570 (270-1040)	1020 (480-2280)	1020 (490-2280)
18-39 Jahre	278 (184-383)	405 (243-765)	1050 (480-2340)	990 (480-2210)
40-59 Jahre	304 (202-402)	500 (275-853)	1080 (520-2340)	1080 (520-2340)
60 Jahre und älter	266 (143-419)	775 (440-1380)	1170 (600-2100)	1140 (580-2100)
Deutschschweiz	281 (165-407)	480 (270-1110)	1230 (580-2520)	1185 (565-2460)
französische Schweiz	275 (159-392)	420 (240-870)	840 (415-1710)	840 (420-1680)
italienische Schweiz	303 (197-414)	600 (305-960)	960 (465-1830)	960 (420-1920)
Gesamtaktivität (MET-Min.)	-	2670 (1440-4960)	5520 (2640-11670)	5400 (2640-11640)
weiblich	-	2400 (1320-4360)	5760 (2880-11520)	5580 (2740-11280)
männlich	-	3000 (1560-5760)	5280 (2520-11940)	5280 (2520-12180)
18-39 Jahre	-	2230 (1300-4800)	5520 (2520-11790)	5280 (2480-11560)
40-59 Jahre	-	2640 (1440-4400)	5520 (2640-12170)	5480 (2640-12240)
60 Jahre und älter	-	3320 (1920-7200)	5760 (2880-10680)	5640 (2760-10680)
Deutschschweiz	-	2560 (1120-5280)	6240 (3000-12960)	6000 (2880-12720)
französische Schweiz	-	2040 (1000-3840)	4320 (2000-8880)	4320 (2000-8760)
italienische Schweiz	-	3240 (1760-5160)	4560 (2240-9180)	4400 (2100-10360)
Intensive Aktivitäten (Min.)	4 (0-45)	120 (0-240)	240 (60-600)	240 (60-600)
weiblich	2 (0-31)	50 (0-180)	210 (20-540)	210 (30-540)
männlich	7 (0-59)	150 (0-300)	270 (60-720)	270 (90-720)
18-39 Jahre	15 (0-68)	133 (60-300)	240 (90-630)	240 (90-600)
40-59 Jahre	4 (0-66)	90 (0-240)	240 (60-600)	240 (60-600)
60 Jahre und älter	0 (0-5)	0 (0-180=)	180 (0-600)	180 (0-630)
Deutschschweiz	1 (0-18)	75 (0-180)	270 (80-270)	240 (90-240)
französische Schweiz	0 (0-11)	0 (0-120)	180 (0-480)	180 (0-480)
italienische Schweiz	14 (0-90)	180 (30-360)	138 (0-480)	180 (0-660)
Moderate Aktivitäten (Min.)	248 (158-340)	360 (150-765)	690 (300-1440)	660 (300-1380)
weiblich	248 (137-340)	390 (190-780)	810 (370-1560)	765 (360-1515)
männlich	248 (169-356)	300 (120-765)	600 (245-1260)	585 (240-1260)
18-39 Jahre	221 (156-313)	210 (110-480)	630 (269-1410)	600 (260-1350)
40-59 Jahre	277 (164-361)	360 (145-673)	674 (300-1440)	660 (300-1410)
60 Jahre und älter	254 (129-409)	720 (330-1320)	815 (419-1440)	780 (385-780)
Deutschschweiz	254 (156-360)	420 (165-960)	770 (345-1560)	720 (330-1500)
französische Schweiz	265 (135-365)	360 (180-720)	540 (240-1080)	525 (255-1020)
italienische Schweiz	235 (161-323)	300 (120-720)	600 (270-1170)	600 (240-1080)
Sitzen (Std.)	62.5 (56.5-67.5)	49.0 (31.5-67.5)	28.0 (17.5-49.0)	18.0 (19.3-56.0)
weiblich	60.0 (54.5-65.2)	45.5 (28.0-66.5)	28.0 (17.5-42.0)	28.0 (17.5-42.0)
männlich	65.1 (59.1-69.1)	56.0 (35.0-70.0)	31.5 (21.0-56.0)	31.5 (21.0-56.0)
18-39 Jahre	64.6 (58.0-69.5)	49.0 (35.0-70.0)	35.0 (21.0-56.0)	35.0 (21.0-56.0)
40-59 Jahre	62.5 (55.3-66.7)	56.0 (35.0-70.0)	28.0 (17.5-49.0)	28.0 (21.0-56.0)
60 Jahre und älter	60.8 (55.2-66.0)	42.0 (28.0-70.0)	28.0 (17.5-35.0)	28.0 (17.5-35.0)
Deutschschweiz	61.6 (56.4-67.1)	49.0 (31.5-70.0)	28.0 (21.0-49.0)	28.0 (21.0-56.0)
französische Schweiz	61.6 (56.7-65.8)	49.0 (35.0-70.0)	28.0 (14.0-42.0)	28.0 (14.0-49.0)
italienische Schweiz	63.0 (56.4-68.1)	49.0 (31.5-70.0)	21.0 (14.0-42.0)	21.0 (14.0-35.0)

