

Definitions and determinants: Use of Participatory System Dynamics to understand de-implementation, consensus, and buy-in

Lauren Caton

lcaton@emory.edu

Emory University

Kristen Hassmiller Lich

University of North Carolina at Chapel Hill

Sherri Green

University of North Carolina at Chapel Hill

Larissa Jennings Mayo-Wilson

University of North Carolina at Chapel Hill

Carmen Gutierrez

University of North Carolina at Chapel Hill

Meghan Shanahan

University of North Carolina at Chapel Hill

Christine Tucker

University of North Carolina at Chapel Hill

Method Article

Keywords: Participatory system dynamics, implementation science, systems thinking, de-implementation, health policy, buy-in, consensus building, micro-interlocutor analysis, maternal health, substance use

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10146583/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

The emerging field of policy implementation science needs rigorous methods to understand how practitioners temporally view underlying causes of complex problems such as policy de-implementation. Participatory system dynamics—a system science method for developing shared understanding of dynamic, complex, problems—may therefore aid in 1) shared definitions of, 2) consensus on, and 3) buy-in towards policy de-implementation, which has yet to be contextualized in the literature.

Methods

Here we demonstrate PSD as a method through a case study to evaluating its ability to measure changes to de-implementation actions, policy consensus, and participant buy-in. We engaged twenty-one community practitioners (clinicians, clinic directors, social workers, and policy advocates) from North Carolina (NC) and Kentucky (KY). Two, two-hour participatory system dynamics modeling sessions were conducted in November and January of 2025 with pre- and post-surveys for evaluation.

Results

PSD demonstrated ability to capture changes to multi-dimensional determinants and in the process, more cohesive and measurable changes in de-implementation conceptualizations. Participants ranked more “structural” action items (funding, policy change, evidence translation) post-session than pre-, indicating the method’s ability to achieve knowledge transfer. Participants recommended making policies less punitive through modifications rather than total removal; indicating they may see these changes as more amenable. This included changing enforcement and funding mechanisms to dis-incentivize harmful or nonevidence-based practices. In cases where individuals are connected to treatment, participants had mixed views as to whether mandated reporting and testing policies are inherently punitive. Participation resulted in non-significant increases in de-implementation buy-in along belief, motivation (NC) and self-efficacy (KY) domains.

Conclusions

This paper advances PSD as an emerging methodological approach in implementation science as the first to use it to understand policy IS and de-implementation. Practitioners viewed de-implementation of punitive policies as complex and influenced by political context and role to implementation. Practitioners may hold nuanced views and prioritize modifications to certain policies (e.g., through enforcement or enabling of treatment exceptions) rather than full dismantling. Future de-implementation efforts should integrate systems-based, practitioner-engaged methods to foster buy-in and better conceptualize de-implementation within IS. Lack of practitioner consensus or buy-in may otherwise thwart de-implementation efforts.

Contributions to the Literature

- This study advances participatory system dynamics, newer to implementation science, as a method to understand changed practitioners’ perspectives of policy de-implementation.
- We introduce novel methodological components to PSD, including micro-interlocutor analysis, modeling as persuasion, and model acceptability to evaluate PSD-driven information transfer and consensus.
- Practitioners may hold nuanced views to punitive policy de-implementation that differ by role, political context, and proximity to implementation, prioritizing modifications to, rather than full dismantling of, certain policies.
- We provide a replicable framework for an effective implementation science approach to understand policy implementation dynamics, increase buy-in, and develop consensus on shared problem definitions (de-implementation).

Introduction

The field of policy implementation science needs methods to understand how practitioners view causes of complex problems and how these dynamics change over time. (1–4) Alone, implementation strategies often don’t document “how” determinants account for temporality of determinants, or how changes to policymaker perspectives over time might (feedback) impact implementation.(1, 3, 5) Participatory system dynamics (PSD), a systems science method that integrates temporal, feedback, and cyclical understanding of complex problems, may better capture these determinants. However, even with a call to combine IS and systems science disciplines (3, 6, 7), few projects outline in detail exactly how PSD projects are applied.(8, 9) This applied use of PSD to understand factors affecting de-implementation of maternal substance use policies addresses the need in both fields for methodological clarity and increased replicability.

Participatory system dynamics (PSD)—a system dynamics (SD) method for developing shared understanding of complex problems—may be an apt implementation approach to address factors impacting policy de-implementation.(10) PSD guides participants in the co-creation of a causal loop diagram on factors affecting the topic.(11) It unpacks the trends, structures, and assumptions that reinforce views and has been shown to develop shared understanding (e.g., mental model shifts), knowledge transfer, and needed buy-in for action.(7, 12–15) While increased buy-in is an oft-cited determinant of PSD and implementation success, it is infrequently measured quantitatively or theoretically or quantitatively.(16–19) While motivation and belief are known precursors

to de-implementation, and a buy-in constructs, few studies explicit provide measures of these as facilitators of de-implementation.(20–23) Consensus is another primary aim of IS and PSD that remains under-defined in the field, needing more robust methods.(24–26) While PSD has been used in implementation settings, it has not been used widely to interrogate policy adoption or de-implementation.(6)

