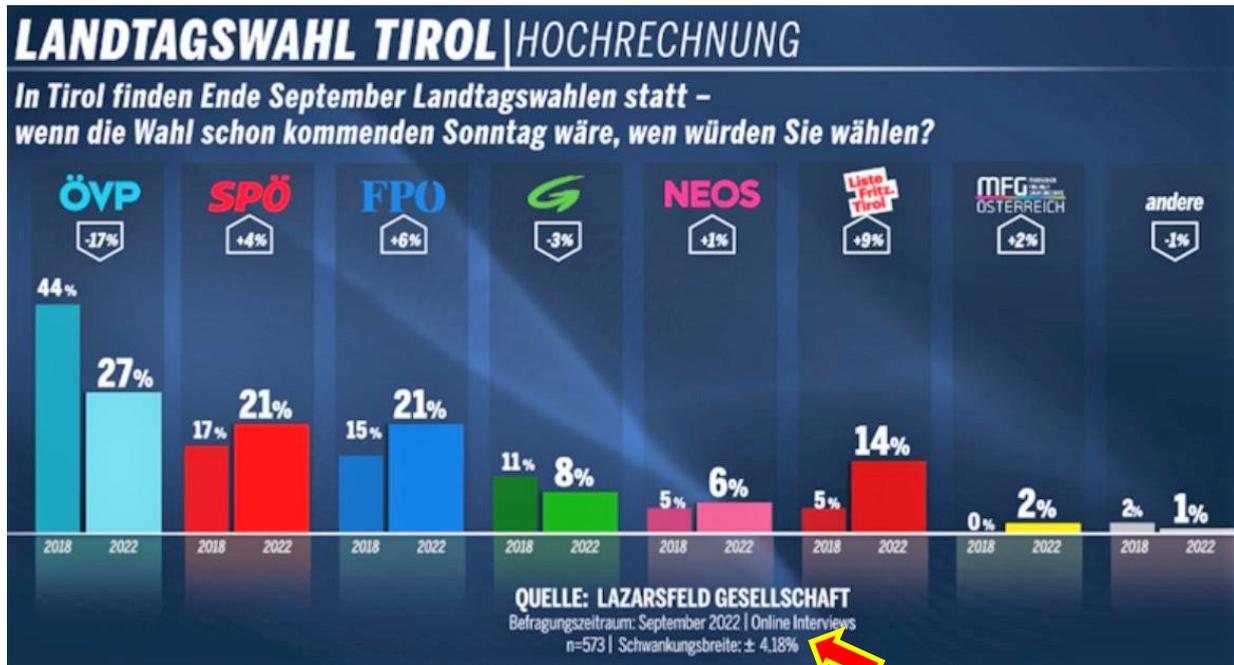


**Unsinn in den Medien – Vom allzu sorglosen Umgang mit Daten:  
Repräsentativität von Stichproben**



(<https://www.oe24.at/oesterreich/politik/landtagswahl-tirol/tirol-wahl-umfragen-sagen-erdbeben-voraus/531147769>; Zugegriffen: 03.10.2022)

**Kommentar:** Im TV-Sender oe24.TV werden regelmäßig Ergebnisse von Umfragen präsentiert. Vor der Landtagswahl in Tirol waren darunter auch hochgerechnete Umfrageergebnisse für diese Wahl („Sonntagsfrage“). Eine aus diesen Ergebnissen explizit gezogene Schlussfolgerung war, dass die ÖVP nach 44,3 % bei der vorangegangenen Wahl diesmal massive Verluste einfahren und unter 30 % fallen wird (siehe Grafik). Das tatsächliche Wahlergebnis von 34,7 % wenige Tage später zeitigte für die ÖVP einen relativen Anteilsverlust von 9,6 %-Punkten. Die Prognose eines massiven Verlusts war korrekt, der vorhergesagte Absturz auf unter 30 % blieb jedoch deutlich aus. Die Ursachen dafür können mannigfaltig sein (z. B. sogar dass sich wegen der Vorhersagen kurz vor der Wahl frühere ÖVP-Wähler:innen doch noch zur Wahl „ihrer“ Partei entschieden haben).

Ohne Zweifel aber handelte es sich bei dieser „Onlinebefragung“ nicht um eine Zufallsauswahl aus den Wahlberechtigten. Dieses Faktum entzog der Anwendung der Wahrscheinlichkeitstheorie auf die Stichprobenergebnisse den dafür nötigen Boden. Dass, obwohl auf diese Weise keine gleichen Auswahlwahrscheinlichkeiten für alle Wahlberechtigten vorliegen konnten, mit der **Formel für einfache Zufallsauswahlen** mit

$$u_{1-\alpha/2} \cdot \sqrt{\frac{\pi \cdot (1-\pi)}{n}} \approx 2 \cdot \sqrt{\frac{0,25}{573}} = 0,0418$$

(Stichprobenumfang n=573, 97,5%-Perzentil der Standardnormalverteilung  $u_{1-\alpha/2} \approx 2$ ;  $\pi = 0,5$ , weil das der in Bezug auf die Genauigkeit schlechteste Fall wäre; siehe etwa: Quatember, A. (2020). Statistik ohne Angst vor Formeln) eine maximale Schwankungsbreite von 4,18 %-Punkten (ja, es sind %-Punkte!) für die Stichprobenergebnisse angegeben wurde, ist somit nicht zu rechtfertigen. Es gibt bei Online-Auswahlen ohne Möglichkeit der Bestimmung der Auswahlwahrscheinlichkeiten einfach keine solche Formel. Bei Angabe von Schwankungsbreiten ist man immer verpflichtet, alle in der Berechnung implizit unterstellten Annahmen zu beschreiben. In diesem Fall wäre dabei also hinzuzufügen: Unter Annahme einfacher Zufallsauswahl der Befragten aus allen Wahlberechtigten. Was es nicht war ...

(Für den Kommentar verantwortlich: Andreas Quatember, IFAS)