Tabelle 3.3: Verhältnis der Resultate aus Akzelerometermessung, Validierungsstudie und Omnibus 2011

	Validierungsstudie vs. Akzelerometer	Omnibus 2011 vs. Akzelerometer	Omnibus 2011 vs. Validierungsstudie
Gesamtaktivität (Min.)	2.8	5.1	1.8
weiblich	2.9	5.3	1.8
männlich	2.8	4.9	1.8
18-39 Jahre	2.5	5.1	2.1
40-59 Jahre	2.5	5.1	2.0
60 Jahre und älter	4.2	5.1	1.2
Deutschschweiz	3.0	5.6	1.8
französische Schweiz	2.7	4.4	1.6
italienische Schweiz	2.8	4.3	1.5
Gesamtaktivität (MET-Min.)	-	-	2.0
weiblich	-	-	2.0
männlich	-	-	2.0
18-39 Jahre	-	-	2.1
40-59 Jahre	-	-	2.3
60 Jahre und älter	-	-	1.4
Deutschschweiz	-	-	2.1
französische Schweiz	-	-	1.9
italienische Schweiz	-	-	1.6
Intensive Aktivitäten (Min.)	5.5	15.1	2.8
weiblich	5.0	14.6	2.9
männlich	5.7	15.0	2.7
18-39 Jahre	5.8	12.6	2.2
40-59 Jahre	3.8	13.4	3.5
60 Jahre und älter	12.7	36.5	2.9
Deutschschweiz	6.7	23.0	3.4
französische Schweiz	4.8	22.3	4.6
italienische Schweiz	5.1	9.6	1.9
Moderate Aktivitäten (Min.)	2.5	3.8	1.5
weiblich	2.7	4.2	1.6
männlich	2.3	3.4	1.5
18-39 Jahre	1.9	3.8	2.0
40-59 Jahre	2.3	3.8	1.7
60 Jahre und älter	3.7	3.7	1.0
Deutschschweiz	2.7	4.0	1.5
französische Schweiz	2.6	3.0	1.2
italienische Schweiz	2.3	3.1	1.4
Sitzen (Std.)	0.8	0.6	0.7
weiblich	0.8	0.6	0.7
männlich	0.9	0.6	0.7
18-39 Jahre	0.8	0.6	0.7
40-59 Jahre	0.9	0.6	0.7
60 Jahre und älter	0.8	0.5	0.6
Deutschschweiz	0.8	0.6	0.7
französische Schweiz	0.9	0.6	0.6
italienische Schweiz	0.8	0.5	0.6

4. Diskussion und Folgerungen

Dass die Angaben aus dem GPAQ des Omnibus 2011 deutlich höher liegen als die Akzelerometermessungen aus der Validierungsstudie konnte angesichts jener Befunde sowie weiterer Untersuchungen zum Zusammenhang direkter Messungen und Fragebogenuntersuchungen zwar erwartet werden. Dass aber die Befragungsresultate des Omnibus 2011 diejenigen aus der Befragung der Validierungsstudie rund um das Zweifache übertreffen, ist doch sehr erstaunlich. Besonders deutlich sind dabei die Abweichungen beim Umfang der intensiven Aktivitäten, welche auch die Gesamtschätzung für die MET-Werte in die Höhe treiben. Aber auch bei den moderaten Aktivitäten und beim Sitzen, wo der Omnibus eine geringere Dauer ausweist als die Validierungsstudie, finden sich substantielle und signifikante Unterschiede.

Es gibt verschiedene mögliche Erklärungen für die Abweichungen, die im folgenden kurz diskutiert werden sollen:

- Stichprobenverzerrungen: Während die Stichprobe des Omnibus 2011 repräsentativ für die 15- bis 74-Jährige Bevölkerung der Schweiz ist, dürfte die Validierungsstudie tatsächlich einige Verzerrungen aufweisen. Für eine genauere Analyse wären jedoch zusätzliche Angaben notwendig. Wahrscheinlich scheinen Verzerrungen bei den Resultaten aus der Sonderstichprobe des Tessins, die nicht nur stark vom Omnibus 2011 abweichen, sondern sich auch erheblich von den Resultaten der Validierungsstudie für die Deutsch- und Westschweiz unterscheiden. Selbst wenn die Stichproben der Validierungsstudie relativ klein und allenfalls etwas verzerrt sind, erklärt dies die grossen Abweichungen zwischen den beiden Untersuchungen jedoch nicht.
- Befragungszeitraum: Die Datenerhebung zum Omnibus 2011 fand zwischen Mai und Juli 2011 statt, während die Datenerhebung zum ETH-Ernährungspanel zwischen März und Mai 2014 und die Validierungsstudie zwischen August 2014 und März 2015 stattfanden. Dies dürfte einen Teil der gefundenen Abweichungen erklären, da ausserhäusliche Bewegungsaktivitäten aus klimatischen Gründen im Winter eher seltener sind als im Sommer. Dass der Bewegungsumfang im Sommer das Doppelte desjenigen im Winter beträgt, scheint uns aber eher unwahrscheinlich.
- Befragungsinstrument und -technik: Während im Omnibus 2011 telefonisch befragt wurde, basierte die Validierungsstudie auf einer schriftlichen Befragung mittels Papierfragebogen. Geht man die beiden Fragebogen durch, so sind keine Unterschiede zwischen den Fragenformulierungen erkennbar, welche die grossen Unterschiede zwischen den beiden Studien erklären würden. Allerdings gilt es darauf hinzuweisen, dass der GPAQ von der WHO für persönliche Interviews entwickelt wurden, wobei der Schritt vom persönlichen zum telefonischen Interview grundsätzlich kleiner ist als derjenige vom persönlichen zum schriftlichen Fragebogen.⁶