The practice of de-implementation—removing nonevidence-based or harmful practices—is emerging in implementation science (IS) and remains poorly understood in the realm of health policy.(27, 28) Most literature investigates phasing out low-value care in clinical settings rather than removing harmful practices.(29, 30) However, this prioritization has implications for the field of substance use treatment, where historical reification of punitive policies toward pregnant persons of color might make them more recalcitrant to change.(20, 30–34) Acknowledging this, de-implementation research documents the efficacy of partner-driven, bottom-up approaches, and politically-backed de-implementation.(19, 30, 35, 36) However, how practitioners themselves define these structural determinants of de-implementation and factors that reinforce policies remain unclear.(29, 30, 35–37) Further, reviews of de-implementation frameworks name a lack of methods to measure “complex relationships”, “policy feedback” and “incrementalism” which are directly addressed through feedback and temporality aims of PSD can provide.(20) More work needs to test the transferability of implementation science frameworks and their needed modifications or relevance for de-implementation.(20)

This study provides a reproducible approach (PSD) for adding needed temporal and cyclical understanding of de-implementation, while also creating buy-in and consensus for needed policy change. In this methodological paper, we describe PSD’s ability to 1) accurately model factors affecting policy de-implementation for punitive maternal substance use, and create measurable changes in participant 2) definitions of, 3) consensus on, and 4) buy-in towards policy. Because policy de-implementation requires collective action, we report outcomes using the Theoretical Domains Framework (TDF).

Methods

We provide replicable components of a PSD application and evaluate its ability as a method to measure implementation constructs and facilitate knowledge transfer. For each state, two, two-hour systems modeling sessions were conducted in November and January of 2025 with pre-/and post-Qualtrics surveys to assess changes to de-implementation actions, policy consensus, and participant buy-in (Supplemental file 2). All methods are reported in alignment with COREQ reporting guidelines.(38)

Recruitment

Participants were sourced from North Carolina (NC) and Kentucky (KY) to compare different political contexts. North Carolina is historically a non-Medicaid expansion state with no punitive policies in place (Fig. 4). Kentucky is a Medicaid-expansion state with a more mixed landscape with mandated reporting and testing policies in place alongside priority treatment and discrimination protections. The study’s purposive, snowball sampling design targeted 8 to 10 practitioners per state along the child welfare reporting chain (child welfare and substance use agency staff, clinic leadership, clinicians, nurses, social workers, and advocates). Participants were recruited via email for a study on “policy change for maternal substance use policies”, so as not to prime for de-implementation. The final, bistate sample of 21 individuals met the criteria for qualitative data saturation and for PSD research, as more than 10 individuals in a session is not manageable (Table 1).(39, 40) Eleven of the 15 individuals who expressed interest in North Carolina and 10 of the 17 in Kentucky were able to attend the final, scheduled sessions. Participants were compensated \$125 upon completion of the sessions and pre-/post- surveys. All activities were deemed exempt by the [REDACTED] Institutional Review Board (IRB #24-1587).

Table 1
Final Participant List and Roles by State

North Carolina	Kentucky
SUD Agency Staff Member	Title V Agency Staff Member
County-level Child Welfare Staff Member	Clinic Director 1
Clinic Director 1	Clinic Director 2
Clinic Director / MD Provider 1	Clinical MD Provider 1
Clinic Director / MD Provider 2	Inpatient Social Worker 1
Clinic Director / non-MD Provider 1	Inpatient Social Worker 2
Inpatient Social Worker 1	Nurse Practitioner
Inpatient Social Worker 2	Peer Recovery Coach 1
Outpatient Social Worker	Peer Recovery Coach 2
Licensed Practical Nurse	Policy Advocate
Policy Advocate	

Modeling Sessions

Two trained, female-identifying facilitators led each session with PSD scripts from the standardized field repository (Supplemental file 1)(Scriptapedia).(14) One was the first-author doctoral student with experience in mixed methods and formal PSD training from the Systems Science for Social Impact (SSSI). The

other was a PhD-level faculty member experienced in systems research, modeling, and facilitation. Though known individuals were contacted for recruitment, there were no prior relationships between participants and facilitators.

There are three core stages of PSD; reference mode and variable elicitation, connection circles, and causal loop diagrams, designed to foster cyclical, temporal, and interdependent thinking (Fig. 1).(7, 14, 41) The first session included reference mode building, variable elicitation, and connection circles as well as a primer in systems thinking to establish shared language. The “reference mode”, or key outcome trend, can be used to remind the group of the primary outcome of interest and guide future PSD stages.

