

National Survey Investigating Hallucinogenic Trends (NSIHT)

Study Design and Methodology

Technical Methods Report

Rocky Mountain Poison & Drug Safety, a Division of Denver Health and Hospital
Authority, Denver CO

Last Updated: September 2026

Purpose

To describe the design, measurement framework, sampling, data quality procedures, fused-survey weighting, statistical inference, and ongoing validation approach used by NSIHT.



National Survey Investigating Hallucinogenic Trends (NSIHT): Study Design and Methodology

Acknowledgements and Funding

NSIHT is owned and operated by Denver Health's Rocky Mountain Poison and Drug Safety (RMPDS). Data collection and investigators are supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health (NIH) under Award No. R01DA064612. Data reports and topical expertise are conducted in collaboration with the Substance Abuse and Mental Health Services Administration (SAMHSA), as well as previous methodological development under annual grants #75S20124P00003, #75S20123P00004, and #HHSS283201700075I/75S20321F42002. The views expressed on this website are those of RMPDS and not NIH, NIDA, nor SAMHSA.

Authoring Organization

This report was authored by the National Survey Investigating Hallucinogenic Trends (NSIHT) team at RMPDS, a division of Denver Health and Hospital Authority (DHHA). For questions about the report or requests for data, please contact info@rmpds.org.

Location: 1391 North Speer Blvd, Suite #600, Denver, CO 80204

Mailing Address: 777 Bannock Street, M/C 0180 Denver, CO 80204

Telephone: (866) 871-4980

Public Domain Notice

All material appearing in this report are in the public domain and may be used with appropriate citation to the source material. This material may not be directly reproduced or distributed without the specific written authorization of RMPDS.

Recommended Citation

Rocky Mountain Poison & Drug Safety. (2026). *National Survey Investigating Hallucinogenic Trends (NSIHT): Study Design and Methodology* (RMPDS Publication No. RMPDS-2026-I-02). Rocky Mountain Poison & Drug Safety, Denver Health and Hospital Authority. Denver, CO. <https://www.rmpds.org/nsiht>

Publication Number: RMPDS-2026-I-02

Date Published: September 2026

Table of Contents

1. Study Overview and Objectives	4
1.1 Overarching Measurement Framework	4
2. Target Population and Study Design	5
2.1 Sampling Frame and Recruitment	6
2.2 Eligibility and Analytic Inclusion	6
2.3 Substances Covered	6
2.4 Guardrail Quota Sampling	7
2.5 Anchor Survey Recruitment and Sampling	7
3. Questionnaire Development and Measures	7
3.1 Recall Period and Adaptive Structure	8
3.2 Questionnaire Measurement Reliability	8
4. Data Quality and Processing	9
5. Weighting and Calibration	10
5.1 Transport Calibration of the Enriched Sample	11
5.2 Population Calibration	11
5.3 Analytic Weight and Pooled Launces	11
6. Statistical Analysis	11
6.1 Variance Estimates and Design Effects	11
6.2 Subgroup and Comparative Analyses	11
7. Missing Item-level Data and Logical Assignment	12
8. Uncertainty, Reporting, and Suppression Rules	12
9. Disclosure and Privacy Protections	12
10. Ongoing Survey Maintenance and Validation	13
11. Strengths and Limitations	13
11.1 Strengths	13
11.2 Limitations	13
12. Reproducibility and Data Availability	14
13. Funding and Required Disclosures	14
14. References	15

1. Study Overview and Objectives

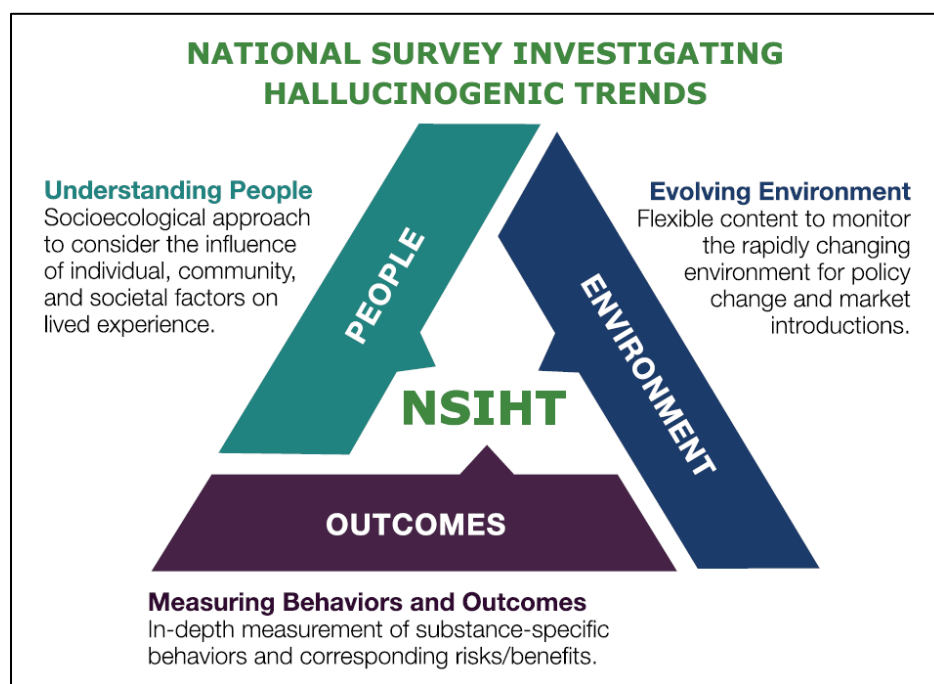
The National Survey Investigating Hallucinogenic Trends (NSIHT) is a repeated cross-sectional, fused survey surveillance study designed to monitor psychedelic and related substance use among adults in the United States (U.S.). The study was developed by Rocky Mountain Poison & Drug Safety (RMPDS), a division of Denver Health and Hospital Authority. Development began in 2023 and the survey was first fielded in 2024. NSIHT is a self-administered online questionnaire and uses a sampling strategy enriched for adults in the U.S. reporting use of psychedelic, hallucinogenic, or dissociative substances in the prior 12 months, a group which will be referred to as “psychedelics” hereafter. Data from the enriched survey are fused with a large general-population anchor survey and calibrated to support nationally representative population inference.^{1,2}