Die Aufzeichnungen des mit der Omnibusbefragung betrauten Befragungsinstituts deuten jedoch darauf hin, dass der vergleichsweise lange Einleitungstext zu den Fragen nach den körperlichen Aktivitäten, in dem der Unterschied zwischen intensiven und moderaten Aktivitäten erklärt wird, von verschiedenen Befragten als kompliziert und schwer verständlich wahrgenommen wurde (vgl. Anhang). Damit scheint es möglich, dass Verständnisprobleme zu einer systematischen Verzerrung der Antworten geführt haben. Dieses Problem dürfte in einer schriftlichen Befragung geringer sein, weil die befragten Personen hier die Möglichkeit haben, den Einleitungstext während der Beantwortung der nachfolgenden Fragen erneut zu konsultieren. Allerdings erklären Verständnisproblem

⁶ Entsprechend ist die Empfehlung in der Validierungsstudie kritisch zu beurteilen, es müsste "basierend auf der von der WHO entwickelten Originalversion des GPAQ (persönliche Interviews) und der hier verwendeten schriftlichen Version [...] von Experten eine für die Schweiz sinnvolle Telefoninterview-Version ausgearbeitet werden". Für den Omnibus 2011 wurde bereits eine telefonische Version entwickelt, die sich eng an der Originalversion orientiert. U.E. gibt es keinen Grund, bei einer allfälligen Optimierung des Befragungsinstruments den Umweg über eine neue schriftliche Version zu nehmen, welche die internationale Vergleichbarkeit weiter erschwert. Allerdings müsste bei telefonischen und persönlichen Befragungen in Einklang mit den internen Empfehlungen des Befragungsinstituts, das den Omnibus 2011 durchführte, erneut am Einleitungstext der Frage gearbeitet werden (siehe Text).

die deutlich höheren Angaben im Omnibus nicht, denn es wäre auch denkbar, dass solche Schwierigkeiten zu unvollständigen und damit tieferen Angaben führen.

- Befragungskontext: Interessant sind in diesem Zusammenhang jedoch die Befunde zum Vergleich zwischen erster Befragung im Rahmen des Ernährungspanels und der Befragung anlässlich der Akzelerometermessung in der Validierungsstudie (vgl. Wanner 2015: 27ff., 69ff.): In der ursprünglichen Befragung lagen die GPAQ-Resultate signifikant höher bzw. die Angaben zur Sitzdauer tiefer als in der zweiten Befragung, was auf einen "Sensibilisierungseffekt" bei den Befragten und eine geringe Reliabilität der Messung hinweist. Die Abweichungen betreffen dabei insbesondere die intensiven körperlichen Aktivitäten, welche im Omnibus ja ebenfalls deutlich höher ausfallen als in der Validierungsstudie. Allerdings liegen die Angaben zur körperlichen Aktivität aus der ersten Befragung deutlich unter den Befunden des Omnibus 2011, so dass dieser Effekt höchstens einen Teil der Unterschiede erklärt.⁷
- Daten-, Codierungs- und Berechnungsfehler: Angesichts der grossen Abweichungen zwischen der Validierungsstudie und dem Omnibus 2011 wurden die Daten und Codierungen der verwendeten Variablen noch einmal geprüft. Mit Blick auf die Rohdaten wurde das BFS kontaktiert, das – abgesehen von den oben erwähnten Problemen mit dem Einleitungstext zum GPAQ – keine Unregelmässigkeiten bei der Datenerhebungen festgestellt hatte. Auch unsere eigenen Kontrollen der Rohdaten deuten darauf hin, dass bei der Datenerhebung und Datenerfassung keine systematischen Fehler passiert sind. Dies gilt auch für die verwendeten Codierungen und Berechnungen, die im Omnibus 2011 exakt den Analysevorgaben der WHO (ohne Jahr) entsprechen.⁸ Über die Datenqualität und Berechnungen in der Validierungsstudie sind zwar keine Aussagen möglich. Da die GPAQ-Werte der Validierungsstudie aber insgesamt näher bei internationalen Vergleichsdaten (vgl. <https://apps.who.int/infobase/Indicators.aspx>) liegen als diejenigen aus dem Omnibus, gehen wir davon aus, dass die Codierungen und Berechnungen in der Validierungsstudie ebenfalls korrekt sind.

