

Swiss
Insights
News



#4

Dirty on the go?

Dirty on the go?

Datenqualität in Online-Surveys zwischen PC und Mobilgeräten



Raffael Meier
Co-founder und CTO
onlineumfragen.com

Unsere Vergleichsanalyse anonymisierter mobiler und nicht mobiler Teilnahmen einer grossen Schweizer Kundenbefragung zeigt: Mobile Antworten fallen bezüglich Datenqualität leicht ab. Straightliner und kürzere Texte sind eine Herausforderung.

Das Internet in der Tasche mit sich tragen - was vor 25 Jahren eine utopische Vorstellung war, ist heute trivial. Die Einführung des Smartphones hat unser Leben verändert, nicht zuletzt auch die Markt- und Meinungsforschung.

Mit immer besseren Internetverbindungen, grösseren Bildschirmen und einer schnelllebigeren Gesellschaft mit permanenter Aufmerksamkeitsknappheit^[5] stieg das Bedürfnis im letzten Jahrzehnt stark an, Umfragen auch auf Mobilgeräten anbieten und ausfüllen zu können. Diese erhöhte Nachfrage mobiler oder zumindest mobilkompatibler Umfragen schafft nicht nur neue Herausforderungen für das Design und die Handhabung von Onlineumfragen, sondern muss auch hinsichtlich Datenqualität differenziert betrachtet werden.

Anhand anonymisierter Meinungsdaten einer Kundenbefragung eines grossen Schweizer Retailers, die 2014 lanciert wurde und seither täglich neue Rückläufe verzeichnet, analysieren wir die Entwicklung der mobilen Teilnahmen und deren Auswirkung auf die Datenqualität.

Datenqualität - Was ist das eigentlich?

70 Millionen Suchergebnisse bei Google demonstrieren die Relevanz des Begriffs "Data Quality" eindrücklich. Während im Alltag der Begriff "Datenqualität" häufig mit der Richtigkeit (Präzision) von Daten gleichgesetzt wird,

ist Datenqualität in der Markt- und Meinungsforschung komplexer. Hier stehen am Ende der Datenverarbeitung Auftraggebende oder Forschende als Consumer der Daten. Deshalb ist es besonders wichtig, zu verstehen, dass Datenqualität aus Consumer-Sicht oft über die Präzision von Werten hinausgeht und auch Aspekte wie beispielsweise Glaubwürdigkeit, Relevanz, Vollständigkeit, Interpretierbarkeit, Konsistenz und Zugänglichkeit von Daten für Consumer berücksichtigt^[18]. Relevanz zum Beispiel umfasst, dass Daten für ihren beabsichtigten Zweck tatsächlich geeignet sind, denn eine schlechte Datenqualität kann zu fehlerhaften Entscheidungen und ineffizienten Geschäftsprozessen führen, während eine hohe Datenqualität eine fundierte Entscheidungsfindung, bessere Geschäftsprozesse und letztendlich bessere Geschäftsergebnisse unterstützt.

Datenqualität, die in der Literatur mit zahlreichen verschiedenen Modellen beschrieben wird^[6], kann durch geeignete Massnahmen optimiert werden. Einerseits vor der Feldphase, mittels elaboriertem Sampling, stringenter Fragebogenkonzeption und Pretesting. Andererseits in der Analyse und der technischen Bereinigung der bereits gewonnenen Daten. Hinzu kommen Meta-Aspekte der Datenqualität wie Passung der Daten zur Forschungsfrage, Prozesstransparenz, faire Datengewinnung oder proaktiver Datenschutz, die Empowerment für Kundinnen und Kunden bewirken^[10].



Nina Gwerder
Consultant
onlineumfragen.com

Muss bei mobilen Umfragen mit tieferer Datenqualität gerechnet werden?

Die spezifischen Effekte einer mobilen Umfrageteilnahme auf die Datenqualität lassen sich in zwei Bereiche gliedern:

- Effekte, die auf die Natur des Eingabegeräts zurückzuführen sind.
- Effekte, die auf die Situation, in der mobil teilgenommen wird, zurückzuführen sind.

Zum ersten Bereich zählt beispielsweise die kleinere Bildschirmgröße auf Smartphones, welche die Wahrnehmung und das Verständnis der Befragung beeinflussen kann^[8]. So können lange Itembatterien mit breiten Skalen auf dem Smartphone nicht immer in der gewohnten Tabellenform übersichtlich dargestellt werden. Die Darstellung erfolgt daher häufig in einer alternativen Form, z.B. eine Aufteilung auf mehrere Ratingfragen oder Drop-Down-Felder. Dies beeinflusst jedoch die Wahrnehmung der Teilnehmenden: Während das einzelne Item in einer Tabellenfrage in den direkten Kontext der Items davor und danach gesetzt wird, sind bei mehreren Einzelfragen die Inhalte der weiter zurückliegenden Fragen oft nicht mehr präsent. Gleichzeitig können Teilnehmende in der Regel die darauf folgenden Items nicht bereits einsehen, was das kognitive Editierverhalten beeinflusst^{[14][15]}. Zudem sind Umfragen, die auf Mobilgeräten grundlegend anders dargestellt werden als auf grossen Bildschirmen, wie bei einigen Anbietern üblich, methodologisch kritisch zu betrachten, da alle Teilnehmenden möglichst gleiche Stimuli erhalten sollten. Weitere Aspekte, in diesem Bereich sind unter anderem die Navigation via Touchscreen und die allenfalls kleinere Rechenleistung^[11].

Dem zweiten Bereich sind Effekte zuzuordnen, die durch die Möglichkeit der Teilnahme "on the go" ausgelöst werden. Dazu gehören situative Einflüsse (z.B. Unterbrüche, weil man bei der Beantwortung im öffentlichen Verkehr umsteigen muss oder laute Hintergrundgeräusche die Konzentration beeinflussen) sowie intrinsische

Einflüsse, wie eine generell tiefere Aufmerksamkeit, wenn das "Scrollen" am Smartphone sich zu einer passiven Nebenbeschäftigung entwickelt^[19], zum Beispiel, wenn Umfragen während anderen Tätigkeiten wie Medienkonsum, Sitzungen, Vorlesungen, an Bushaltestellen etc. ausgefüllt werden.

Die Forschung zur Datenqualität in Mobile-Surveys steckt noch immer in den Kinderschuhen. Seit 2009 wurden zwar diverse Studien zu diesem Thema veröffentlicht^{[1][3][8][9][13]}, diese zeichnen sich jedoch durch partiell widersprüchliche Ergebnisse, kleine Stichproben und teilweise klinische Settings aus. Einig geworden sind sich die Autorinnen und Autoren dabei dennoch in folgenden zwei Punkten:

- Die Abschlussrate ("completion rate") bei mobilen Teilnahmen ist tiefer als bei der Teilnahme über einen PC / Laptop (d.h. mehr Break-Offs)^{[1][3][8][9]}.
- Die Bearbeitungszeit bei mobilen Teilnahmen ist höher als bei der Teilnahme über einen PC / Laptop^{[3][8]}.

Fallstudie: Kundenbefragung eines grossen Schweizer Retailers

Unsere Analyse erhebt erstmals a) eine durchmischte, reale Zielgruppe (Kundinnen und Kunden eines Retailers) mit b) grosser Datenmenge c) für die Schweiz und hebt sich damit deutlich von bisherigen Studien ab. Sie orientiert sich methodologisch am Vorgehen von Schlosser und Mays^[13], die 2018 den Einfluss von mobilen Teilnahmen auf die Datenqualität an einer Gruppe von 820 deutschen Studierenden untersuchten.

Unsere Analyse greift auf vollständig anonymisierte Daten der Kundenbefragung eines Schweizer Retailers zurück, die 2014 in einem Unternehmensbereich lanciert und anschliessend sukzessiv auf weitere Unternehmensbereiche ausgeweitet wurde. Stand Mai 2023 lagen 345'000 Umfragerückläufe mehrsprachig (de, fr, it) schweizweit vor. Der Fragebogen enthält 6 Fragen. Darunter befinden sich eine Frage zur Weiterempfehlungsbereitschaft, eine

Tabelle mit 5 Items und elfstufiger Likert-Skala, eine Frage mit offener Textantwort und zwei Ja-Nein-Fragen. Für die Beantwortung stehen die im Internet verbreiteten Radiobuttons und Textfelder zur Verfügung. Die mittlere Bearbeitungszeit beläuft sich auf rund 2 Minuten. Es werden keine soziodemografischen Daten erhoben. Als Befragungssoftware wird die Umfrageplattform von [onlineumfragen.com](https://www.onlineumfragen.com) genutzt. Die Kundinnen und Kunden werden in einem mixed-mode Verfahren via E-Mail (ca. 95%) und SMS eingeladen (ca. 5%).

Entwicklung der Relevanz mobiler Teilnahmen 2014 bis 2023

In unserem Sample stieg der Anteil der mobilen Teilnahmen, gemessen über Browser Agent Strings und Bildschirmgrösse, seit 2014 stetig. Mobile Teilnahmen haben sich in der Zeit von 2014 bis 2023 mehr als verdreifacht.

Eine Abschwächung der Entwicklung in den letzten fünf bis sechs Jahren könnte mit einer Sättigung der Zielgruppe zu tun haben: Wer potenziell mobil teilnehmen möchte, verfügt nun auch über die Möglichkeit.