NSIHT was created to address limitations of other surveillance systems. Large general-population surveys provide breadth of data across many health metrics and substance use behaviors but have limited questionnaire space and relatively few observations for low-prevalence, substance-specific behaviors. Targeted surveys can provide depth but commonly lack an explicit basis for generalization based on targeted sampling. The fused survey design combines these strengths while minimizing associated limitations by oversampling the population of interest with an in-depth survey and then correcting compositional differences through transport weighting and population calibration.^{1,2}

1.1 Overarching Measurement Framework

NSIHT is organized around three complementary domains that guide surveillance priorities, questionnaire development, and interpretation of the study as an ongoing public-health monitoring system (Figure 1).¹

Figure 1. Complementary Three-Domain Measurement Framework



These three domains guide both research priorities and survey content. The purpose and examples of assessment for each domain are outlined in Table 1.

Table 1. NSIHT Three-Domain Measurement Framework

Domain	Purpose	Examples of Assessment
Measuring Behaviors and Outcomes	Characterize substance-specific behaviors and corresponding perceived risks, benefits, and health outcomes.	Prevalence, frequency, recency, reasons for use, set/setting, concomitant substance use, experiences, adverse events, and self-perceived outcomes. Use of self-reported validated scales.
Understanding People	Use a socioecological perspective to understand how individual, peer, community, and societal factors relate to psychedelic use and lived experience.	Demographics, physical and mental health, substance use, well-being, peer/community factors, and contextual characteristics.
Evolving Environment	Maintain flexible surveillance capacity as policy, markets, products, terminology, and patterns of use change.	State-regulated access, emerging substances and products, terminology, market changes, policy-relevant behaviors, and exploratory modules.

2. Target Population and Study Design

The target population for population-level inference is the U.S. adult population (ages 18+ years). The enriched NSIHT survey specifically recruits adults who report use of at least one psychedelic substance during the prior 12 months. The repeated, cross-sectional survey is fielded semiannually in the summer and winter resulting in ongoing surveillance over time.

The cross-sectional observational study design combines intensive measurement of psychedelic use with population generalization through a transport weighting design. The NSIHT survey provides detailed data on people who use psychedelics. It is fused to the Survey of Non-Medical Use of Prescription Drugs (NMURx), which supplies a large general-population anchor sample, including adults who do and do not report psychedelic use. NMURx has been shown to provide externally valid drug use estimates in the U.S.³ The two sources are linked analytically through data fusion and calibration weighting rather than through person-level linkage (see [Section 5](#) for more details).²

Interpretation of “nationally-representative”

In NSIHT, nationally-representative inference is produced analytically after data fusion and calibration. External validity represents the generalizability of findings outside the study sample into the target population from which the sample was recruited; a survey with good external validity is considered “representative”.⁴ Estimates are meant to be nationally representative of all adults in the U.S. The unweighted enriched psychedelic use sample is intentionally not a sample of all U.S. adults; the fused survey design is intended to provide sufficient observations and detailed data for study of psychedelic use.

2.1 Sampling Frame and Recruitment

The study design intentionally separates intensive measurement of the target behavior measuring psychedelic use from population generalization. The NSIHT survey provides detailed data on people who use psychedelics, whereas the NMURx survey supplies a large general-population anchor sample.