Insgesamt reichen die einzelnen diskutieren Erklärungsansätze nicht aus, um die deutlichen Unterschiede zwischen Validierungsstudie und Omnibus 2011 vollständig zu erklären. Die verschiedenen potentiellen Probleme könnten aber insgesamt zu einer Kumulation verschiedener Überschätzungen im Omnibus und Unterschätzungen in der Validierungsstudie und damit zu den im Aggregat doch sehr beträchtlichen Unterschieden geführt haben. Selbst wenn der Befund einer befriedigenden Validität des GPAQ aus der Validierungsstudie durch die vorliegenden Befunde nicht grundsätzlich in Frage gestellt wird, muss ein Fragezeichen hinter die Objektivität und Zuverlässigkeit (Reliabilität) der Messungen gesetzt werden.

Mit Blick auf die *Objektivität* gilt es zunächst festzustellen, dass die Abweichungen zwischen Akzelerometermessungen und Befragungen nicht überbewertet werden sollten, messen die beiden Zugänge doch einerseits eine vermeintlich objektives Verhalten und andererseits subjektive Einschätzungen. Dabei gilt es zunächst zu beachten, dass auch die Akzelerometermessung mit verschiedenen Problemen behaftet sein kann (vgl. Bürgi 2010). So vermögen die verwendeten Geräte beispielsweise Wege mit dem Fahrrad oder Bewegungsepisoden im Wasser nicht abzubilden. Zudem hängt die Messung davon ab, ob das Gerät ständig und korrekt getragen wird. Und schliesslich werden die Resultate von den verwendeten Aufzeichnungsintervallen und Grenzwerten für geringe, moderate und intensive Aktivitäten beeinflusst. So werden kurze Bewegungsepisoden (z.B. das Benutzen der Treppe

⁷ Gemäss Wanner et al. (2015) liegt beispielsweise die Schätzung für die gesamte körperliche Aktivität pro Woche in der Erstbefragung bei 940 Minuten (Validierungsstudie gemäss Anhangtabelle 7 im Bericht: 821 Min./Woche; Omnibus 2011: 1534 Min./Woche). Bei den intensiven Aktivitäten weist die Erstbefragung einen Wert von 222 Min./Woche aus (Validierungsstudie gemäss Anhangtabelle 7 im Bericht: 132 bzw. 197 Min./Woche; Omnibus 2011: 542 Min./Woche).

⁸ Die Ausnahme bildet die Angabe zur Sitzdauer, welche von der WHO als täglicher Wert, in der Validierungsstudie jedoch pro Woche ausgewiesen wird. Die Angaben pro Tag wurden für den Omnibus jedoch ebenfalls auf der Grundlage der WHO-Richtlinien berechnet.

statt des Lifts), die gerade als Unterbrechung von Sitzepisoden durchaus gesundheitsförderlich sein können, in den Messungen tendenziell unterschätzt.⁹ Mit Blick auf die Intensität von Bewegungs-episoden brauchen Befragungen dabei nicht grundsätzlich schlechter abzuschneiden als Akzelerometermessung, da eine Aktivität von verschiedenen Personen tatsächlich unterschiedlich erlebt werden kann: Eine untrainierte Person mag einen zügigen Spaziergang bereits als anstrengend empfinden, während ein trainierter Sportler bei dieser Aktivität kaum ausser Atem gerät. Vor diesem Hintergrund wäre es heikel, die höheren Werte aus der Befragung generell als "Überschätzung" des Bewegungsverhaltens zu interpretieren. Die Vermutung, dass der Bewegungsumfang mittels Akzelerometermessung möglicherweise unterschätzt wird, legt den Schluss nahe, dass der "wahre Wert" wohl zwischen den Angaben aus der Messung und der Befragung liegen dürfte. Wo der "wahre Wert" tatsächlich liegt, lässt sich beim aktuellen Kenntnisstand jedoch nicht sagen.

Bezüglich der *Zuverlässigkeit* der Angaben muss jedoch die Frage gestellt werden, was mit einem Instrument, das trotz befriedigender Validitätswerte je nach Anwendung und Stichprobe sehr unterschiedliche Resultate erbringt, angefangen werden kann, wenn es um die Einschätzung des Bewegungsverhaltens geht. Vor diesem Hintergrund teilen wir die einleitend erwähnte Einschätzung der Validitätsstudie nicht, dass der GPAQ zwar durchaus verwendet werden könne, aber die "absoluten Werte, welche aus dem GPAQ resultieren, vorsichtig interpretiert werden" müssten (Wanner et al. 2015: 34). Die vorliegenden Resultate deuten darauf hin, dass die Befunde aus einer telefonischen Befragung ungefähr durch zwei bzw. durch fünf geteilt werden müssten, um zu vergleichbaren Befunden zu gelangen, wie sie aus schriftlichen Befragungen bzw. Akzelerometermessungen resultieren. Weil aber nicht klar ist, welches der korrekteste Wert ist, wie gross die Abweichungen bei neuen Studien sind und ob allenfalls andere als die hier dargestellten unabhängigen Variablen (Geschlecht, Alter, Sprachregion) einen Einfluss auf die Einschätzung haben (z.B. Arbeitskontext, Familiensituation), kann dieser "Korrekturfaktor" nicht zuverlässig angegeben werden. Insgesamt muss der GPAQ zum gegenwärtigen Zeitpunkt eher als unbefriedigendes Instrument bezeichnet werden, wenn es nicht nur darum geht, Rangfolgen der körperlichen Aktivität zu untersuchen, sondern auch Aussagen über effektive Aktivitätsniveaus zu machen.