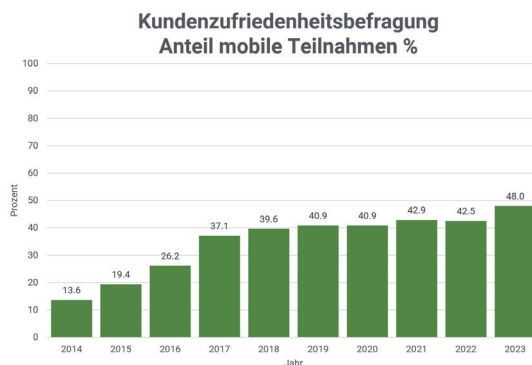


Bild 1: Anteil mobile Teilnahmen Kundenzufriedenheitsbefragung im Retail-Sektor, 2014 bis 2023

Auswirkungen des mobilen Modus auf die Datenqualität

Um Datenqualität in konkreten Aspekten zu messen, wurde ein Subsample aus insgesamt 46'581 aktuellen Teilnahmen herangezogen. Die Eingrenzung erfolgte nach Datum jünger als 01.01.2022. Folgende Kriterien der Datenqualität wurden untersucht:

- Reaktionszeit zur Umfrage-Einladung
- Bearbeitungszeit für eine Frage

- Bearbeitungszeit ganze Umfrage
- Abschlussrate
- Item nonresponse
- Straightliner
- Extrem Response Style (ERS)
- Länge der Antworten auf offene Fragen

Benötigen mobile Teilnehmende mehr Zeit?

Zunächst wurde die **Reaktionszeit** zur Umfrage-Einladung ermittelt. Sie wurde über die Differenz des Zeitstempels der Einladung zur Umfrage (Versandzeitpunkt Einladungs-E-Mail) zum Zeitstempel des Aufrufs der Umfrage (Klick auf Umfragelink) berechnet und auf Ausreisser bereinigt, indem nur Teilnahmen, die innert 7 Tagen ab Einladung erfolgten, berücksichtigt wurden. Da die Teilnehmenden auf Mobilgeräten daran gewöhnt sind, ihr Mobiltelefon ständig bei sich zu tragen und erreichbar zu sein, gehen wir davon aus, dass diese Befragten schneller reagieren als Befragte, die die Einladung über ihren PC erhalten (Hypothese 1).

Auf Mobilgeräten betrug die mittlere Reaktionszeit 13 Stunden ($M=13.15$; $SD=25.43$; $n_{mob}=23'685$), auf Nichtmobilgeräten 22.75 Stunden ($M=22.75$; $SD=34.26$; $n_{pc}=20'481$), die Abweichung ist gemäss Welch-Test $t(37137)=32.86$ signifikant mit $p<0,001$ und die Effektstärke mit Cohen's $d=0.32$ entspricht einem mittleren Effekt.

Wurden zusätzlich Teilnahmen zwischen 7 und 14 Tagen nach Einladung berücksichtigt, ergab sich mobil eine mittlere Reaktionszeit von 17.5 Stunden ($M=17.53$; $SD=40.11$; $n_{mob}=24'169$), auf Nichtmobilgeräten 31 Stunden ($M=31.09$; $SD=53.81$; $n_{pc}=21'335$), die Abweichung ist gemäss Welch-Test $t(38947)=30.08$ signifikant mit $p<0,001$ und die Effektstärke mit Cohen's $d=0.29$ zeigt einen knapp mittleren Effekt.

Die Teilnahmen durch Personen, die Mobilgeräte für die Umfrage nutzen, erfolgen also im Mittel rund 9.6 (resp. 13.6) Stunden früher und damit deutlich näher am Zeitpunkt der Einladung (Hypothese 1 bestätigt). Dies könnte je nach Umfrage einen Einfluss auf Themen wie Erinnerungsleistung, Teilnahmemotivation oder die Emotionalität der Rückmeldungen haben.

Die **mittlere Beantwortungszeit für eine Frage** wurde als Differenz zwischen dem Zeitpunkt der abgeschlossenen Anzeige im Browser und dem Klicken auf den Button "Speichern - nächste Frage" bei einer Tabellenfrage bestehend aus 5 Items mit elfstufiger Likert-Skala zur Zufriedenheit bestimmter Aspekte des Einkaufserlebnisses gemessen. Die Frage wurde mobil und nicht mobil methodologisch identisch präsentiert und nur im Seitenverhältnis der Tabelle, der Breite und der Schriftgrösse im Sinne einer responsiven Darstellung mobil optimiert. Für die mobilen Teilnahmen rechnen wir mit einer erhöhten Beantwortungszeit für komplexere Tabellenfragen^[4], da diese mobil schwieriger zu erfassen sind, die mobile Befragungssituation oft konzentriertes Beantworten erschwert, Fragen etwas weniger leicht zu lesen sind und die technische Erfassung der Antworten (Touchscreen) etwas anspruchsvoller und fehleranfälliger ist (Hypothese 2).

Das arithmetische Mittel der Bearbeitungszeit wurde auf Ausreisser bereinigt (Zeit > 10 Sekunden und < 180 Sekunden) und liegt bei den mobilen Teilnahmen bei 40 Sekunden ($M=40.09$; $SD=23.01$; $n_{mob}=20'333$), bei den nicht mobilen Geräten bei 36 Sekunden ($M=35.70$; $SD=20.53$; $n_{pc}=19'402$).

Die Abweichung ist gemäss Welch-Test $t(39555)=20.09$ signifikant mit $p<0,001$ und die Effektstärke mit Cohen's $d=0.2$ zeigt einen kleinen, aber vorhandenen Effekt. Die Bearbeitungszeit ist auf Nichtmobilgeräten etwas geringer (Hypothese 2 bestätigt).

Die **Bearbeitungszeit für die ganze Umfrage** wurde aus der Differenz der Anzeige der ersten Frage am Bildschirm und des Klickens auf den Button "Speichern" bei der letzten Frage berechnet.

Weil das Ausfüllen einer Umfrage über Mobilgeräte mühsamer und störanfälliger sein kann und auch die mobile Befragungssituation in der Regel mehr Ablenkungen ausgesetzt ist, könnte die Gesamtbearbeitungszeit mobil höher ausfallen. Andererseits könnte durch genau diese situativen Faktoren die Motivation, die Umfrage elaboriert "in Ruhe" und genau auszufüllen, abgeschwächt sein, und es findet vermehrt eine

schnellere, oberflächlichere kognitive Verarbeitung statt, speziell bei Fragen, die sich nicht mit dem individuellen inhaltlichen Feedback-Kern decken und eher als Ballast empfunden werden. Daher vermuten wir in Abwägung dieser Überlegung für die mobilen Teilnahmen eine etwas kürzere Gesamtbearbeitungszeit (Hypothese 3).

Das arithmetische Mittel der Gesamtbearbeitungszeit wurde auf Ausreisser bereinigt (Zeit > 10 Sekunden und < 300 Sekunden) und liegt bei den mobilen Teilnahmen bei 118 Sekunden ($M=118.17$; $SD=63.68$; $n_{mob}=17'435$), bei den nicht mobilen Geräten bei 116 Sekunden ($M=115.73$; $SD=65.40$; $n_{pc}=16'915$).

Die Abweichung ist gemäss Welch-Test $t(34237)=3.50$ signifikant mit $p<0,001$ und die Effektstärke mit Cohen's $d=0.038$ zeigt einen sehr kleinen Effekt. Die Bearbeitungszeit ist auf Mobilgeräten somit minimal länger.

Ein spannendes Bild zeigt sich bei einer weniger starken Bereinigung von Ausreissern (Zeit > 10 Sekunden und < 3600 Sekunden). Dann liegt die Gesamtbearbeitungszeit bei den mobilen Teilnahmen bei 173 Sekunden ($M=172.50$; $SD=237.03$; $n_{mob}=19'619$), bei den nicht mobilen Geräten bei 185 Sekunden ($M=184.69$; $SD=256.90$; $n_{pc}=19'518$). Diese Abweichung ist gemäss Welch-Test $t(38851)=4.88$ signifikant mit $p<0,001$ und die Effektstärke mit Cohen's $d=0.05$ zeigt einen kleinen Effekt, aber nun dauern die PC-Teilnahmen länger. Dies könnte damit zusammenhängen, dass es auf PCs eine höhere Anzahl an Teilnehmenden gibt, die die Umfrage unterbrechen und nach einer Pause (z.B. Mittagspause, Telefongespräch, etc.) fortsetzen. Damit ergeben sich schnell sehr lange (aber nicht andauernd von Aktivität geprägte) Bearbeitungszeiten. Unter Ausschluss von Gesamtbearbeitungszeiten über 5 Minuten sind mobile Teilnahmen also geringfügig langsamer und weisen weniger Pausen auf (Hypothese 3 teilweise abgelehnt). Weshalb bei Einschluss von Gesamtbearbeitungszeiten bis zu einer Stunde die PC-Teilnahmen länger dauern, müsste weiter untersucht werden und könnte auch mit einem höheren Anteil an älteren und weniger IT-affinen Personen in der PC-Gruppe zusammenhängen, sowie auch damit, dass am

PC generell etwas längere Texte bei Textantworten erfasst werden – jedoch werden diese oft auch schneller getippt (siehe Hypothese 8).

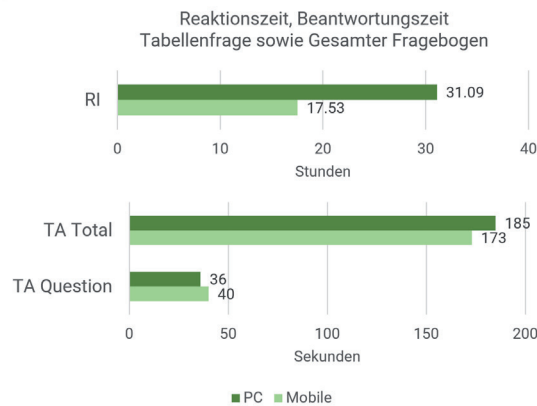


Bild 2: Reaktionszeit (RZ), Beantwortungszeit Tabellenfrage (BZ Frage) sowie Beantwortungszeit Gesamter Fragebogen (BZ Gesamt)

Brechen mobile Teilnehmende häufiger ab?

