

Wissenschaftliches Gutachten von Katharina Schüller:
Qualitäts-Anforderungen an Studien zur Ableitung von Regulierungsmaßnahmen:
Kritische Evaluation des Glücksspiel-Surveys 2021

Zusammenfassung des Gutachtens seitens der Auftraggeber:
Bundesverband deutscher Spielbanken gegr. 2008 als BupriS e.V. (BupriS),
Deutscher Online Casinoverband e.V. (DOCV),
Deutscher Sportwettenverband e.V. (DSWV),
Die Deutsche Automatenwirtschaft e.V. (DAW)

I. Hintergrund

Im März 2022 haben Sven Buth, Gerhard Meyer und Jens Kalke ihren Forschungsbericht „Glücksspielteilnahme und glücksspielbezogene Probleme in der Bevölkerung – Ergebnisse des Glücksspiel-Surveys 2021“ vorgelegt. Es handelt sich um den ersten von drei Surveys, die im Abstand von jeweils zwei Jahren durchgeführt werden sollen. Projektträger sind das ISD Institut für interdisziplinäre Sucht- und Drogenforschung in Hamburg und die Arbeitseinheit Glücksspielforschung der Universität Bremen; gefördert wird das Projekt vom Deutschen Lotto- und Totoblock (DLTB).

Ziel der Surveys ist die Erhebung einer aktuellen validen Zahlenbasis über das Glücksspielverhalten in Deutschland, über Art und Ausmaß von problematischem Glücksspielverhalten und über Zusammenhänge u.a. zwischen Glücksspiel und Alkoholkonsum, allgemeinen psychischen Belastungen sowie Werbung für Glücksspiel. Tatsächlich haben der Survey und die Aufbereitung seiner Ergebnisse durch die Autoren im großen Umfang Eingang in aktuelle wissenschaftliche, politische und gesellschaftliche Debatten über das Glücksspiel in Deutschland gefunden. Insbesondere die kontraintuitive Ergebnisinterpretation der Survey-Autoren, wonach 8 % der 18-bis-70-jährigen Bevölkerung „problematisches Spielverhalten“ aufweisen sollen (vgl. S. 6 des Surveys; die letzte Erhebung der BZgA 2019 erhob noch einen Bevölkerungsanteil von Problem Spielern von lediglich 0,73 %), wurde politisch und medial breit rezipiert; etwa der Beauftragte für Sucht- und Drogenfragen der Bundesregierung und der Senator für Inneres der Freien Hansestadt Bremen leiteten hieraus Forderungen nach einer restriktiveren Glücksspielregulierung ab.

Ein Peer-Review des Surveys, d.h. eine Begutachtung der Qualität der Arbeit durch unabhängige Wissenschaftler, hat bis heute ebenso wenig stattgefunden wie eine kritische Diskussion in der Fachliteratur. Beides gehört zu den Selbstverständlichkeiten des transparenten Wissenschaftsbetriebs, zumal wenn Forschungsergebnisse Einfluss auf politische Entscheidungen nehmen. Gerade dann sind besonders hohe Qualitätsmaßstäbe anzulegen. Die Glücksspielanbieter und ihre Branchenverbände veranlasste dieses Fehlen einer wissenschaftlichen Qualitätskontrolle zur Beauftragung der vorliegenden

Begutachtung des Surveys durch die Statistikerin Katharina Schüller, Vorstandsmitglied der Deutschen Statistischen Gesellschaft.

II. Ergebnisse

Das Fazit der Begutachtung ist eindeutig: Der Glücksspiel-Survey 2021 ist zur Erreichung der Mehrheit seiner Ziele wegen des untauglichen Untersuchungsdesigns nicht geeignet. Seine Ergebnisse bilden aufgrund fehlender Präzision und Intransparenz keine belastbare Entscheidungsgrundlage hinsichtlich der Bewertung und Anpassung gesetzlicher Regelungen im Glücksspielbereich:

- 1) **Gescheiterter methodischer Neustart:** Die gewählte Stichprobe im Mixed-Mode-Design (Telefon plus Online-Befragung) führt weder zu repräsentativen noch zu präzisen explorativen Daten. Nach aktuellem Forschungsstand ist eine Verschlechterung der Datenqualität gegenüber klassischen Telefonbefragungen, wie sie in Vorgängerstudien der BZgA durchgeführt wurden, höchst wahrscheinlich. Eindeutiges Indiz für die unzureichende Qualität ist die Nonresponse-Quote, d.h. der Anteil der Antwortverweigerer, die bei der Online-Befragung des Surveys deutlich höher (90 %) als bei der telefonischen Befragung (73 %) ist. Dies wird im Survey allerdings nicht weiter diskutiert. Stattdessen zitieren die Autoren selektiv Passagen aus der Literatur bezüglich der Vorteile von Online-Befragungen, obwohl die zitierte Literatur selbst von diesen abrät, wenn das Ziel bevölkerungsrepräsentative und präzise Daten sein sollen.
- 2) **Handwerkliche Fehler in Datenerhebung und -auswertung:** Die Analyse der methodischen Fehler und Irrtümer des Surveys anhand des „Total Survey Error Frameworks“, ein allgemein anerkannter Rahmen zur Prüfung der Qualität von Erhebungen, belegt: Die durch handwerkliche Fehler bedingten Verzerrungen der erhobenen Daten bilden die Realität keinesfalls präzise, sondern allenfalls unscharf ab. Zufallsbedingte Schwankungen der Ergebnisse – der sog. „Stichprobenfehler“, der üblicherweise in Form von Konfidenzintervallen angegeben wird – werden von den Survey-Autoren nicht hinreichend kommuniziert und tendenziell unterschätzt. Zudem werden Teilnehmer bei der Online-Befragung nicht zufällig ausgewählt, sondern entlang ihres individuellen Glücksspiel-Interessen und ggf. durch monetäre Anreize für die Befragung gewonnen: Hierdurch ergibt sich methodisch bedingt eine künstliche Erhöhung der Zahl von Befragten, die an Glücksspielen teilnehmen und Probleme aufweisen. Die genutzte, sehr simple Gewichtung durch Poststratification zum Ausgleich solcher Verzerrungen ist, wenn überhaupt, nur zur Korrektur des telefonischen Teils der Befragung geeignet; auch die Mode-Gewichtung im Survey überzeugt aufgrund der Heranziehung überholter Begründungen und genutzter Aggregatdaten nicht.
- 3) **Methodisch unzulässige Kausalschlüsse:** Alle im Survey aufgestellten und dem Rezipienten nahegelegten Kausalitätsvermutungen u.a. in Bezug auf die Identifikation von „riskanten Glücksspielformen“, in Bezug auf Alkoholkonsum oder allgemeine psychische Probleme sind

unzulässig. Querschnittuntersuchungen wie der Survey können nur die Äußerungen der Befragten am Tag der Befragung registrieren und Korrelationen verschiedener Sachverhalte aufzeigen. Kausalitäten können statistisch grundsätzlich nur in aufwendigen Längsschnittuntersuchungen untersucht werden.

- 4) **Falscher Messzeitpunkt:** Als valide „Nullpunktmessung“ zur Erfassung der Auswirkungen des GlüStV 2021 ist der Survey gänzlich ungeeignet, da die Befragung Daten vor und nach dessen Inkrafttreten mischt (Befragung zwischen August und Oktober 2021 in Bezug auf das Glücksspielverhalten der letzten zwölf Monate).

- 5) **Absenkung der Schwellenwerte für problematisches Spielverhalten (nach DSM-5):** Das Diagnosemanual der Amerikanischen Psychiatrischen Gesellschaft ist das führende Klassifikationssystem zur Definition psychischer Erkrankungen. Aktuell gilt die Version DSM-5. Die bisher im deutschen Sprachraum unter Verwendung der DSM-Kriterien durchgeführten Untersuchungen gingen von problematischem Glücksspielverhalten (präklinische und klinische Störung) aus, wenn mindestens vier der neun DSM-5-Kriterien erfüllt waren. Der Glücksspielsurvey reduziert diesen Schwellenwert für die Annahme problematischen Spielverhaltens entgegen allen vergleichbaren vorherigen Untersuchungen von der Erfüllung von mindestens vier auf ein Kriterium, wodurch sich der Anteil der Personen in Deutschland zwischen 18 und 70 Jahren, die der Survey als problematische Spieler einstuft, von 2,3 auf 8,0 Prozent (auf rund 5 Mio. Menschen) vervielfacht. Die Gruppe der neu als problematisch eingestuften Spieler bezeichnen die Survey-Autoren als „riskante Spieler“ (5,7 %). Die aus der Literatur entnommene Kategorie des „riskanten Spielers“ ist wissenschaftlich umstritten. Die Autoren des Surveys führen keine eigenen Belege zur Rechtfertigung eines solch niedrigen Schwellenwertes an. Wegen seiner Anlage als Querschnittuntersuchung kann mit dem Survey auch auf keinen Fall belegt werden, dass sich aus riskantem Glücksspielverhalten gestörtes entwickelt.