Dagegen stimmen wir mit der Empfehlung der Validierungsstudie überein, dass an den bisher verwendeten Merkmalen zur Erfassung der körperlichen Aktivität in bestehenden Studien (insbesondere den Aktivitätsfragen in der Schweizerischen Gesundheitsbefragung SGB) mit Blick auf die Beschreibung von Zeittrends unbedingt festgehalten werden sollte. Gerade mit Blick auf die SGB würden wir als Ergänzung aber weniger den GPAQ, die lange Version des IPAQ (International Physical Activity Questionnaire) oder den Single Item Questionnaire, als vielmehr die kurze Version des IPAQ bevorzugen, weil diese bereits im schriftlichen Teil SGB 2012 verwendet wurde und im internationalen Rahmen häufiger benutzt wird als die doch recht unhandliche lange Version.

Für die Verwendung des kurzen IPAQ und der konventionellen Bewegungsfragen aus der SGB sprechen überdies die in Tabelle 4.1 dargestellten Vergleichsdaten aus der SGB 2012. Der kurze IPAQ unterscheidet ebenso wie der GPAQ zwischen intensiven und moderaten Aktivitäten sowie zwischen den Wegstrecken zu Fuss (nicht aber mit dem Fahrrad), nimmt aber keine Unterscheidung zwischen Arbeits- und Freizeitbereich vor. Die älteren Bewegungsfragen aus der SGB sind nur mit Blick auf die moderaten Aktivitäten mit den Angaben aus GPAQ und IPAQ vergleichbar.¹⁰ Aus der Tabelle wird jedoch deutlich, dass die Angaben aus der SGB 2012 deutlich tiefer liegen als diejenigen, welche sich im Omnibus 2011 mittels GPAQ ermittelten liessen. Vergleicht man die Angaben in Tabelle 4.1 mit denjenigen in Tabelle 3.1, so wird überdies deutlich, dass die Resultate deutlich in die Nähe der

⁹ Umgekehrt entfällt beim Akzelerometer jedoch die Einschränkung im GPAQ- und IPAQ-Fragebogen, nur Episoden von mindestens zehn Minuten Länge zu berücksichtigen, woraus sich bei den Akzelerometermessungen eine Überschätzung der Aktivitätsdauer ergeben könnte. Ob die Daten in der Validierungsstudie auf 10-Minuten Intervalle beschränkt wurden, entzieht sich unserer Kenntnis.

¹⁰ Die Frage nach den intensiven Aktivitäten wird explizit auf die Freizeit eingeschränkt und bezieht sich nur auf die Anzahl Tage mit "Schwitzepisoden", nicht aber auf die Dauer dieser Episoden.

Befunde aus der Validierungsstudie rücken. Der IPAQ dürfte damit die körperliche Aktivität weniger stark überschätzen als der GPAQ und damit letzten Endes zuverlässigere Resultate liefern.

Tabelle 4.1: Vergleich von Kennwerten der körperlichen Aktivität aus der SGB 2012 und dem Omnibus 2011 des BFS

	SGB 2012* Bewegungsfragen	SGB 2012* IPAQ	Omnibus 2011 GPAQ gewichtet
Gesamtaktivität (Min.)	-	882 (867-898)	1534 (1487-1581)
Gesamtaktivität (MET Min.)	-	4018 (3947-4090)	8305 (8028-8583)
Intensive Aktivitäten (Min.)	-	239 (233-245)	542 (514-570)
Moderate Aktivitäten ohne Wegstrecken (Min.)	359 (348-369)	372 (363-380)	809 (778-840)
Moderate Aktivitäten mit Wegstrecken (Min.)**	-	660 (648-673)	991 (957-1025)
Sitzen (Std. pro Tag)	-	5.4 (5.3-5.4)	5.2 (5.1-5.3)

* 15- bis 74 Jährige; ** IPAQ: Wegstrecken zu Fuss; GPAQ: Wegstrecken zu Fuss oder mit dem Fahrrad.

Bereits anlässlich des Detailberichts zu den Resultaten des GPAQ aus dem Omnibus 2011 wurde im übrigen festgehalten, dass dieser die tatsächliche Bewegungsaktivität wohl deutlich überschätzt (vgl. Wiegand et al. 2012). Damals wurde argumentiert, dass die differenzierte Erfassung verschiedener Intensitäten über unterschiedliche Bereiche hinweg (Arbeit, Freizeit, Wegstrecken) möglicherweise zu einer Kumulation von kleineren Überschätzungen in den einzelnen Teildimensionen führe. Dies könnte auch erklären, weshalb der Single Item Questionnaire in der Validierungsstudie von Wanner et al. (2015) noch besser abschneidet als die längeren Instrumente: Aufgrund der sehr limitierten Information, welche hier erfasst wird, existiert wenig Spielraum für (kumulierte) Überschätzungen. Aufgrund der summarischen Erhebung des Bewegungsverhaltens – es muss lediglich die Anzahl Tage mit mindestens 30 Minuten körperlicher Aktivität pro Woche angegeben werden¹¹ – eignet sich dieses Instrument u.E. jedoch nur bedingt für die Erhebung des körperlichen Aktivitätsniveaus, lässt sich mit ihm doch nicht einmal die Einhaltung der aktuellen Empfehlungen für gesundheitswirksame Bewegung rekonstruieren.¹²