Weiter wurde die **Abschlussrate** berechnet. Sie bezeichnet die Anzahl der Teilnehmenden, die den Fragebogen bis zum Schluss ausgefüllt haben (letzte Frage wurde beantwortet).

Da das Ausfüllen der Umfrage über mobile Geräte weniger bequem sein kann, in "mobilen Situationen" oftmals vermehrt Ablenkungen auftreten und eine Umfrage auch nebenbei beantwortet werden könnte^{[1][3][8]}, sollte die Abbruchquote in der Mobil-Gruppe höher sein als in der PC-Gruppe (Hypothese 4).

Auf Mobilgeräten betrug mit einer Stichprobengrösse von $n_{mob}=24'613$ die Complete-Rate 82.3% ($n_{comp_mob}=20'259$), sowie mit $n_{pc}=21'968$ auf Nichtmobilgeräten 89.6% ($n_{comp_pc}=19'679$). Die Abweichung hat ein Odds Ratio^[16] von 0.5412 mit $p<0,001$ (entspricht Cohen's d von rund 0.33 als mittlerer Effekt^[2]), der Unterschied ist gemäss Fisher's Exact Test signifikant mit $p<0,001$. Die Abschlussrate ist also auf Mobilgeräten rund 7,3 Prozentpunkte tiefer. Umgekehrt betrachtet wurde eine Abbruchquote (Break-Off Rate) von mobil 17.7% gegenüber nicht mobil 10.4% beobachtet. Dies entspricht einer doch deutlichen Erhöhung um zwei Drittel (Hypothese 4 bestätigt).

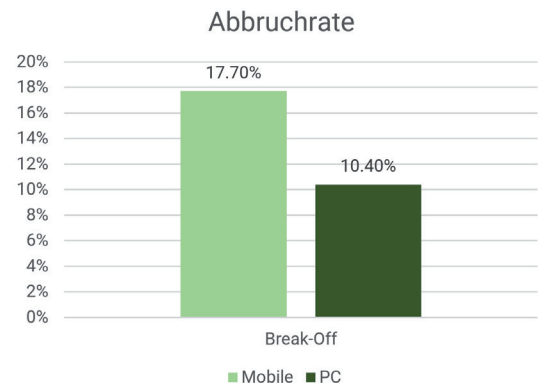


Bild 3: Abbruchrate Mobil vs. PC

Antworten mobile Teilnehmende weniger aufmerksam?

Der Begriff **Item Non-Response** beschreibt das Nichtbeantworten von Fragen, oder – bei Pflichtfragen, die bei Auslassen nochmals gestellt werden wie in unserem Fragebogen – die Auswahl einer Ausweichkategorie wie zum Beispiel "weiss nicht" oder "keine Antwort" bei Single- und Multiple-Choice-Fragen oder Tabellen mit Likert-Skalen. Untersucht haben wir dazu die im Fragebogen enthaltene Tabellenfrage mit 5 Items/Zeilen, die eine elfstufige Likert-Skala von höchst zufrieden bis höchst unzufrieden sowie die Ausweichkategorie "nicht beurteilbar" anbietet. Unsere Berechnung zeigt, wie oft die Ausweichkategorie ausgewählt wurde. Aus denselben Gründen wie bei Hypothese 4 vermuten wir, dass die Häufigkeit von Item Non-Response mobil höher ist als am PC (Hypothese 5). Limitierend für diese Studie ist anzumerken, dass Pflichtfragen eingesetzt wurden und daher das Kriterium "Item Non-Response" keine eigentlichen Nicht-Antworten erfasst, sondern lediglich die Nutzung der Ausweichkategorie, und diese zudem mit "nicht beurteilbar" statt typischerweise "weiss nicht/keine Antwort" beschriftet ist.

Auf Mobilgeräten betrug mit einer Gesamtantwortanzahl von $n_{mob}=107'131$ der Anteil an "nicht beurteilbar"-Antworten 3.39% ($n_{w_mob}=3'626$), auf Nichtmobilgeräten mit $n_{pc}=101'225$ Antworten 4.48% ($n_{w_pc}=4'535$). Diese Abweichung hat ein Odds Ratio^[16] von 1.3385 mit $p<0,001$ (entspricht Cohen's d von 0.16 als schwacher Effekt^[2]), der Unterschied ist gemäss Fisher's Exact Test signifikant mit

$p < 0,001$. Der Prozentsatz an "nicht beurteilbar"-Antworten ist entgegen unserer Vermutung auf Mobilgeräten damit um rund einen Viertel tiefer (absolut 1.09%). Die Hypothese 5 wird damit vorläufig abgelehnt.

Dies könnte auf eine höhere Datenqualität hinweisen, könnte aber auch daran liegen, dass die Ausweichkategorie mobil auf Grund des kleinen Bildschirms der Position ganz rechts als marginal wahrgenommen wird, oder dass Teilnehmende auf Mobilgeräten zu bequem sind, die vorhandene Ausweichkategorie überhaupt erst auszuwählen, und daher sogenannte Trash-Antworten hinterlegen und die Ausweichkategorie schlichtweg nicht akkurat benutzen. Zum Beispiel, indem eine Spalte mit immer gleichen Antworten ausgewählt wird ("herunterkreuzeln" ohne nachzudenken).

Um diese Art der Verschmutzung genauer zu untersuchen, wurden nachfolgend auch einige Typen von **Straightlining** untersucht. Es handelt sich dabei um ein Null-Varianz-Antwortverhalten, bei dem ein immer gleicher Skalenpunkt unabhängig von der Skalenbreite, -ausrichtung und Frageformulierung für alle Zeilen einer Skalentabelle gewählt wird, was häufig bei unmotivierten Teilnehmenden auftritt^[7]. Straightlining kann unter gewissen Umständen dennoch valide sein, zum Beispiel wenn eine Item-Batterie eine hohe interne Konsistenz aufweist und alle Items in dieselbe Richtung formuliert sind^[12]. Bei unserem Vergleich zwischen mobilen und nicht mobilen Teilnahmen ist dies besonders spannend, da in beiden Gruppen die Zahl der validen Straightliner konstant sein müsste (da sich die eigentliche Meinung, auch wenn sie über die 5 Items hinweg einheitlich ist, von mobil und nicht mobil teilnehmenden Personen bei so grossen Stichproben nicht unterscheiden dürfte) und lediglich die Zahl der auf Grund der Geräteverschiedenheit unterschiedlich agierenden Teilnehmenden, also die nicht validen Straightliner, variieren dürfte. Dieses sozusagen geräteinduzierte Straightlining stellt eine Datenverschmutzung dar.

Wir vermuten, dass mobile Teilnehmende häufiger Straightlining aufweisen als Teilnehmende am PC (Hypothese 6).

Auf Mobilgeräten (Stichprobengrösse von $n_{mob}=21'500$) betrug der Anteil an

Teilnehmenden mit Straightlining 35.47% ($n_{st_mob}=7627$), auf Nichtmobilgeräten (Stichprobengrösse mit $n_{pc}=20'252$) 31.19% ($n_{st_pc}=6316$), diese Abweichung hat ein Odds Ratio^[16] von 1.2131 mit $p < 0,001$ (entspricht Cohen's d von rund 0.10 als schwacher Effekt^[2]), der Unterschied ist gemäss Fisher's Exact Test mit $p < 0,001$ signifikant. Straightlining ist damit bei Mobilgeräten etwas häufiger problematisch, insbesondere, wenn man davon ausgeht, dass valides Straightlining (bewusstes, elaboriertes Entscheiden für immer dieselbe Skalausprägung in allen Zeilen der Tabellenfragen) bei beiden Vergleichsgruppen theoretisch gleich häufig sein müsste. Eine allfällig doch vorhandene Differenz müsste demnach ausschliesslich den Anteil "verschmutzter Daten" auf Grund von demotivational bedingtem Straightlining widerspiegeln. Damit dürfte der tatsächliche Effekt grösser sein als der gemessene Effekt. Beispielsweise würden nach Abzug von angenommen 30% validen Straightliner für die nicht validen Straightliner mobil 5.47% und nicht mobil 1.19% "übrig bleiben", also schon fast 5 mal mehr (Hypothese 6 bestätigt).

Der Begriff **Extreme Response Style (ERS)** bezeichnet ein spezifisches Antwortverhalten, bei dem in Tabellenfragen mit Likert-Skalen die Extrempunkte übermässig oder ausschliesslich genutzt werden. Wir klassifizieren für diese Studie Teilnehmende, die ausschliesslich Skalenendpunkte genutzt haben und mindestens eine Zeile mit einer diametral anders gepolten Antwort ausgewählt haben (z.B. 4 mal "höchst zufrieden" und 1 mal "höchst unzufrieden", 3 mal "höchst zufrieden" und 2 mal "höchst unzufrieden", 2 mal "höchst zufrieden" und 3 mal "höchst unzufrieden" etc.).

Wir vermuten, dass ESR auf Grund der Seltenheit des Phänomens bei mobilen Teilnahmen nicht signifikant häufiger auftritt als am PC (Hypothese 7).

Auf Mobilgeräten (Stichprobengrösse von $n_{mob}=21'500$) betrug der Anteil an Teilnehmenden mit ESR 0.35% ($n_{esr_mob}=75$), auf Nichtmobilgeräten (Stichprobengrösse mit $n_{pc}=20'252$) 0.38% ($n_{esr_pc}=77$), diese Abweichung hat ein Odds Ratio^[16] von 1.3993 mit $p=0,62$, der Unterschied ist gemäss Fisher's Exact Test nicht signifikant. Es gibt damit keinen signifikanten

(allenfalls nur zufälligen) Unterschied in der Häufigkeit von Extreme Response Style (ERS) zwischen Mobil- und Nichtmobilgeräten (Hypothese 7 bestätigt).