The enriched psychedelic sample is recruited from the YouGov commercial online panel. Panel membership is maintained independently from participation in this study, and panelists receive opportunities to participate in surveys spanning many topics. YouGov is responsible for ongoing recruitment and maintenance of the broader commercial panel. The broader panel is recruited through multiple channels and therefore the pool this psychedelic study recruits from is neither a community of psychedelic enthusiasts nor critics. Invitations are presented generically rather than identifying the survey as a psychedelic-focused study before eligibility screening, reducing preferential entry by individuals specifically seeking drug-related surveys.^{1,2}

Eligible respondents can enter and exit the online questionnaire during the open field period, subject to consent, eligibility screening, and unfilled sampling quotas. The commercial panels employed a range of quality assurance methods, such as ongoing panel validation, payment fraud detection, location checks, and bot removal during data collection to ensure respondent quality and evasion of large language model infiltration.^{5,6} Upon completion of the survey, respondents receive points based on survey length that can be redeemed for monetary compensation (approximately equivalent to 4 to 5 USD).

2.2 Eligibility and Analytic Inclusion

Eligibility and analytic inclusion criteria are applied during fielding and during post-collection data processing. Survey respondents included in the NSIHT sample must meet the following conditions:

- Agree to the implied consent language at the beginning of the survey
- Be 18 to 110 years of age
- Report use of at least one psychedelic drug during the prior 12 months – specific drugs may change over time
- Not complete more than one NSIHT survey within a calendar year
- Fall within a region or age-by-sex sampling stratum through quota sampling (see [Section 2.4](#))
- Complete the questionnaire in its entirety
- Pass post-hoc data-quality checks for careless or inattentive responding (see [Section 4](#))

2.3 Substances Covered

NSIHT eligibility requires the use of at least one psychedelic substance during the prior 12 months, based on lifetime and past 12-month report of the following listed substances. These substances are presented to respondents in a randomized order. They are subject to change over time.

- Ayahuasca
- DMT (N,N-dimethyltryptamine)
- Esketamine

- Ibogaine
- Illicit ketamine
- LSD (Lysergic acid diethylamide)
- MDA (3,4-methylenedioxyamphetamine)
- MDMA (3,4-methylenedioxymethamphetamine)
- PCP (Phencyclidine)
- Peyote
- Prescription ketamine
- Psilocybin or psilocin
- *Salvia divinorum*
- San Pedro
- Synthetic mescaline
- 2C-B (4-bromo-2, 5-dimethoxyphenethylamine)
- 5-MeO-DMT (5-O-methyl-bufotenin-DMT)

2.4 Guardrail Quota Sampling

NSIHT uses guardrail quota sampling to maintain demographic and geographic variation during enrollment. Geographic quotas are specified independently, while age and sex are cross-classified. The geographic strata collection targets are updated annually and are based on the most recent available U.S. Census Bureau 1-year estimates.⁷ Colorado and Oregon are oversampled separately from other Western states because of their distinct policy environments. The sex-by-age strata collection targets are centered around the median age of psychedelic use as determined by the NMURx population estimates in the most recent year. Current sex quotas are approximately 60% male and 40% female, and age quotas are balanced within sex around median age of 35 years. Each survey launches targets approximately 2,000 eligible respondents. Geographic strata have guardrail minimum quotas of 5% of the target strata sample; age and sex strata have guardrail minimum quotas of 10% of the target. Guardrail maximum quotas are set 10% above the corresponding target strata. Once a stratum reaches its maximum, additional respondents in that stratum cannot complete the survey. Fielding closes after minimum quota requirement and the total target sample size are achieved.

2.5 Anchor Survey Recruitment and Sampling

The NMURx anchor survey is recruited separately through Kantar. NMURx is a large national online survey of the adult population and is fielded semiannually in the spring and fall. It includes demographic and substance-use measures aligned with NSIHT and contains an identifiable analytic subpopulation reporting past-12-month psychedelic use. NMURx has previously been validated for national drug-use estimation.³

3. Questionnaire Development and Measures

NSIHT questionnaire development combined subject-matter expertise, focus groups findings of people with lived experiences,⁸ language preference testing,⁹ content-validity work, and other psychometric evaluation. The questionnaire went through a multi-step validation process including expert review, pilot testing with content validation, revision, and a reliability test-retest sub-study.¹⁰

The questionnaire uses best practices for online surveys including plain language,¹¹ direct recall question prompting,¹² modern slang terms, personalized time references,¹³ adaptive routing so that respondents receive detailed follow-up modules for the substances and experiences relevant to them. Question order and response-option order are randomized where appropriate to reduce order effect bias.¹⁴ The survey interface also uses a device-responsive design.

The core questionnaire focuses on past-12-month psychedelic use. For substances endorsed in the 12 months prior to survey completion, respondents enter substance-specific loops that collect information about topics such as: frequency and recency of use, reasons for use, companions while using, source of substances, concomitant use with other substances, set and setting context, and experienced or perceived outcomes. Additional sections measure demographics, other substance use (including prescription and non-prescription substances), health conditions, and factors at the individual, peer, community, and societal levels.