Insgesamt scheint sich der GPAQ trotz akzeptabler Validitätswerte somit nicht zu bewähren. Selbst wenn auch die anderen, aktuell verwendeten Befragungsinstrumente (Fragen aus der SGB, IPAQ) nicht ohne Probleme sind, sollte u.E. vorderhand an diesen Fragen festgehalten werden, weil sie zumindest eine Vielfalt an Vergleichsmöglichkeiten mit anderen Ländern sowie über die Zeit erlauben.

¹¹ Die englische Version der Frage lautet: "In the past week, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate? This may include sport, exercise, and brisk walking or cycling for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job." (vgl. Wanner et al. 2012)

¹² Aktuell werden in der Schweiz mindestens 150 Minuten moderater Bewegung pro Woche bzw. 75 Minuten intensiver Bewegung empfohlen (vgl. Bundesamt für Sport et al. 2013). Die ältere Empfehlung enthielt die Zusatzempfehlung, sich täglich mindestens 30 Minuten zu bewegen. Die ältere Empfehlung hätte sich somit mit dem Single Item Questionnaire abbilden lassen, nicht aber die neue, die auch dann erfüllt ist, wenn nur an einem Tag 75 Minuten intensiv trainiert wird.

Literaturhinweise

Bundesamt für Sport BASPO, Bundesamt für Gesundheit BAG, Gesundheitsförderung Schweiz, bfu – Beratungsstelle für Unfallverhütung, Suva, Netzwerk Gesundheit und Bewegung Schweiz (2013). Gesundheitswirksame Bewegung. Magglingen: BASPO.

Bürgi, R. (2010): Vergleich von zwei Fragebögen zur Erfassung körperlicher Aktivität. Unveröffentlichte Masterarbeit zur Erlangung des Masters of Science ETH in Bewegungswissenschaften. Zürich: ETH:

Stamm, H., D. Wiegand, R. Bürgi, M. Lamprecht: Sitzender Lebensstil. Sonderanalyse des Omnibus 2011 und der Schweizerischen Gesundheitsbefragung 2012 des Bundesamts für Statistik. Studie im Auftrag des Bundesamts für Gesundheit. Bern: BAG.

Wanner, M., N. Probst-Hensch, S. Kriemler, F. Meier, A. Bauman, B.W. Martin (2012): What physical activity surveillance needs: validity of a single-item questionnaire. British Journal of Sport Medicine (doi: 10.1136).

Wanner, M., E. Martin-Diener, G. Pestoni, C. Hartmann, M. Siegrist, B. Martin (2015): Validierung des Global Physical Activity Questionnaire (GPAQ) in der Schweiz. Zürich: Institut für Epidemiologie, Biostatistik und Prävention. (Version vom 28.7.2015).

WHO (ohne Jahr): Global Physical Activity Questionnaire (GPAQ). Analysis Guide. Genf: WHO.

Wiegand, D., H. Stamm und M. Lamprecht (2012): Analyse von Fragen zum Bewegungsverhalten im Omnibus 2011 des Bundesamtes für Statistik. Studie im Auftrag des Bundeamtes für Gesundheit. Bern: BAG.

Anhang: Fragen des Global Physical Activity Questionnaire (GPAQ) im Omnibus 2011 des Bundesamts für Statistik

Q.301.00

Jetzt kommen wir zu den körperlichen Aktivitäten, wo Sie innerhalb von einer normalen Woche verrichten.

Dabei unterscheiden wir zwischen...

o INTENSIVEN KÖRPERLICHEN AKTIVITÄTEN, d.h. Aktivitäten von mindestens 10 Minuten, bei denen Atmung und Puls stark zunehmen, z.B. Tragen und Heben von schweren Lasten, Bauarbeiten, im Garten oder beim Sport, ... und...

o GEMÄSSIGTEN KÖRPERLICHEN AKTIVITÄTEN, wo mindestens 10 Minuten dauern und bei denen Atmung und Herzschlag nur leicht zunehmen, z.B. zügiges Laufen, Tragen von leichten Lasten, Staubsaugen.

Q.301.01

Zuerst reden wir über die körperlichen Aktivitäten, wo Sie während dem "Schaffen" verrichten. Wenn wir vom "Schaffen" reden, meinen wir ALLES, WO SIE erledigen MÜSSEN, also ALLES, WO NICHT ZUR FREIZEIT ODER zur FORTBEWEGUNG gehört.