Sind mobile Textantworten kürzer?

Die **Länge der Textantworten auf offene Fragen** kann ebenfalls ein Qualitätskriterium sein, weil durch kürzere Antworten oft weniger substantielle oder ungenauere Aussagen für Auftraggebende herausgearbeitet werden können. Bei mobilen Teilnahmen gehen wir aufgrund der umständlicheren Eingabetechnologie, der eingeschränkten Platzverhältnisse und der mobilen Befragungssituation, die weniger elaborierte und zeitlich limitierte Reflexion begünstigen, von deutlich kürzeren Eingaben aus (Hypothese 8). Dies wurde bereits von Mavletova^[9] sowie Toepoel and Lugtig^[17] berichtet.

Das ausreisserbereinigte arithmetische Mittel der Textlängen grösser als 0 und kleiner als 500 Zeichen auf die Frage "Was (...) hat Sie am meisten gefreut (...) oder verärgert?" liegt bei den mobilen Teilnahmen bei 71.57 Zeichen ($M=71.57$; $SD=78.20$; $n_{mob}=16'156$), bei den nicht mobilen Geräten bei 100.95 Zeichen ($M=100.95$; $SD=97.34$; $n_{mob}=15'427$).

Die Abweichung ist gemäss Welch-Test $t(29572)=29.48$ signifikant mit $p<0,001$ und die Effektstärke mit Cohen's $d=0.33$ zeigt einen mittleren Effekt. Die Teilnehmenden, die nicht mobil geantwortet haben, hinterlegten also deutlich längere Texte (Hypothese 8 bestätigt).

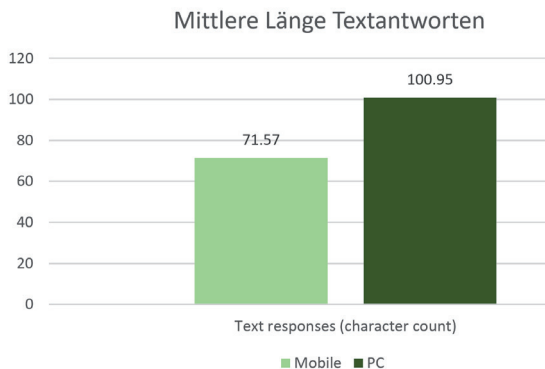


Bild 4: Arithmetisches Mittel (ausreisserbereinigt) Länge Textantworten Mobil vs. PC

	Mobil	Nichtmobil	Effekt
Reaktionszeit Umfrage-Einladung mobil kürzer	17.53 h	31.09 h	++
Abschlussrate mobil tiefer	82.3%	89.6%	++
Item nonresponse mobil tiefer	3.39%	4.48%	+
Straightliners mobil erhöht	35.47%	31.19%	+(+)
Extreme Response Style (ERS) keine Aussage	0.35%	0.38%	n.v.
Länge der Antworten auf offene Fragen mobil kürzer	72 Zeichen	101 Zeichen	++
Bearbeitungszeit für eine Frage mobil länger	40 s	36 s	+
Bearbeitungszeit ganze Umfrage mobil kürzer	173 s	185 s	(+)
Bearbeitungszeit ganze Umfrage mobil länger (Pausen exkludiert)	118 s	116 s	(+)

Tabelle 1: Zusammenfassung der Ergebnisse und Effektstärke

Fazit und Empfehlungen für die Praxis

Datenqualität in Onlineumfragen wurde im Vergleich zwischen mobilen und nicht mobilen Teilnahmen bisher wenig untersucht, wobei Ergebnisse vorangehender Studien mit einem erstmals sehr grossen Sample aus einer aktuellen Schweizer Kundenbefragung in einer realen für die Marktforschung relevanten Zielgruppe im Retail-Sektor weitgehend bestätigt werden. Mobile Teilnahmen zeichnen sich in der vorliegenden Untersuchung im Unterschied zu nicht mobilen Teilnahmen aus durch...

- eine **erhöhte Abbruchrate** (=tiefere Abschlussrate),
- eine **höhere Datenverschmutzung durch nicht valides Straightlining**,
- **deutlich kürzere offene Textantworten** mit möglicherweise weniger substantiellen oder detaillierten Aussagen,
- eine **schnellere Reaktionszeit** auf Umfrageeinladungen,
- etwas **weniger häufige Auswahl der Ausweichkategorie** "nicht beurteilbar" in

Likert-Skalen (was in Bezug auf Datenqualität unklar ist, da Teilnehmende unter Umständen mobil vorziehen, Trash-Antworten zu hinterlegen, anstatt akkurat die Ausweichkategorie zu benutzen, oder diese technisch auf Mobilgeräten zu wenig salient platziert ist),

- uneindeutige Ergebnisse zur Antwortdauer auf einzelne Fragen und des gesamten Fragebogens, wobei auf nicht mobilen Geräten mehr Pausen gemacht werden.

Grössere Unterschiede zwischen mobilen und nicht mobilen Teilnahmen bezüglich Datenqualität sind mit zunehmender Verbreitung technisch hochstehender Smartphones und Tablets und responsiven Fragebögen grundsätzlich nicht zu erwarten, da die noch vor wenigen Jahren beschriebenen Eingabehürden (schlechte Prozessorgeschwindigkeit, sehr kleine Screens, Usability, mangelhafte Netze^[11]) weitgehend aus dem Weg geräumt wurden.

Eine zentrale Ausnahme ist die zu erwartende geringere Textmenge und die damit möglicherweise weniger elaborierten und ausführlichen Angaben bei Fragen mit offenen Textfeldern, die gerade bei Fragebögen mit einer gewissen qualitativen Orientierung die Datenqualität ganz wesentlich beeinträchtigen können.

Eine in mobilen Teilnahmen erhöhte Zahl an Straightliner - wie in dieser Studie gezeigt - kann durch ausgeklügelte Methoden^{[7][10]} auch nachträglich in Survey Daten bereinigt werden.

Zusammenfassend empfehlen wir, obenstehende Implikationen für zukünftige geplante Umfragen regelmässig zu reflektieren sowie potenzielle Vor- und Nachteile transparent zu kommunizieren. Darüber hinaus empfehlen wir für alle Onlineumfragen intensives Pretesting in gemischten Zielgruppen (mobil und nicht mobil), eine entsprechende und gezielte Analyse der Pretest-Daten vor Feldstart auf für das Projekt wesentliche und in diesem Artikel beschriebene Parameter hin sowie das vorsorgliche Einbinden fachlicher Beratung durch Experten, gerade bei sensitiven Projekten.

Die Autoren

Raffael Meier, MA, MSc – Mitgründer/CTO von onlineumfragen.com und Pionier der deutschsprachigen Onlineumfragetechnologie. Er befasst sich mit gesellschaftlichen und methodologischen Aspekten von Daten und berät Kundinnen und Kunden mit dem Ziel «Empowerment».

Nina Gwerder, MA – Lead Consultant bei onlineumfragen.com und spezialisiert auf die Beratung namhafter nationaler und internationaler Unternehmen rund um das Thema Onlineumfragen und deren effektive Auswertung.

Kontakt

Raffael Meier, +41 44 500 5137
 raffael.meier@onlineumfragen.com

Nina Gwerder, +41 44 500 5140
 nina.gwerder@onlineumfragen.com

Dirty on the go?

Data quality in online surveys between PC and mobile devices



Raffael Meier
Co-founder and CTO
onlineumfragen.com

Our comparative analysis of anonymized mobile and non-mobile participation in a large Swiss customer survey from 2022 and 2023 reveals: Mobile responses fall slightly in terms of data quality. Straightliners and too short text responses pose a challenge.

Carrying the internet in your pocket - what was a utopian idea 25 years ago is trivial today. The introduction of the smartphone has changed our lives, including market and opinion research.



Nina Gwerder
Consultant
onlineumfragen.com

With increasingly better internet connections, larger screens and a faster-paced society with a permanent shortage of attention^[5], the need to be able to offer and complete surveys on mobile devices has risen sharply in the last decade. This increased demand for mobile or at least mobile-compatible surveys not only creates new challenges for the design and handling of online surveys, but must also be considered in a differentiated way in terms of data quality.

Using anonymized opinion data from a customer survey of a large Swiss retailer, launched in 2014 and receiving new responses every day, we analyze the development of mobile participation and its impact on data quality.

Data quality - what is it actually?

70 million search results on Google impressively demonstrate the relevance of the term "data quality". While in everyday life the term "data quality" is often equated with the correctness (precision) of data, data quality in market and opinion research is more complex. At the end of the data processing, the client or researcher is the consumer of the data. It is

therefore particularly important to understand that data quality from a consumer perspective often goes beyond the precision of values and also takes into account aspects such as credibility, relevance, completeness, interpretability, consistency and accessibility of data for consumers^[18]. Relevance, for example, involves ensuring that data is actually fit for its intended purpose, as poor data quality can lead to flawed decisions and inefficient business processes, while high data quality supports informed decision-making, better business processes and ultimately better business outcomes.

Data quality, which is described in the literature with numerous different models^[6], can be optimized through suitable measures. On the one hand, prior to the field phase, by means of elaborate sampling, stringent questionnaire design and pretesting. On the other hand, in the analysis and technical cleaning of the data already obtained. In addition, there are meta-aspects of data quality such as the fit of the data to the research question, process transparency, fair data collection or proactive data protection, which bring about empowerment for clients^[10].

Should we expect lower data quality in mobile surveys?

The specific effects of mobile survey participation on data quality can be divided into two categories:

- Effects that are due to the nature of the input device.
- Effects that are due to the situation in which mobile participation takes place.