Validated instruments used in NSIHT include the Drug Abuse Screening Test-10 (DAST-10),^{15,16} Patient Health Questionnaire-9 (PHQ-9),¹⁷ Generalized Anxiety Disorder-7 (GAD-7),^{18,19} Short Warwick-Edinburgh Mental Wellbeing Scale (S-WEMWBS),^{20,21} and an adapted shortened version of the Swiss Psychedelic Side Effects Inventory,²² among other health measures.

3.1 Recall Period and Adaptive Structure

The principal surveillance window is the past 12 months at the time of survey completion. This interval is intended to measure psychedelic behaviors that may be episodic or infrequent within the defined reference period while maximizing reliability and minimizing recall bias.²³ Survey adaptive structure is integral to the measurement model, where irrelevant questions are not displayed. Therefore, some downstream values are logically assigned from upstream responses.

3.2 Questionnaire Measurement Reliability

NSIHT validation included a formal test-retest assessment of single-item substance-use measures and related behaviors. In a 2024 test-retest sub-study, all respondents completing the relevant NSIHT launch were invited to re-consent for a re-test.¹⁰ The same question wording was used, substance order was independently randomized, and test and re-test data were linked through a hashed, nonidentifying participant identifier. The analysis was pre-registered, and the final re-test sample included 690 retest respondents.²⁴

Reliability was evaluated using overall agreement, average positive agreement, average negative agreement, and Gwet's first-order agreement coefficient (AC1).²⁵ Gwet's AC1 was selected because low-prevalence binary behaviors can produce misleadingly low kappa values despite high observed agreement. Across the 17 evaluated substances, AC1 values were in the substantial-to-excellent range from an interpretation perspective (Table 2).²⁶ Positive agreement decreased as substance endorsement became rarer, an important qualification when interpreting high overall agreement for low-prevalence items. However, the AC1 remained high for these groups, suggesting the question is reliable overall.

Table 2. Test-retest Reliability of Past 12-month Substance Use Measures*

Note: Adapted with permission from Rockhill et. al., Reliability of Self-Reported Real-World Psychedelic Use – A Test-Retest Study in a Population Sample, Psychedelics (2026)

Substance	Average Overall agreement, %	Average Positive agreement, %	Gwet's AC1
Psilocybin or psilocin	86.4	86.8	0.73
MDMA	85.1	77.9	0.73
Prescription ketamine	90.0	79.7	0.84
LSD	88.3	61.6	0.84
2C-B	88.3	60.9	0.84
DMT	90.4	61.2	0.88
Esketamine	90.4	46.1	0.89
Ibogaine	90.9	50.4	0.89
PCP	91.9	43.0	0.91
<i>Salvia divinorum</i>	92.2	46.0	0.91
MDA	91.7	41.3	0.90
Ayahuasca	91.7	30.0	0.91
Synthetic mescaline	92.5	38.2	0.91
San Pedro	92.2	37.2	0.91
5-MeO-DMT	92.5	35.0	0.92
Peyote	92.6	34.2	0.92
Illicit ketamine	92.8	53.0	0.92

* Negative agreement, not shown, was uniformly high driving high overall agreement values; positive agreement is retained here to make reliability of affirmative endorsements visible.

Reasons for use and recency measures were evaluated using analyses that account for survey adaptive questioning and respondent multiple substance endorsement. Mixed-effects logistic regression models included drug and response option as fixed effects and a respondent-level random intercept. After adjustment, estimated probability of a reliable response for reasons for use and recency were all above 0.95.¹⁰

4. Data Quality and Processing

Data quality procedures operate at both the panel-administration and researcher stages. During fielding, the survey administration company uses ongoing panel validation, payment-fraud detection, location checks, bot-removal procedures, and related controls intended to reduce low-quality or automated responding. Completed survey data are transferred through secure hosting and stored on RMPDS secure servers before creating the analysis dataset.

After collection, RMPDS applies post-hoc checks for careless or inattentive response patterns.²⁷ Current procedures include evaluation of implausibly brief response times, long sequences of identical answers to substance-endorsement items, and weak internal correspondence among selected mental-health responses. Respondents meeting exclusion criteria are removed before final weighting and analysis. These techniques have been shown to meaningfully change estimates of low-prevalence behaviors.²⁸