Wir meinen damit also auch alle AUSBILDUNGSAKTIVITÄTEN, alle AUFGABEN IM HAUSHALT ODER im GARTEN, FAMILIENAUFGABEN aber selbstverständlich auch alle BEZAHLTE, UNBEZAHLTE oder FREIWILLIGE ARBEIT, d.h. z.B. auch Kursleitungen, Gemeindearbeit.

Q.301.10

Wenn Sie jetzt an all diese verschiedenen Arbeiten denken, wo Sie in einer normalen Woche machen: Beinhaltet Ihre Arbeit dann auch INTENSIVE KÖRPERLICHE AKTIVITÄTEN?

INT: INTENSIVEN KÖRPERLICHEN AKTIVITÄTEN, d.h. Aktivitäten von mindestens 10 Minuten, bei denen Atmung und Puls STARK zunehmen

<1> Ja, verrichte intensive körperliche Arbeit

<2> Nein, verrichte KEINE intensive körperliche Arbeit

<7> Bin behindert, kann keine intensive körperliche Arbeit verrichten

<8> Weiss nicht

<9> Keine Angabe

Q.302.00

An wie vielen Tagen in einer NORMALEN WOCHE machen Sie intensive körperliche Arbeit?

<1> 1 Tag pro Woche mit intensiver körperlicher Arbeit

<2> 2 Tage pro Woche

<3> 3 Tage pro Woche

<4> 4 Tage pro Woche

<5> 5 Tage pro Woche

<6> 6 Tage pro Woche

<7> 7 Tage pro Woche mit intensiver körperlicher Arbeit

<8> Weiss nicht

<9> Keine Angabe

Q.303.00

Und im Durchschnitt, wie viel Zeit verbringen Sie an einem NORMALEN TAG mit intensiver körperlicher Arbeit?

<hh><mm> Stunden | Minuten

<98><98> Weiss nicht

<99><99> Keine Angabe

Q.304.00

Und verrichten Sie während dem Schaffen, also während den erwähnten Tätigkeiten, auch GEMÄSSIGTE KÖRPERLICHE AKTIVITÄTEN?

INT: SCHAFFEN = ALLES AUSSER FORTBEWEGUNG UND FREIZEIT!

INT: GEMÄSSIGTE AKTIVITÄTEN = Aktivitäten wo mind. 10 Min. dauern und bei denen Puls und Atmung LEICHT zunehmen.

<1> Ja, verrichte gemässigte körperliche Arbeit

<2> Nein, verrichte KEINE gemässigte körperliche Arbeit

<7> Bin behindert, kann keine körperliche Arbeit verrichten

<8> Weiss nicht

<9> Keine Angabe

Q.305.00

An wie vielen Tagen in einer NORMALEN WOCHE machen Sie gemässigte körperliche Arbeit?

- <1> 1 Tag pro Woche mit gemässigter körperlicher Arbeit
 - <2> 2 Tage pro Woche
 - <3> 3 Tage pro Woche
 - <4> 4 Tage pro Woche
 - <5> 5 Tage pro Woche
 - <6> 6 Tage pro Woche
 - <7> 7 Tage pro Woche mit gemässigter körperlicher Arbeit
-
- <8> Weiss nicht
 - <9> Keine Angabe

Q.306.00

Und im Durchschnitt, wie viel Zeit verbringen Sie an einem NORMALEN TAG mit gemässigter körperlicher Arbeit?

<hh> <mm> Stunden | Minuten

-
- <98> <98> Weiss nicht
 - <99> <99> Keine Angabe

Q.310.00

Bei den nächsten Fragen geht es jetzt nur um die FORTBEWEGUNG, das heisst, wie Sie in einer normalen Woche von einem Ort zum anderen kommen, z.B. von zu Hause zur Arbeit, zur Ausbildung, zum Einkaufen, zum Markt oder zur Kirche.

Bitte zählen Sie alle Aktivitäten, wo Sie mir vorher schon beim "Schaffen" berichtet haben, NICHT dazu.

Q.311.00

Machen Sie in einer NORMALEN WOCHE Fusswege oder Velowege von mindestens zehn Minuten, um von einem Ort zum anderen zu kommen?

-
- <1> Ja, mache Fusswege oder Velowege von mindestens 10 Minuten
 - <2> Nein, mache KEINE Fusswege oder Velowege von mind. 10 Minuten
-
- <7> Bin behindert, kann nicht zu Fuss gehen / Velo fahren
 - <8> Weiss nicht
 - <9> Keine Angabe

Q.312.00

An wie vielen Tagen in einer NORMALEN WOCHE gehen Sie mindestens zehn Minuten zu Fuss oder fahren Sie mindestens zehn Minuten mit dem Velo, um von einem Ort zum anderen zu kommen?

- <1> 1 Tag pro Woche mit Fuss-/Velowegen von mind. 10 Min. Dauer
 - <2> 2 Tage pro Woche
 - <3> 3 Tage pro Woche
 - <4> 4 Tage pro Woche
 - <5> 5 Tage pro Woche
 - <6> 6 Tage pro Woche
 - <7> 7 Tage pro Woche mit Fuss-/Velowegen von mind. 10 Min. Dauer
-
- <8> Weiss nicht
 - <9> Keine Angabe

Q.313.00

Und im Durchschnitt, wie viel Zeit verbringen Sie an einem NORMALEN TAG mit Fusswegen oder Velowegen, um von einem Ort zum anderen zu kommen?