The first category includes, for example, the smaller screen size on smartphones, which can influence the perception and understanding of the survey^[8]. For example, long item batteries with wide scales cannot always be presented clearly in the usual table form on a smartphone. They are therefore often presented in an alternative form, e.g. a division into several rating questions or drop-down fields. However, this influences the perception of the participants: While the individual item in a matrix question (likert item battery) is placed in the direct context of the items before and after it, the contents of the questions further back are often no longer present in several individual questions. At the same time, participants usually cannot already see the items that follow, which influences the cognitive editing behavior^{[14][15]}. In addition, surveys that are presented in a fundamentally different way on mobile devices than on large screens, as is common with some providers, are to be viewed methodologically sceptically, since all participants should receive the same stimuli as far as possible. Other aspects that can be assigned to this category include navigation via touchscreen and the possibly smaller computing power^[11].

The second category includes effects that are triggered by the possibility of participating "on the go". These include situational influences (e.g. interruptions because one has to change trains while answering on public transport or loud background noises influence concentration) as well as intrinsic influences, such as a generally lower level of attention when "scrolling" on the smartphone develops into a passive secondary activity^[19], for example when surveys are completed during other activities such as media consumption, meetings, lectures, at bus stops etc.

Research on data quality in mobile surveys is still in its early stages. Although various studies on this topic have been published since 2009^{[1][3][8][9][13]}, they are characterized by

partially contradictory results, small samples and partly clinical settings. The authors nevertheless agree on the following two points:

- The completion rate for mobile participation is lower than for participation via a PC / laptop (i.e. more break-offs)^{[1][3][8][9]}.
- The completion time for mobile participation is higher than for participation via a PC / laptop^{[3][8]}.

Case study: Customer survey of a large Swiss retailer

Our analysis is the first to cover a) a mixed, real target group (retail customers) with b) a large amount of data c) for Switzerland and thus clearly stands out from previous studies. Methodologically, it is based on the approach of Schlosser and Mays^[13], who examined the influence of mobile participation on data quality in a group of 820 German students in 2018.

Our analysis draws on fully anonymized data from the customer survey of a Swiss retailer, which was launched in 2014 in one division and subsequently successively expanded to other divisions. As of May 2023, 345,000 survey responses have been received in several languages (de, fr, it) throughout Switzerland. The questionnaire contains 6 questions. These include a question on the willingness to recommend, a five-item matrix table with an eleven-point Likert scale, a question with an open text response and two yes/no questions. The radio buttons and text boxes widely used on the Internet are used for answering the questions. The average completion time is about 2 minutes. No sociodemographic data is collected. The survey software used is the survey platform provided by onlineumfragen.com. The customers are invited in a mixed-mode procedure via e-mail (approx. 95%) and SMS (approx. 5%).

Development of the relevance of mobile participation 2014 to 2023

In our sample, the percentage of mobile participation, measured by browser agent strings and screen size, has risen steadily since 2014. Mobile engagement increased more than threefold in the period from 2014 to 2023.

A slowdown in the growth over the last five to six years could be related to a saturation of the target group: Those who potentially want to participate on mobile devices now also have the option.

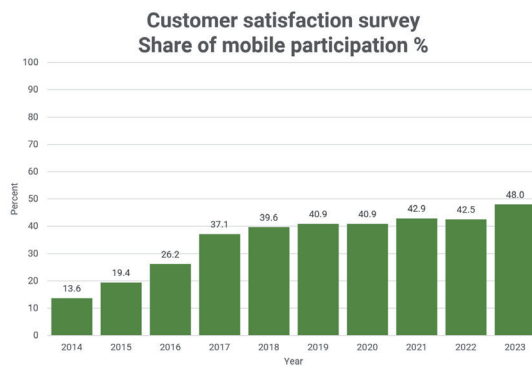


Figure 1: Share of mobile participation in customer satisfaction surveys in the retail sector, 2014 to 2023

Impact of mobile mode on data quality

In order to measure data quality in specific aspects, a subsample with a total of 46,581 current participations was used. The data was narrowed down according to participation dates younger than 01.01.2022. The following data quality criteria were examined:

- Response time to survey invitation
- Processing time for one question
- Processing time for entire survey
- Completion rate
- Item non-response
- Straightliner
- Extrem Response Style (ERS)
- Length of responses to open-ended questions

Do mobile participants need more time?

First, the **response time to the survey invitation** was calculated. It was determined by the difference between the time stamp of the invitation to the survey (time of sending the invitation

email) and the time stamp of the call to the survey (click on the survey link) and adjusted for outliers by only taking into account participations that took place within 7 days after the invitation.

As participants on mobile devices are used to having their mobile phone with them at all times and being reachable, we assume that these respondents react faster than respondents who receive the invitation via their PC (hypothesis 1).

On mobile devices the mean response time was 13 hours ($M=13.15$; $SD=25.43$; $n_{mob}=23'685$), on non-mobile devices 22.75 hours ($M=22.75$; $SD=34.26$; $n_{pc}=20'481$), the deviation is significant according to Welch's test $t(37137)=32.86$ with $p<0,001$ and the effect size with Cohen's $d=0.32$ corresponds to a medium effect.

If participation between 7 and 14 days after invitation was also taken into account, the mean response time on mobile devices was 17.5 hours ($M=17.53$; $SD=40.11$; $n_{mob}=24'169$), on non-mobile devices 31 hours ($M=31.09$; $SD=53.81$; $n_{pc}=21'335$), the deviation is significant according to Welch's test $t(38947)=30.08$ with $p<0,001$ and the effect size with Cohen's $d=0.29$ indicates a somewhat medium effect.

Participation by people who use mobile devices for the survey therefore takes place on average around 9.6 (or 13.6) hours earlier and thus significantly closer to the time of the invitation (hypothesis 1 confirmed). Depending on the survey, this could have an influence on topics such as memory strength, participation motivation or the emotionality of the feedback. The **mean response time for a question** was measured as the difference between the time of the completed display in the browser and the click on the button "Save - next question" for an item battery presented in table form consisting of 5 items with an eleven-point Likert scale on the satisfaction of certain aspects of the shopping experience. The question was presented methodologically identical for mobile and non-mobile and only optimized for mobile in terms of the aspect ratio of the table, the width and the font size in the sense of a responsive presentation.

For the mobile participations, we expect an

increased response time for more complex matrix questions^[4], as these are more difficult to grasp on the move, the mobile survey situation often makes concentrated answering more difficult, questions are somewhat less easy to read and the technical recording of the answers (touch screen) is somewhat more demanding and prone to errors (hypothesis 2).

The arithmetic mean of the processing time was adjusted for outliers (time > 10 seconds and < 180 seconds) and is 40 seconds ($M=40.09$; $SD=23.01$; $n_{mob}=20'333$) for mobile participants and 36 seconds ($M=35.70$; $SD=20.53$; $n_{pc}=19'402$) for non-mobile participants.

The difference is significant according to Welch's test $t(39555)=20.09$ with $p<0,001$ and the effect size with Cohen's $d=0.2$ points to a small but present effect. The processing time is slightly lower on non-mobile devices (hypothesis 2 confirmed).

The **processing time for the whole survey** was calculated using the difference between the time when the first question was displayed on the screen and the time when the last question was answered by clicking on the "Save" button. Because filling out a survey on mobile devices can be more tedious and more prone to interference, and the mobile survey situation is also usually exposed to more distractions, the total processing time could be higher on mobile devices. On the other hand, precisely these situational factors could weaken the motivation to fill out the survey elaborately "in peace" and accurately, and a faster, more superficial cognitive processing takes place, especially with questions that do not coincide with the individual content-related feedback core and are rather perceived as overhead or ballast. Therefore, in weighing up this consideration, we assume a somewhat shorter total processing time for mobile participation (hypothesis 3).

The arithmetic mean of the total processing time was adjusted for outliers (time > 10 seconds and < 300 seconds) and is 118 seconds ($M=118.17$; $SD=63.68$; $n_{mob}=17'435$) for mobile participants and 116 seconds ($M=115.73$; $SD=65.40$; $n_{pc}=16'915$) for non-mobile participants. The deviation is significant according

to Welch's test $t(34237)=3.50$ with $p<0,001$ and the effect size with Cohen's $d=0.038$ shows a very small effect. The processing time is thus minimally longer on mobile devices.

An exciting picture emerges with a less pronounced adjustment for outliers (time > 10 seconds and < 3600 seconds). In this case, the total processing time for mobile participants is 173 seconds ($M=172.50$; $SD=237.03$; $n_{mob}=19'619$), for non-mobile participants 185 seconds ($M=184.69$; $SD=256.90$; $n_{pc}=19'518$). This deviation is significant according to Welch's test $t(38851)=4.88$ with $p<0,001$ and the effect size with Cohen's $d=0.05$ shows a small effect, but now the PC participations last longer.

This could be associated with the fact that there are a higher number of participants on PCs who interrupt the survey and continue after a break (e.g. lunch hour, telephone call, etc.). This rapidly results in very long (but not continuously active) processing times. If total processing times of more than 5 minutes are excluded, mobile participations are therefore slightly slower and have fewer breaks (hypothesis 3 partially rejected). More research would have to be done to find out why PC participation takes longer when total processing times of up to one hour are included. This could also be related to a higher proportion of older and less IT-savvy people in the PC group, as well as to the fact that text responses are generally longer on the PC - however, these are often typed more quickly (see hypothesis 8).

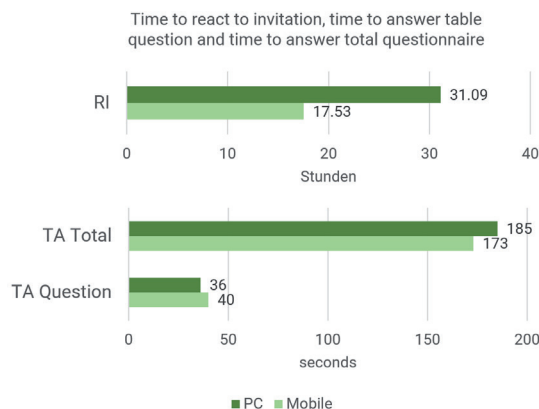


Figure 2 Time to react to invitation (RI), time to answer table question (TA question) and time to answer total questionnaire (TA total):

Do mobile participants break off more often?