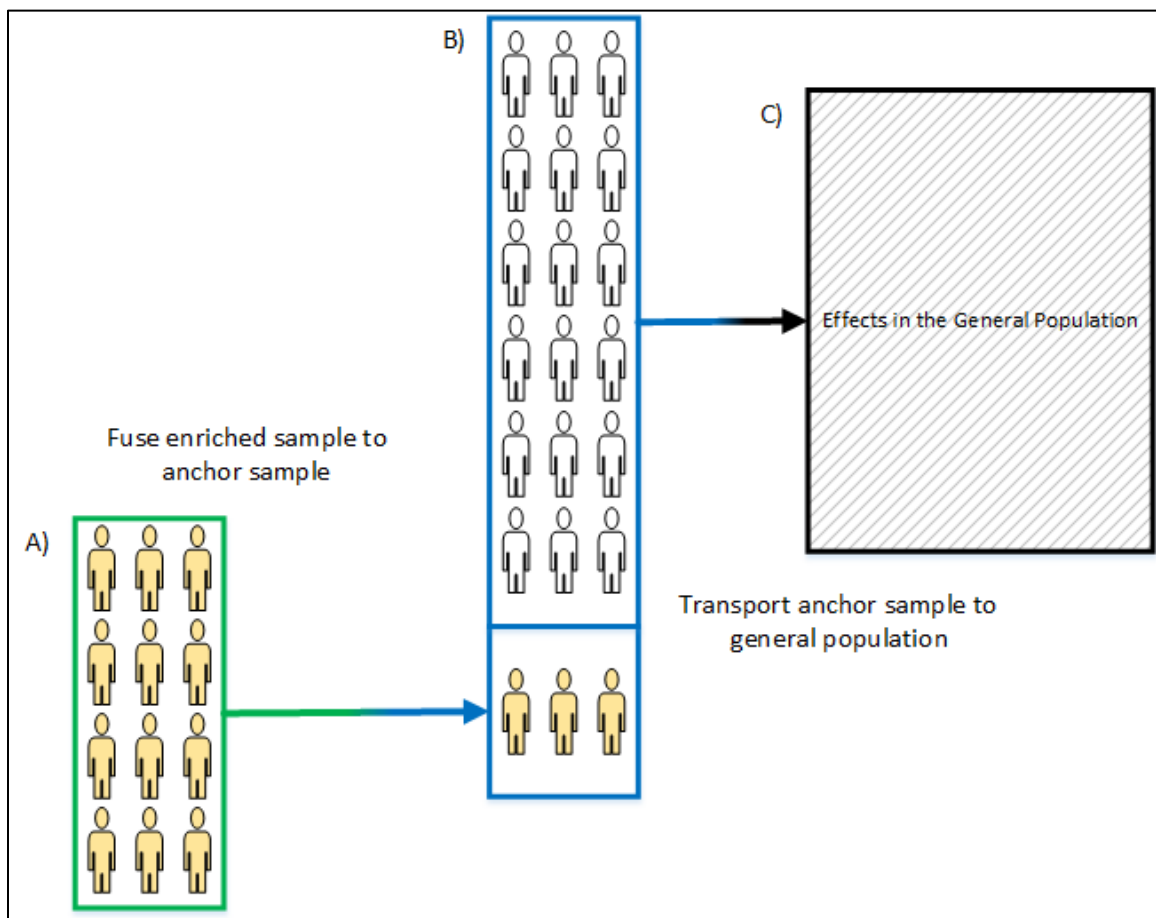
5. Weighting and Calibration

NISHT utilizes data fusion and calibration weighting to generate nationally representative estimates of psychedelic use behaviors among adults in the U.S.

All population-level NSIHT analyses require the person-level analytic weight. The weighting system is applied separately to each survey launch after respondents have completed the questionnaire and passed post-hoc quality checks. The approach is a two-step fused survey design that separates inter-survey selection-bias correction from population generalization (Figure 2). The approach was pre-registered;²⁹ a validation study has been published elsewhere,² but is briefly described here.

Figure 2. Conceptual Diagram of Fusing an Enriched Sample with an Anchor Sample to Create Generalizable Estimates

Note: Adapted with permission from Rockhill et al., Fusing Specialized Surveys of Rare Populations to Larger Surveys for Generalized Inference: Cross-Sectional Survey Study, J Med Internet Res. (2026)



5.1 Transport Calibration of the Enriched Sample

In the first step, the enriched psychedelic sample is calibrated to the weighted past 12-month psychedelic-use subpopulation in the NMURx anchor survey. Generalized raking is used to adjust individual weights so that selected marginal distributions align between the enriched and anchor samples.³⁰ This step produces transport weights intended to correct composition differences arising from the different recruitment and sampling strategies. Following transport calibration, the enriched psychedelic sample replaces the past 12-month psychedelic-use subsample from the anchor survey. Anchor survey respondents who did not report past 12-month psychedelic use remain in the fused sample and retain their original analytical weights.

The peer-reviewed validation study evaluated a large set of candidate transport-weighting schemes.² A parsimonious scheme was selected based on low transport bias across independent demographic, health, and substance-use evaluation measures. The validated scheme included core demographic variables and selected health/attitudinal variables which may be updated over time to account for changes in population composition. Replicability of the approach was tested across survey launches and the fusion did not perturb the subpopulation using psychedelic drugs across samples.

5.2 Population Calibration

The combined survey is calibrated to population benchmarks, including state-level population counts. Following the calibration steps, weighted estimates are interpreted for the U.S. adult residential population (i.e., the target population). During validation procedures, survey weights were stable in the tails. If extreme weights are detected across launches, weights are trimmed to the 99th percentile.