- 6) **Behinderung kritischer Auseinandersetzung mit dem Survey:** Die Bereitstellung der detaillierten Survey-Daten, des Feldberichts und des Fragebogens haben die Survey-Autoren trotz begründeter Anfrage der Gutachterin verweigert. Sie verstoßen damit fundamental gegen gute wissenschaftliche Praxis und den Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ der Deutschen Forschungsgemeinschaft, der explizit die Replizierbarkeit von Forschung durch Herausgabe der Forschungsdaten und relevanten Materialien betont sowie die Bedeutung einer kritischen wissenschaftlichen Diskussion über Forschungsergebnisse und ihre Limitationen hervorhebt.

III. Schlussfolgerung

Die Survey-Autoren haben auf eine transparente und umfangreiche Diskussion der Aussagekraft sowie der Limitationen ihrer Arbeit verzichtet. Hinzu kommen gewichtige handwerkliche Fehler sowie undifferenzierte Aussagen in der den Survey begleitenden Kommunikation. Akkurates wissenschaftliches

Vorgehen nach aktuellem Forschungsstand und die transparente Darlegung der Einschränkungen der eigenen Arbeit sind jedoch insbesondere bei einer an die Öffentlichkeit gerichteten Publikation essenziell. Dem Laien, d.h. dem Nicht-Statistiker/-Wissenschaftler, wird eine sachgerechte Rezeption des Surveys hierdurch nahezu verunmöglicht. Die Folge sind zahlreiche statistische Fehlinterpretationen, inkorrekte und unterkomplexe Rezeptionen des Surveys durch gesellschaftliche, politische und mediale Akteure, die diesen nicht anzulasten sind: Die Rezipienten bzw. Multiplikatoren sind gleichsam die Opfer fehlerbehafteter Forschung und defizitärer Kommunikation seitens der Survey-Autoren.

Wegen seiner evidenten Intransparenz und der nachgewiesenen methodischen Fehler und daraus resultierender Verzerrungen ist von der Nutzung des Glücksspiel-Surveys 2021 im Zusammenhang mit gesellschaftspolitischen Bewertungen und erst recht im Zusammenhang mit glücksspielrechtlicher Regulierung dringend abzuraten. Er liefert keine valide und präzise Datenbasis in Bezug auf das Glücksspielverhalten in Deutschland.

IV. Ausblick

Für zukünftige Surveys sind aus wissenschaftlicher Sicht grundsätzlich Längsschnittstudien zu empfehlen, die anders als der gewählte, teils nichtprobabilistische Mixed-Mode-Ansatz auf einer reinen Wahrscheinlichkeitsstichprobe (Probability Sampling) mit wohldefiniertem Stichprobenrahmen basiert. Solche Studien sollten von Beginn an die Erhebung verschiedener Hilfsvariablen zur Korrektur von Antwortverweigerung und weiteren Verzerrungen einplanen; die Gewichtung selbst muss sich, wie auch die übrigen genutzten Methoden, an aktueller methodischer Literatur orientieren. Ebenso ist ein Pretest des Fragebogens für Screening-Zwecke aufgrund des sensiblen Inhalts der Fragen dringend zu empfehlen. Selbstverständlich sollten auch Empfehlungen guter wissenschaftlicher Praxis vollumfänglich umgesetzt werden, also Limitationen transparent diskutiert und die Daten sowie alle relevanten Informationen (Fragebogen, Feldbericht, Code) proaktiv öffentlich zur Verfügung gestellt werden.