The completion rate has also been calculated. This refers to the number of participants who completed the questionnaire to the end (last question was answered).

Since completing the survey on mobile devices can be less convenient, distractions often increase in "mobile situations" and a survey could also be answered in parallel^{[1][3][8]}, the break-off rate should be higher in the mobile group than in the PC group (hypothesis 4).

On mobile devices, with a sample size of $n_{mob}=24'613$, the complete rate was 82.3% ($n_{comp_mob}=20'259$), and with $n_{pc}=21'968$ on non-mobile devices it was 89.6% ($n_{comp_pc}=19'679$). The difference has an odds ratio^[16] of 0.5412 with $p<0,001$ (corresponds to Cohen's d of around 0.33 as a medium effect^[2]), the difference is significant according to Fisher's Exact Test with $p<0,001$. The completion rate is thus around 7.3 percentage points lower on mobile devices. Conversely, a break-off rate of 17.7% was observed on mobile devices compared to 10.4% on non-mobile devices. This corresponds to a significant increase of two thirds (hypothesis 4 confirmed).

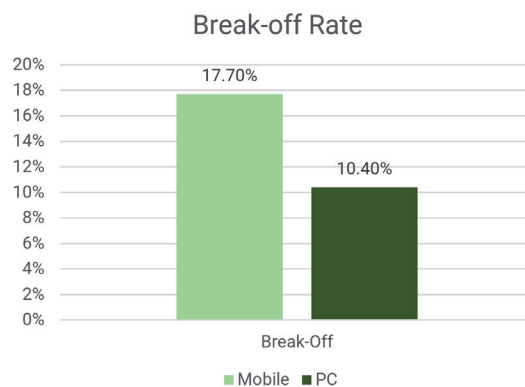


Bild 3: Break-off rate mobile vs. PC

Do mobile participants respond less attentively?

The term **item non-response** describes the non-answering of questions or - in the case of compulsory questions which are asked repeatedly if omitted, as in our questionnaire - the selection of an evasive category such as "don't

know" or "no answer" in the case of single and multiple-choice questions or tables with Likert scales. For this purpose, we examined the table question with 5 items/lines included in the questionnaire, which offers an eleven-point Likert scale from highly satisfied to highly dissatisfied as well as the alternative category "not assessable". Our analysis shows how often the alternative category was chosen.

For the same reasons as for hypothesis 4, we assume that the frequency of item non-response is higher on mobile devices than on PCs (hypothesis 5). A limiting factor for this study is that compulsory questions were used and therefore the criterion "item non-response" does not record actual non-answers, but only the use of the alternative category, which is furthermore labelled "not assessable" instead of the typical "don't know/no answer".

On mobile devices, with a total number of responses of $n_{mob}=107'131$ the proportion of "not assessable" responses was 3.39% ($n_{w_mob}=3'626$), on non-mobile devices, with $n_{pc}=101'225$ responses, the proportion was 4.48% ($n_{w_pc}=4'535$).

This deviation has an odds ratio^[16] von 1.3385 with $p<0,001$ (corresponds to Cohen's d of 0.16 as a small effect^[2]), the difference is significant according to Fisher's Exact Test with $p<0,001$. Contrary to our assumption, the percentage of "not assessable" responses is approximately a quarter lower on mobile devices (1.09% in absolute terms). Hypothesis 5 is therefore rejected for the moment.

This could indicate a higher data quality, but could also be due to the fact that the fallback category is perceived as marginal on mobile devices due to the small screen and the position on the far right, or that participants on mobile devices are too comfortable to select the available fallback category in the first place and hence deposit so-called trash answers and simply do not use the fallback category accurately. For example, by selecting a column with repeatedly the same answers ("ticking down" without reflection). In order to examine this type of pollution

in more detail, some types of straightliners were also examined below. **Straightlining** is a zero-variance response behaviour in which an always identical scale point is chosen for all rows of a scale table, regardless of scale width, scale orientation and question wording, which often occurs with unmotivated participants^[7].

Straightlining can nevertheless be valid under certain circumstances, for example if an item battery has a high internal consistency and all items are formulated in the same direction^[12]. This is particularly exciting in our comparison between mobile and non-mobile participation, since in both groups the number of valid straightliners should be constant (because the actual opinion, even if it is uniform across the 5 items, should not differ between mobile and non-mobile participants in such large samples) and only the number of participants who act differently due to the device difference, i.e. the non-valid straightliners, should vary. This device-induced straightlining, so to speak, represents data pollution. We assume that mobile participants are more likely to exhibit straightlining than participants using a PC (hypothesis 6).

On mobile devices (sample size of $n_{mob}=21'500$) the proportion of participants with straightlining was 35.47% ($n_{st,mob}=7627$), on non-mobile devices (sample size with $n_{pc}=20'252$) 31.19% ($n_{st,pc}=6316$), this difference has an odds ratio^[16] of 1.2131 with $p<0,001$ (corresponds to Cohen's d of 0.10 as a weak effect^[2]), the difference is significant according to Fisher's Exact Test with $p<0.001$.

Straightlining is therefore a somewhat more frequent problem with mobile devices, especially if one assumes that valid straightlining (conscious, elaborate decisions for the same scale value in all rows of the table questions) should theoretically be equally frequent in both comparison groups. Any difference that exists should therefore exclusively reflect the proportion of "polluted data" due to demotivational straightlining. Thus, the actual effect should be larger than

the measured effect. For example, after deducting an assumed (purely hypothetical) 30% of valid straightliners, 5.47% of the non-valid straightliners would be mobile and 1.19% non-mobile, i.e. nearly 5 times more (hypothesis 6 confirmed).

The term **Extreme Response Style (ERS)** describes a specific response behavior in which the extreme points are used excessively or exclusively in table questions with Likert scales. For this study, we classify participants who exclusively used scale endpoints and selected at least one row with a diametrically different response (e.g. 4 times "highly satisfied" and 1 time "highly dissatisfied", 3 times "highly satisfied" and 2 times "highly dissatisfied", 2 times "highly satisfied" and 3 times "highly dissatisfied" etc.).

We assume that ESR does not occur significantly more often in mobile participation than on the PC due to the rarity of the phenomenon (hypothesis 7).

On mobile devices (sample size of $n_{mob}=21'500$) the proportion of participants with ESR was 0.35% ($n_{esr,mob}=75$), on non-mobile devices (sample size with $n_{pc}=20'252$) 0.38% ($n_{esr,pc}=77$), this difference has an odds ratio^[16] von 1.3993 with $p=0,62$, the difference is not significant according to Fisher's Exact Test. There is thus no significant (at best only random) difference in the frequency of Extreme Response Style (ERS) between mobile and non-mobile devices (hypothesis 7 confirmed).

Are mobile text responses shorter?

The length of text answers to open questions can also be a quality criterion, because shorter answers can often produce less substantive or more inaccurate statements for clients.

For mobile participation, we assume significantly shorter statements due to the more inconvenient input technology, the limited space and the mobile interview situation, which favor less elaborate and time-limited reflection (hypothesis 8). This has already been reported by

Mavletova^[9] and Toepoel and Lugtig^[17].

The outlier-adjusted arithmetic mean of the text lengths greater than 0 and less than 500 characters to the question "What (...) made you most happy (...) or angry?" is 71.57 characters for mobile participants ($M=71.57$; $SD=78.20$; $n_{mob}=16'156$), and 100.95 characters for non-mobile participants ($M=100.95$; $SD=97.34$; $n_{mob}=15'427$).

The deviation is significant according to Welch's test $t(29572)=29.48$ with $p<0,001$ and the effect size with Cohen's $d=0.33$ indicates a medium effect. The participants who did not answer by mobile phone therefore deposited significantly longer texts (hypothesis 8 confirmed).

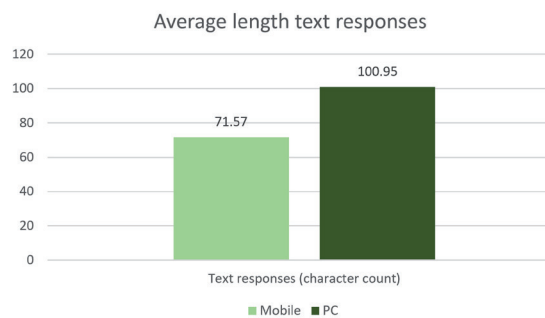


Figure 4: Arithmetic mean (adjusted for outliers) length of text responses mobile vs. PC

	Mobil	Non-mobile	Effect
Response time Survey invitation mobile shorter	17.53 h	31.09 h	++
Completion rate mobile lower	82.3%	89.6%	++
Item nonresponse mobile lower	3.39%	4.48%	+
Straightliners mobile increased	35.47%	31.19%	+(+)
Extreme Response Style (ERS) no statement	0.35%	0.38%	n.s.
Length of answers to open questions mobile shorter	72 Zeichen	101 Zeichen	++
Processing time for one question mobile longer	40 s	36 s	+
Processing time for whole survey mobile shorter	173 s	185 s	(+)
Processing time for whole survey mobile longer (breaks excluded)	118 s	116 s	(+)

Table 1: Summary of results and effect size

Conclusion and recommendations for practice

Data quality in online surveys has been rarely investigated in comparison between mobile and non-mobile participation, whereas results of previous studies are confirmed to a considerable extent for the first time with a very large sample from a current Swiss customer survey in a real target group in the retail sector that is relevant for market research.

In contrast to non-mobile participation, mobile participation in this study is characterized by...