5.3 Analytic Weight and Pooled Launches

The analytic dataset includes a person-level weight representing the number of adults in the target population (all 50 states and Washington D.C.) represented by each respondent. When multiple launches are pooled into one analysis, the person-level weight is divided by the total number of launches being combined so that the sum of the adjusted weights remains aligned with an average population total across launches.

6. Statistical Analysis

6.1 Variance Estimates and Design Effects

The primary descriptive estimands are weighted prevalence proportions and weighted population counts. Analyses use statistical software capable of incorporating the NSIHT person-level weights, 95% confidence intervals (CI) may be calculated using Taylor series expansion or bootstrap CI depending on the analysis.

6.2 Subgroup and Comparative Analyses

NSIHT supports subgroup analyses when the unweighted sample size and precision of the weighted estimate are sufficient. Analytic domains may include demographic groups, geography, substance-specific use, frequency or reason-of-use groups, health characteristics, or other pre-specified subpopulations. Because many psychedelic behaviors are low

prevalence, subgroup estimates require particular attention to the suppression and instability criteria described (see [Section 8](#)).

Comparisons across survey launches should use launch-specific weights or appropriately rescaled pooled weights. Repeated cross-sectional data support population trend comparisons but should not be interpreted as longitudinal follow-up of the same respondents; participants are restricted from completing more than one NSIHT survey within a calendar year, although individuals may participate in different calendar years.

7. Missing Item-level Data and Logical Assignment

The online questionnaire requires responses and displays items based on previous responses via adaptive structures; therefore, there is no conventional item-specific nonresponse. Downstream items may be logically assigned based on previous answers.

8. Uncertainty, Reporting, and Suppression Rules

NSIHT applies disclosure and statistical-stability rules to all reported estimates. These rules are intended to reduce confidentiality risk and prevent presentation of estimates that are too unstable for responsible interpretation (Table 3).

Table 3. Current NSIHT Suppression and Instability Rules

Criterion	Action	Current Reporting Rule
Unweighted numerator <5	Suppress	No count or summary statistic may be shown; this criterion reduces privacy-breach risk.
Coefficient of variation >0.80	Suppress	No count or summary statistic may be shown because of statistical instability in the estimate.
Coefficient of variation ≥ 0.33 and ≤ 0.80	Flag	The estimate may be shown but should be identified as potentially unstable in reporting and interpreted with caution.
Estimated prevalence <0.0005	Suppress	The estimate may not be shown because of statistical instability.

9. Disclosure and Privacy Protections

NSIHT is administered confidentially through the commercial survey panel infrastructure. Personally identifiable information beyond county of residence is not collected in the research questionnaire, and information held by the panel administrators that could identify respondents is not made available to RMPDS researchers. Study datasets transferred for research use are deidentified; only Census Region geographies are included.

Third party access to limited data is governed through an approved request and data-use arrangement with RMPDS and DHHA. Public reporting is additionally constrained by the suppression rules above. These controls are intended to protect participant confidentiality while allowing scientifically useful population-level analysis.

10. Ongoing Survey Maintenance and Validation

NSIHT is designed as an evolving surveillance platform rather than a static questionnaire. Methodological maintenance is therefore part of the study design. Survey content, terminology, constructs, and analytic procedures are reviewed as patterns of psychedelic use, products, policy environments, and public language change over time.^{1,9}

Ongoing maintenance includes qualitative assessment of whether important behaviors remain represented, review of question comprehension and terminology, continued content-validity work, psychometric evaluation of novel measures, monitoring of data-quality indicators, and periodic reassessment of weighting and external validity. Space is reserved on the questionnaire for emerging topics while preserving stable core measures needed for surveillance trends, supporting the third NSIHT measurement domain: Studying the Evolving Environment.

Multiple best practices have been implemented to support target validity of the survey. Measurement reliability evaluates consistency of respondents' answers; data-quality procedures reduce careless or automated responses; and the fused-survey validation framework evaluates whether weighted estimates are internally consistent and externally aligned with national benchmarks. These forms of target validation evidence address different methodological concerns and are interpreted together rather than as interchangeable forms of validation.

11. Strengths and Limitations

11.1 Strengths

- Enriched sampling produces a substantially larger analytic sample of people who use psychedelics in the past 12-months than would be feasible in a similarly sized general-population survey, supporting detailed substance-specific analyses.
- The fused survey design provides an explicit framework for transporting enriched-sample estimates to a nationally representative target population.
- Open science principles were followed, including pre-registration of studies and peer-reviewed publication of findings. Multiple types of validation and reliability are published in scientific literature. This “methods suite” of publications supports the overall target validity of the survey.^{1,2,8-10}
- The survey combines detailed behavioral measurement with health, demographic, and socioecological measures, allowing context and outcomes to be examined alongside prevalence.
- Questionnaire development incorporates lived-experience input, subject-matter expertise, comprehension work, adaptive design, and formal reliability testing.
- Selection bias, measurement quality, and estimate stability are addressed through multiple distinct procedures rather than through weighting alone.