<hh> <mm> Stunden | Minuten

-
- <98> <98> Weiss nicht
 - <99> <99> Keine Angabe

Q.320.00

Und jetzt geht es um die restliche Zeit in einer normalen Woche, d.h. um Ihre Freizeit.

Bitte zählen Sie alle Aktivitäten, wo Sie mir vorher schon beim "Schaffen" und bei der Fortbewegung berichtet haben, NICHT dazu.

Q.321.00

Betreiben Sie in der Freizeit INTENSIVE KÖRPERLICHE AKTIVITÄTEN oder SPORT, d.h. Aktivitäten oder Sport von mindestens zehn Minuten, bei der Atmung und Puls stark zunehmen.

- <1> Ja, betreibe intensive körperliche Aktivität/Sport in Freizeit
- <2> Nein, betreibe KEINE int. körperliche Aktivität/Sport in Freizeit

<7> Bin behindert, kann keine körperliche Aktivität verrichten

<8> Weiss nicht

<9> Keine Angabe

Q.322.00

An wie vielen Tagen in einer NORMALEN WOCHE machen Sie intensive körperliche Aktivität oder Sport in der Freizeit?

- <1> 1 Tag pro Woche mit intensiver körperlicher Aktivität/Sport
- <2> 2 Tage pro Woche
- <3> 3 Tage pro Woche
- <4> 4 Tage pro Woche
- <5> 5 Tage pro Woche
- <6> 6 Tage pro Woche
- <7> 7 Tage pro Woche mit intensiver körperlicher Aktivität/Sport

<8> Weiss nicht

<9> Keine Angabe

Q.323.00

Und im Durchschnitt, wie viel Zeit verbringen Sie an einem NORMALEN TAG mit intensiver körperlicher Aktivität oder Sport in der Freizeit?

<hh> <mm> Stunden | Minuten

<98> <98> Weiss nicht

<99> <99> Keine Angabe

Q.324.00

Und wenn Sie jetzt nur an GEMÄSSIGTE KÖRPERLICHE AKTIVITÄTEN oder SPORT denken, d.h. an Aktivitäten von mindestens zehn Minuten, wo Atmung und Puls leicht zunehmen.

Betreiben Sie dann in der Freizeit (auch) GEMÄSSIGTE KÖRPERLICHE AKTIVITÄTEN oder SPORT?

- <1> Ja, betreibe gemässigte körperliche Aktivität/Sport in Freizeit
 - <2> Nein, betreibe KEINE gemässigte körperliche Aktivität/Sport in Freizeit
-
- <7> Bin behindert, kann keine körperliche Aktivität verrichten
 - <8> Weiss nicht
 - <9> Keine Angabe

Q.325.00

An wie vielen Tagen in einer NORMALEN WOCHE machen Sie gemässigte körperliche Aktivität oder Sport in der Freizeit?

- <1> 1 Tag pro Woche mit gemässigter körperlicher Aktivität/Sport
- <2> 2 Tage pro Woche
- <3> 3 Tage pro Woche
- <4> 4 Tage pro Woche
- <5> 5 Tage pro Woche
- <6> 6 Tage pro Woche
- <7> 7 Tage pro Woche mit gemässigter körperlicher Aktivität/Sport

<8> Weiss nicht

<9> Keine Angabe

Q.326.00

Und im Durchschnitt, wie viel Zeit verbringen Sie an einem NORMALEN TAG mit gemässigter körperlicher Aktivität oder Sport in der Freizeit?

<hh> <mm> Stunden | Minuten

<98> <98> Weiss nicht

<99> <99> Keine Angabe

Q.330.00

Und jetzt geht es noch um die Zeit, wo Sie mit SITZEN oder RUHEN verbringen, sei das bei der Arbeit, zu Hause, unterwegs oder bei Freunden. Denken Sie dabei z.B. ans Sitzen am Schreibtisch, ans Zusammensitzen mit Freunden, im Auto, im Bus oder im Zug, ans Kartenspielen oder ans Fernsehen. Bitte zählen Sie die Zeit, wo Sie mit Schlafen verbringen, NICHT dazu.

Wie viel Zeit verbringen Sie an einem NORMALEN TAG mit Sitzen oder Ruhen?

<hh> <mm> Stunden | Minuten

<98> <98> Weiss nicht

Zusatzfragen:

Q.331.00

Können Sie mir sagen, wie gross Sie ohne Schuhe sind?

<XXX> cm (3-stellig)

<998> weiss nicht

<999> keine Antwort

Q.332.00

Und wie schwer sind Sie?

INT: FÜR SCHWANGERE FRAUEN:

Wie schwer waren Sie zu Beginn der Schwangerschaft?

INT: Auf Nachfrage => Gewicht ohne Kleidung angeben

<XXX> kg (3-stellig)

<998> weiss nicht

<999> keine Antwort