- an **increased break-off rate** (=lower completion rate),
- higher data pollution due to **non-valid straightlining**,
- significantly **shorter open text responses** with possibly less substantive or detailed statements,
- a **faster response time to survey invitations**,
- slightly **less frequent selection of the fallback category "not assessable"** in Likert scales (which is unclear in terms of data quality, as participants may prefer to leave trash answers on mobile rather than accurately using the fallback category, or it may not be technically placed salient enough on mobile devices),
- ambiguous results on the duration of answers to individual questions and the entire questionnaire, with more pauses being made on non-mobile devices.

Major differences between mobile and non-mobile participation regarding data quality are basically not to be expected with the increasing spread of technically advanced smartphones and tablets and responsive questionnaires, as the input hurdles described just a few years ago (poor processor speed, very small screens, usability, poor networks^[11]) have largely been overcome.

A key exception is the expected smaller amount of written text and the resulting possibly less elaborate and detailed information in questions with open text fields, which can significantly

affect data quality, especially in questionnaires with a certain qualitative orientation.

An increased number of straightliners in mobile participations - as shown in this study - can also be subsequently remedied in survey data by sophisticated methods^{[7][10]}.

In summary, we advise to regularly reflect on the above implications for future planned surveys and to transparently communicate potential advantages and disadvantages. Furthermore, we recommend intensive pretesting in mixed target groups (mobile and non-mobile) for all online surveys, a corresponding and systematic analysis of the pretest data before the field start with regard to parameters that are essential for the project and described in this article, as well as the precautionary inclusion of specialist advice from experts, particularly for sensitive projects.

The authors

Raffael Meier, MA, MSc - Co-founder/CTO of onlineumfragen.com and pioneer of online survey technology and consulting in Germany, Switzerland and Austria. He engages with societal and methodological aspects of data and advises clients with the goal of "empowerment".

Nina Gwerder, MA - Lead Consultant at onlineumfragen.com and specialist in advising well-known national and international companies on all aspects of online surveys and their in-depth analysis.

Kontakt

Raffael Meier, +41 44 500 5137
raffael.meier@onlineumfragen.com

Nina Gwerder, +41 44 500 5140
nina.gwerder@onlineumfragen.com

Literatur

- [1] Bosnjak, M., Poggio, T., Becker, K.R., Funke, F., Wachenfeld, A., & Fischer, B. (2013). *Online survey participation via mobile devices*. Conference Paper, The American Association for Public Opinion Research (AAPOR) 68th Annual Conference, 2013, Boston, MA.
- [2] Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Converting Among Effect Sizes*. In *Introduction to Meta-Analysis* (eds M. Borenstein, L.V. Hedges, J.P.T. Higgins and H.R. Rothstein).
- [3] de Brujine, M., & Wijbant, A. (2013) Comparing survey results obtained via mobile devices and computers: An experiment with a mobile web survey on a heterogeneous group of mobile devices versus a computer-assisted web survey. *Social Science Computer Review*, 31, S.482-504.
- [4] Couper, M. P., & Peterson, G. (2016). Why do web surveys take longer on smartphones? *Social Science Computer Review*. First published on February 11, 2016. doi:10.1177/0894439316629932
- [5] Franck, G. (1998). *Ökonomie der Aufmerksamkeit*. Ein Entwurf. Dtv, München 2007, ISBN 978-3-423-34401-2.
- [6] Haug, A. (2021). Understanding the differences across data quality classifications: a literature review and guidelines for future research. *Industrial Management and Data Systems*, 121(12), 2651–2671. <https://doi.org/10.1108/IMDS-12-2020-0756>
- [7] Jandura, O. (2018). Fake Data? Zur Trennung von sauberen und verschmutzten Daten bei selbstadministrierten Befragungsmodi. In *Rössler P. & Rossman, C. (Hrsg.), Kumulierte Evidenzen*. Wiesbaden: Springer VS, S. 207-223.
- [8] Mavletova, A. (2013). Data Quality in PC and mobile web surveys. *Social Science Computer Review*, 31, S.725-743.
- [9] Mavletova, A., & Couper, M. P. (2015). A meta-analysis of breakoff rates in mobile web surveys. In *D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), Mobile research methods: Opportunities and challenges of mobile research methodologies* (pp. 81–98). London, England: Ubiquity Press.
- [10] Meier, R., & Gwerder, N. (2022). Dirty Data in Online Surveys. Wie Datenqualität vor und nach der Feldphase verbessert werden kann. In Swiss Data Insights Association (Hrsg.), *Swiss Insights Report 2022*. Swiss Insights. <https://swiss-insights.ch/wp-content/uploads/2023/01/Annual-Report-2022-klein.pdf>
- [11] Peytchev, A., & Hill, C.A. (2010). Experiments in mobile web survey design. Similarities to other modes and unique considerations. *Social Science Computer Review*, 28, S.319-355.
- [12] Reunig, K., & Plutzer E. (2020). Valid vs. Invalid Straightlining: The Complex Relationship Between Straightlining and Data Quality. *Survey Research Methods*, 14(5), S.439-459
- [13] Schlosser, S., & Mays, A. (2018). Mobile and Dirty: Does Using Mobile Devices Affect the Data Quality and the Response Process of Online Surveys?, *Social Science Computer Review*, 36(2), S.212-230. DOI: 10.1177/0894439317698437
- [14] Schwarz, N. (2007). Cognitive aspects of survey methodology. *Appl. Cognit. Psychol.*, 21: 277-287. <https://doi.org/10.1002/acp.1340>
- [15] Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93–105. <https://doi.org/10.1037/0003-066X.54.2.93>
- [16] Sheskin, DJ (2004). *Handbook of parametric and nonparametric statistical procedures*. 3rd ed. Boca Raton: Chapman & Hall /CRC.
- [17] Toepoel, V., & Lugtig, P. (2014). What happens if you offer a mobile option to your web panel? Evidence from a probability-based panel of internet users. *Social Science Computer Review*, 32, 544–560.
- [18] Wang, R.Y., & Strong, D.M. (1996). Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems*, 12(4), S.5-33.
- [19] Wells, T., Bailey, J.T., & Link, M.W. (2014). Comparison of smartphone and online computer survey administration. *Social Science Computer Review*, 32, S.238-255.

SWISS INSIGHTS

Corporate Member

About You Services SA

passage du Cardinal 14, 1700 Fribourg
+41 79 301 11 37 / info@about-you.app
about-you.app

amPuls Market Research

Hirschengraben 49, 6000 Luzern 7
+41 41 612 14 14 / info@ampuls.ch
www.ampuls.ch

amrein+heller MarktforschungTreuhand AG

Südweid 7, 6274 Eschenbach
+41 748 63 70 / contact@ah-feedback.ch
www.ah-feedback.ch

Bilendi Schweiz AG

Reinhardstrasse 19, 8008 Zürich
+41 79 801 88 80 / contact.ch@bilendi.com
www.bilendi.ch

Boomerang Ideas AG

Sihlquai 131, 8005 Zürich
+41 44 500 88 60 / raphael@boomerangideas.com
www.boomerangideas.com

BSI Business Systems Integration AG

Täferweg 1, 5405 Baden
+41 58 255 90 00, info@bsi-software.com
www.bsi-software.com

Constant Dialog

Alte Steinhäuserstrasse 33, 6330 Cham
+41 41 310 05 40 / info@constant-dialog.ch
www.constant-dialog.ch

DemoSCOPE Data + Research

Klusenstrasse 17, 6043 Adligenswil
+41 41 375 40 00 / demoscope@demoscope.ch
www.demoscope.ch

Gallup AG

Reinhardstrasse 19, 8008 Zürich,
+41 78 891 31 15 / office@gallup.swiss
www.gallup.swiss

gff Swiss Research Services

Baarerstrasse 25, 6300 Zug
+41 41 560 01 60 / gut@gff.ag
www.gff.ag

GfK Switzerland AG

Suurstoffi 18A, 6343 Rotkreuz
+41 41 632 91 11 / info.ch@gfk.com
www.gfk.ch / www.gfk.com

gfs.bern. Menschen. Meinungen. Märkte.

Effingerstrasse 14, Postfach, 3001 Bern
+41 31 311 08 06 / info@gfsbern.ch
www.gfsbern.ch

gfs-zürich, Markt- & Sozialforschung

Riedtlistrasse 9, 8006 Zürich
+41 44 360 40 20 / gfs@gfs-zh.ch
www.gfs-zh.ch

gfs-befragungsdienst

Schaffhauserstrasse 491, 8052 Zürich
+41 44 360 26 40 / info@gfs-bd.ch
www.gfs-bd.ch

GIM Suisse AG

General-Wille-Strasse 10, 8002 Zürich
+41 44 283 18 18 / info@g-i-m.ch
www.g-i-m.ch

Happy Thinking People AG

Staufacherstrasse 101, 8048 Zürich
+41 44 204 16 26 / contact-zurich@happythinkingpeople.com
www.happythinkingpeople.com

INNOFACT (Schweiz) AG Research & Consulting

Flurstrasse 50, 8048 Zürich
+41 43 931 77 82, Info@innofact.ch
www.innofact.ch

Instight Institute AG

Bergstrasse 138, 8032 Zürich
+41 44 387 90 90 / info@insightinstitute.ch
www.insightinstitute.ch

intervista

Optingenstrasse 5, 3013 Bern
+41 31 511 39 00 / anfragen@intervista.ch
www.intervista.ch

IPSOS Suisse SA

11, Chemin du Château-Bloch, 1219 Le Lignon
+41 22 591 06 00 / Contact_Switzerland@ipsos.com
www.ipsos.com/de-ch

SWISS INSIGHTS

Corporate Member

just-medical!