11.2 Limitations

- The enriched NSIHT sample is recruited from a nonprobability online panel; weighting can reduce measured selection differences but cannot guarantee correction for unmeasured selection mechanisms. Residual differences from external benchmarks may remain because calibration does not necessarily correct for higher-order interactions or unmeasured selection factors.

- Some populations may remain underrepresented in online panels, including Indigenous communities, institutionalized populations, and adults engaged in the justice system. Weighting cannot create information for groups that are sparsely sampled or not sampled at all.
 - Additionally, the survey content is not intended to study Indigenous use practices.
- All substantive behaviors and health measures are self-reported and may be affected by recall, social-desirability, comprehension, or other measurement biases. Test-retest reliability measures consistency, not accuracy.
- Rare outcomes and small subgroups remain subject to stochastic uncertainty even after weighting. Suppression and precision rules reduce but do not eliminate this limitation.

12. Reproducibility and Data Availability

Core elements of the fused survey weighting methodology have been peer reviewed, and methodological validation work has used pre-registration and independent replication across survey waves. NSIHT analysis should be conducted with software capable of survey-weighted estimation, with all analytic choices documented sufficiently to reproduce the target population, survey launch or pooled-launch definition, weight handling, subpopulation definitions, precision assessment, and suppression decisions.

Data Availability

NSIHT microdata are not publicly available at this time but can be accessed under an approved request/data-use arrangement. Reach out for more information at info@rmpds.org.

13. Funding and Required Disclosures

Data collection and investigators are supported by the National Institute on Drug Abuse (NIDA) of the National Institutes of Health (NIH) under Award No. R01DA064612. Data reports and topical expertise are prepared in collaboration with the Substance Abuse and Mental Health Services Administration (SAMHSA), as well as previous methodological development under prior annual grants #75S20124P00003, #75S20123P00004, and #HHSS283201700 075I/75S20321F42002. NSIHT is owned and operated by Denver Health's Rocky Mountain Poison and Drug Safety. The views expressed on this technical document are those of RMPDS and not NIH, NIDA, nor SAMHSA.

RMPDS is a division of nonprofit Denver Health and Hospital Authority, a subdivision of the State of Colorado. Outside of this work, RMPDS provides independent surveillance, research, and reporting services to government agencies, pharmaceutical manufacturers, and non-government organizations. No other organization or party besides the authors participated in the analysis, drafting, or review of this report.

Use of Artificial Intelligence

Generative artificial intelligence (OpenAI ChatGPT 5-series model) was used to assist with the organization, synthesis, and drafting of this methodology document from pre-existing NSIHT study documentation, protocols, reports, and scientific manuscripts. All methodological content, interpretations, and final language were reviewed and approved by the NSIHT study team.

Artificial intelligence was not used to develop the study design, conduct statistical analyses, or generate study findings presented in this report.

14. References

1. Monte AA, Black JC, Rockhill KM, et al. The need for national understanding of psychedelic use behaviors in the United States: development of the National Survey Investigating Hallucinogenic Trends (NSIHT) [unpublished manuscript]. 2026.
2. Rockhill KM, Bemis EA, Schow N, et al. Fusing Specialized Surveys of Rare Populations to Larger Surveys for Generalized Inference: Cross-Sectional Survey Study. *J Med Internet Res*. Apr 27 2026;28:e86059. doi:10.2196/86059. PMID:PMC13120540
3. Black JC, Rockhill K, Forber A, et al. An Online Survey for Pharmacoepidemiological Investigation (Survey of Non-Medical Use of Prescription Drugs Program): Validation Study. *J Med Internet Res*. Oct 25 2019;21(10):e15830. doi:10.2196/15830. PMID:PMC6914238
4. Westreich D, Edwards JK, Lesko CR, Cole SR, Stuart EA. Target Validity and the Hierarchy of Study Designs. *Am J Epidemiol*. Feb 1 2019;188(2):438–443. doi:10.1093/aje/kwy228. PMID:PMC6357801
5. Westwood SJ. The potential existential threat of large language models to online survey research. *Proc Natl Acad Sci U S A*. Nov 25 2025;122(47):e2518075122. doi:10.1073/pnas.2518075122. PMID:PMC12663962
6. Shakespeare S. *Reality Report 2026: Does Your Data Provider Pass the Reality Check*. 2026. Accessed April 22, 2026. <https://yougov.com/reports/does-your-data-provider-pass-the-reality-check>
7. U.S. Census Bureau. Data from: American Community Survey 1-Year Data.
8. Jewell JS, James C, Black JC. Improving Psychedelic Harm Reduction and Therapy: Community Perspectives from Individuals Who Use Recreationally and for Mental Health Purposes. *J Psychoactive Drugs*. Aug 12 2026;1–8. doi:10.1080/02791072.2026.2710054.
9. Lyons FE, Rockhill KM, Fox EJ, et al. Psychedelic Terminology Preference in the 2024 National Survey Investigating Hallucinogenic Trends (NSIHT). *J Psychoactive Drugs*. Mar 16 2026;1–8. doi:10.1080/02791072.2026.2644863.
10. Rockhill KM, Schow N, Garibaldi D, Monte AA, Black JC. Reliability of self-reported real-world psychedelic use – A test-retest study in a population sample. *Psychedelics*. 2026/12/01/2026;4:100040. doi:<https://doi.org/10.1016/j.psyche.2026.100040>.
11. Bauer I, Kunz T, Gummer T. Plain language in web questionnaires: effects on data quality and questionnaire evaluation. *International journal of social research methodology*. 2025;28(1):57–69. doi:10.1080/13645579.2023.2294880.
12. Klungel OH, de Boer A, Paes AH, Herings RM, Seidell JC, Bakker A. Influence of question structure on the recall of self-reported drug use. *J Clin Epidemiol*. Mar 1 2000;53(3):273–7. doi:10.1016/s0895-4356(99)00167-5.
13. Hjorthøj CR, Hjorthøj AR, Nordentoft M. Validity of Timeline Follow-Back for self-reported use of cannabis and other illicit substances--systematic review and meta-analysis. *Addict Behav*. Mar 2012;37(3):225–33. doi:10.1016/j.addbeh.2011.11.025.
14. Krosnick JA, F. AD. An Evaluation of Cognitive Theory of Response-Order Effects in Survey Measurement. *Public Opinion Quarterly*. 1987;51(2):201–219. doi:10.1086/269029.
15. Yudko E, Lozhkina O, Fouts A. A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. *J Subst Abuse Treat*. Mar 2007;32(2):189–98. doi:10.1016/j.jsat.2006.08.002.

16. Bohn M, Babor T, Kranzler H. Validity of the Drug Abuse Screening Test (DAST-10) in inpatient substance abusers. *Problems of drug dependence*. 1991;119:233–235.
17. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. Sep 2001;16(9):606–13. doi:10.1046/j.1525-1497.2001.016009606.x. PMID:PMC1495268
18. Lowe B, Decker O, Muller S, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care*. Mar 2008;46(3):266–74. doi:10.1097/MLR.0b013e318160d093.
19. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. May 22 2006;166(10):1092–7. doi:10.1001/archinte.166.10.1092.
20. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes*. Nov 27 2007;5:63. doi:10.1186/1477-7525-5-63. PMID:PMC2222612
21. Ng Fat L, Scholes S, Boniface S, Mindell J, Stewart-Brown S. Evaluating and establishing national norms for mental wellbeing using the short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): findings from the Health Survey for England. *Qual Life Res*. May 2017;26(5):1129–1144. doi:10.1007/s11136-016-1454-8. PMID:PMC5376387
22. Calder AE, Hasler G. Validation of the Swiss Psychedelic Side Effects Inventory: Standardized assessment of adverse effects in studies of psychedelics and MDMA. *Journal of Affective Disorders*. 2024/11/15/ 2024;365:258–264. doi:<https://doi.org/10.1016/j.jad.2024.08.091>.
23. Clarke PM, Fiebig DG, Gerdtham U-G. Optimal recall length in survey design. *Journal of Health Economics*. 2008/09/01/ 2008;27(5):1275–1284. doi:<https://doi.org/10.1016/j.jhealeco.2008.05.012>.
24. Rockhill KM, Bemis EA, Schow N, et al. Establishment of Internal Validity of the National Survey Investigating Hallucinogenic Trends (NSIHT). Updated 5 December 2024. <https://doi.org/10.17605/OSF.IO/A65XV>
25. Wongpakaran N, T. W, Wedding D, Gewt KL. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients. *BMC Med Res Methodol*. 2013;doi:10.1186/1471-2288-13-61.
26. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276–82. PMID:PMC3900052
27. Ward MK, Meade AW. Dealing with Careless Responding in Survey Data: Prevention, Identification, and Recommended Best Practices. *Annu Rev Psychol*. Jan 18 2023;74:577–596. doi:10.1146/annurev-psych-040422-045007.
28. Black JC, Rockhill KM, Schow N, Monte AA. Importance of proactive bias mitigation when using online panels: Illicitly manufactured fentanyl in the general population. *JAMA Health Forum*. 2025;6:e254011. doi:10.1001/jamahealthforum.2025.4011.
29. Rockhill K, Bemis EA, Schow N, et al. Establishment of External Validity of the National Survey Investigating Hallucinogenic Trends (NSIHT). Center for Open Science. Updated 19 May 2025. <https://doi.org/10.17605/OSF.IO/856DX>
30. Deville JC, E. SC, Sautory O. Generalized raking procedures in survey sampling. *J Am Stat Assoc*. 1993;88(423):1013–1020. doi:10.1080/01621459.1993.10476369.