Blegistrasse 5, 6340 Baar
+41 41 766 11 55 / info@just-medical.com
www.pharmaagentur.ch



SensoPLUS

Industriestrasse 16, 6300 Zug
+41 41 710 71 61 / info@sensoplus.ch
www.sensoplus.ch



Kantar Media Switzerland AG

Bahnhofstrasse 4, 3073 Gümliigen
+41 31 537 79 00 / ch.panel@kantarmedia.com
www.kantarmedia.com



TALK Online Panel

Lindenmoosstrasse 4, 8910 Zürich
+41 43 550 14 62 / rfq@talkonlinepanel.com
www.talk-group.com



LINK

Baslerstrasse 60, 8048 Zürich
+41 41 367 73 73 / zurich@link.ch
www.link.ch



TransferPlus AG Market Research

Haldenstrasse 11, 6006 Luzern
+41 41 618 33 11 / transfer@transferplus.ch
www.transferplus.ch



Marketagent.com Schweiz AG

Seefeldstrasse 19, 8008 Zürich
+41 43 555 06 50 / schweiz@marketagent.com
www.marketagent.com



M.I.S. Trend SA

Pont Bessières 3, 1005 Lausanne
+41 21 320 95 03 / info@mistrend.ch
www.mistrend.ch



NielsenIQ (Switzerland) GmbH

Park 4, 6039 Root D4
+41 41 445 64 64 / nielsen-ch@nielsen.com
www.nielsen.com



onlineumfragen.com

Kernserstrasse 15, 6056 Kägiswil
+41 44 500 50 54 / info@onlineumfragen.com
www.onlineumfragen.com



POLYQUEST AG

Flurstrasse 26, 3014 Bern
+41 31 335 64 00 / info@polyquest.ch
www.polyquest.ch



Publicom AG

Alte Landstrasse 55, 8802 Kilchberg
+41 44 716 55 11 / publicom@publicom.ch
www.publicom.ch



Qualitest AG, Institut für Marketing- und Sozialforschung

Rosenberghöhe 3, 6004 Luzern
+41 41 712 12 21 / qualitest@qualitestag.ch
www.qualitestag.ch





Swiss Insights

Swiss Insights ist der Verband und die Interessensvertretung aller Unternehmen, die Daten und prädiktive Modelle im Rahmen von Marketing, Innovationsprozessen, Kundenservice, Angebotsgestaltung, Kommunikation und Zielgruppendefinitionen erheben, analysieren, einsetzen und daraus Handlungsempfehlungen ableiten.

Die Corporate Mitglieder von Swiss Insights sind Unternehmen, die sich für den fairen Umgang mit Auskunftspersonen sowie Auftraggeber und den Schutz der Privatsphäre engagieren. Corporate Mitglieder sind berechtigt, das Logo "SWISS INSIGHTS - Corporate Member" zu tragen. Corporate Member haben zudem die Möglichkeit, das Label "Market and Social Research by Swiss Insights" und/oder das "Label Data Fairness by Swiss Insights" zu beantragen. Unternehmen, die eines der beiden Label tragen, garantieren, dass sie sich an die strengen Richtlinien und Reglemente von Swiss Insights und Esomar halten.

Das Label "Market and Social Research by Swiss Insights" kann von Unternehmen mit Sitz/Niederlassung in der Schweiz beantragt werden. Sie unterliegen einem strengen Regelwerk von schweizerischen und internationalen Normen und garantieren, dass keine Interviews mit Werbe- oder Verkaufsabsichten durchgeführt, wissenschaftlich abgesicherte Methoden angewendet und die Datenschutzrichtlinien eingehalten werden.

Das Label "Data Fairness by Swiss Insights" kann sowohl von Schweizer wie auch von ausländischen Unternehmen beantragt werden. Mit dem Label zeigen sie, dass sich ihr Unternehmen oder ihre Abteilung für den transparenten und verantwortungsvollen Umgang bei der Bearbeitung und Analyse von grossen Datenmengen verpflichtet.



Herausgeber und Kontakt

Swiss Insights, Swiss Data Insights Association, Gruebengasse 10, 6055 Alpnach, Switzerland
+41 44 3501960, info@swiss-insights.ch, www.swiss-insights.ch

SWISS INSIGHTS

Corporate Member

About You Services SA

passage du Cardinal 14, 1700 Fribourg
+41 79 301 11 37 / info@about-you.app
about-you.app

amPuls Market Research

Hirschengraben 49, 6000 Luzern 7
+41 41 612 14 14 / info@ampuls.ch
www.ampuls.ch

amrein+heller MarktforschungsTreuhand AG

Südweid 7, 6274 Eschenbach
+41 748 63 70 / contact@ah-feedback.ch
www.ah-feedback.ch

Bilendi Schweiz AG

Reinhardstrasse 19, 8008 Zürich
+41 79 801 88 80 / contact.ch@bilendi.com
www.bilendi.ch

Boomerang Ideas AG

Sihlquai 131, 8005 Zürich
+41 44 500 88 60 / raphael@boomerangideas.com
www.boomerangideas.com

BSI Business Systems Integration AG

Täferweg 1, 5405 Baden
+41 58 255 90 00, info@bsi-software.com
www.bsi-software.com

Constant Dialog

Alte Steinhäuserstrasse 33, 6330 Cham
+41 41 310 05 40 / info@constant-dialog.ch
www.constant-dialog.ch

DemoSCOPE Data + Research

Klusenstrasse 17, 6043 Adligenswil
+41 41 375 40 00 / demoscope@demoscope.ch
www.demoscope.ch

Gallup AG

Reinhardstrasse 19, 8008 Zürich,
+41 78 891 31 15 / office@gallup.swiss
www.gallup.swiss

gff Swiss Research Services

Baarerstrasse 25, 6300 Zug
+41 41 560 01 60 / gut@gff.ag
www.gff.ag

GfK Switzerland AG

Suurstoffi 18A, 6343 Rotkreuz
+41 41 632 91 11 / info.ch@gfk.com
www.gfk.ch / www.gfk.com

gfs.bern. Menschen. Meinungen. Märkte.

Effingerstrasse 14, Postfach, 3001 Bern
+41 31 311 08 06 / info@gfsbern.ch
www.gfsbern.ch

gfs-zürich, Markt- & Sozialforschung

Riedtlistrasse 9, 8006 Zürich
+41 44 360 40 20 / gfs@gfs-zh.ch
www.gfs-zh.ch

gfs-befragungsdienst

Schaffhauserstrasse 491, 8052 Zürich
+41 44 360 26 40 / info@gfs-bd.ch
www.gfs-bd.ch

GIM Suisse AG

General-Wille-Strasse 10, 8002 Zürich
+41 44 283 18 18 / info@g-i-m.ch
www.g-i-m.ch

Happy Thinking People AG

Staufacherstrasse 101, 8048 Zürich
+41 44 204 16 26 / contact-zurich@happythinkingpeople.com
www.happythinkingpeople.com

INNOFACT (Schweiz) AG Research & Consulting

Flurstrasse 50, 8048 Zürich
+41 43 931 77 82, Info@innofact.ch
www.innofact.ch

Instight Institute AG

Bergstrasse 138, 8032 Zürich
+41 44 387 90 90 / info@insightinstitute.ch
www.insightinstitute.ch

intervista

Optingenstrasse 5, 3013 Bern
+41 31 511 39 00 / anfragen@intervista.ch
www.intervista.ch

IPSOS Suisse SA

11, Chemin du Château-Bloch, 1219 Le Lignon
+41 22 591 06 00 / Contact_Switzerland@ipsos.com
www.ipsos.com/de-ch

SWISS INSIGHTS

Corporate Member

just-medical!

Blegistrasse 5, 6340 Baar
+41 41 766 11 55 / info@just-medical.com
www.pharmaagentur.ch



Qualitest AG, Institut für Marketing- und Sozialforschung

Rosenberghöhe 3, 6004 Luzern
+41 41 712 12 21 / qualitest@qualitestag.ch
www.qualitestag.ch



Kantar Media Switzerland AG

Bahnhofstrasse 4, 3073 Gümligen
+41 31 537 79 00 / ch.panel@kantarmedia.com
www.kantarmedia.com



SensoPLUS

Industriestrasse 16, 6300 Zug
+41 41 710 71 61 / info@sensoplus.ch
www.sensoplus.ch



LINK

Baslerstrasse 60, 8048 Zürich
+41 41 367 73 73 / zurich@link.ch
www.link.ch



TALK Online Panel

Lindenmoosstrasse 4, 8910 Zürich
+41 43 550 14 62 / rfq@talkonlinepanel.com
www.talk-group.com



Marketagent.com Schweiz AG

Seefeldstrasse 19, 8008 Zürich
+41 43 555 06 50 / schweiz@marketagent.com
www.marketagent.com



TransferPlus AG Market Research

Haldenstrasse 11, 6006 Luzern
+41 41 618 33 11 / transfer@transferplus.ch
www.transferplus.ch



M.I.S. Trend SA

Pont Bessières 3, 1005 Lausanne
+41 21 320 95 03 / info@mistrend.ch
www.mistrend.ch



NielsenIQ (Switzerland) GmbH

Park 4, 6039 Root D4
+41 41 445 64 64 / nielsen-ch@nielsen.com
www.nielsen.com



onlineumfragen.com

Kernserstrasse 15, 6056 Kägiswil
+41 44 500 50 54 / info@onlineumfragen.com
www.onlineumfragen.com



POLYQUEST AG

Flurstrasse 26, 3014 Bern
+41 31 335 64 00 / info@polyquest.ch
www.polyquest.ch



Publicom AG

Alte Landstrasse 55, 8802 Kilchberg
+41 44 716 55 11 / publicom@publicom.ch
www.publicom.ch



QUALINSIGHT Sàrl

Route des Mosses 34, 1613 Maracon
+41 79 911 62 19 / e.seve@qualinsight.ch
www.qualinsight.ch