In variable elicitation, individuals identify factors (variables) influencing the reference mode trends by describing causes increases or decreases to the projected “extent to which policies are punitive / rehabilitative” over time. These were grouped by a second facilitator who shared the thematic results at the end of the activity. Participants then anonymously “voted” (by placing virtual dots) on variables they felt were “most influential” on the reference mode trends; to reduce participation bias and take the most-validated (e.g., most votes) variables to the next stage of modeling.(5) Connection circles further explore relationships between key variables, highlighting influential causal connections and allow participants to start identifying variable inter-connections and feedback loops.(14)

Prior to the second session, the primary facilitator created “seed diagrams”, or thematic, initial causal loop diagrams based on the variable elicitation, connection circles, and session recording. The second two-hour session focused on refining the “seed diagrams” to synthesize into a visual representation of the system. The modeler then clarifies the model after the session by sending to participants for review. Each step builds on the previous one to foster consensus and understanding of feedback mechanisms, policy implications, and leverage points and create a comprehensive systems model.(43)

Outcomes

In this paper, we report on the process-level outcomes of a participatory modeling process used to understand policy de-implementation. We used the several domains of the Theoretical Domains Framework (TDF) to create targeted outcomes, an integrative framework across psychological and organizational domains of behavior change across that has recently been used for de-implementation projects.(Table 6)(20, 44, 45) The three-question pre-survey included questions on 1) needed action items for de-implementation, 2) views toward punitive and rehabilitative policy categorizations (consensus), and 3) a validated, four-item buy-in Likert scale.(16) The post-survey included questions (2) and (3) and a model acceptability score. There was a 100% (11/11; NC) and 90% (9/10; KY) response rate for the post-surveys. Participation in both the sessions and pre-survey was the eligibility criteria for inclusion in the final analyses. For Kentucky, one individual did not complete the post-survey. Their pre-survey responses were kept for policy consensus but removed for buy-in and qualitative analysis. One NC participant only completed the post-survey and was entirely removed for a final participating sample of 10 (NC) and 10 (KY) for the consensus and 10 (NC) and 9 (KY) for the buy-in and qualitative analyses.

Model accuracy

Model acceptability. Measuring model accuracy (e.g., model accurately represented participant understanding of the system) was a primary study aim. This was measured through the post-survey question, “*Do you accept that the finalized model reflects your understanding of the system?*” with Likert responses: 1= “I cannot accept the model diagram”, 2= “I need revisions before I can accept the model diagram”, 3= “I can accept the model diagram, but I have some reservations”, 4= “I am mostly satisfied with the model diagram, I can support it”, and 5= “I am very satisfied with the model diagram and fully support it”.(12, 13, 57, 88) For ratings under 4, participants were prompted to list desired changes. The micro-interlocutor analysis also served as an accuracy check—ensuring that the most “agreed” upon concepts were present in the final model.

Qualitative analysis of action items

In both the pre- and post-survey, participants were asked “*If we wanted to de-implement punitive policies, what are the most important action items?*” We qualitatively coded pre-/post- changes in participant action items using inductive, content analysis to understand how they conceptualized “de-implementation” based on perceived actions.(46) This approach included open coding whereby codes are noted as they emerge and later analyzed for interrelationships through selective coding.(46, 47) Codes were developed from the North Carolina pre-survey by two graduate-level qualitative researchers who independently open-coded the short-answer tasks and then later compared, reconciled, and collapsed codes into a final, 10-item codebook (Table 2). For example, action items identifying a need to “fund systems-level work” and “more funding for treatment programs to keep babies with the mother” were coded as funding. Some action items were double-coded if they met multiple criteria. This codebook was then applied to the North Carolina post-survey and Kentucky pre- and post-survey with no new codes emerging (e.g., “code meaning”), indicating qualitative saturation had been reached.(39, 40)

Table 2
Action Items Codebook and Illustrative Quotes

Code	Illustrative Quote
Advocacy and campaigning	Discussion of advocacy efforts at any level. "Meeting with Legislators"
Removing barriers / increase access	Discussion of removing barriers to care or access. "Reduce barriers to participation in recovery programs for pregnant and parenting people"
Evidence-to-policy pipeline	Integration of evidence and data into political processes. "Subject matter experts need to educate lawmakers"
Funding	References to increased funding or funding mechanisms. "NC and federal legislatures need to adequately fund specialized treatment programs"
Education and training	Discussion of education or credentialing of staff or leadership at any level. "Leadership needs to be re-trained and educated on alternative approaches to addressing maladaptive behaviors"
Less punitive	References to making extant policies less punitive: "Lawmakers need to change policies to decriminalize SUD with pregnant and parenting people"
More rehabilitative	References to making extant policies more rehabilitative: "Policymakers need to support laws that are more supportive of recovery efforts"
Other policies (individual)	Policies addressing individual-level barriers. "More childcare vouchers", "state provides more transportation options"
Other policies (structural)	Policies addressing structural or political determinants. "Remove any barriers to Medicaid access for all health care including abortion"
Stigma reduction	Discussion of changing attitudes or views toward substance use disorders. "Major campaign about how addiction is a disease and use reward for treatment versus punishment for disease."

A well-documented outcome of PSD involvement is knowledge transfer, so we used a method known as "modeling as persuasion" (e.g., the exchange of ideas through the course of the session as a result of exposure to other participant ideas) which can show idea transfer as a result of the sessions.(48, 49, 49–54) The codes were quantified by the total number of codes that emerged for pre- and post-, "retained" codes that emerged in both the pre- and post-survey for that same individual, and "transferred" codes (e.g., codes that emerged in the pre-survey of one individual and the post- survey of another individual) (Table 3).(53, 55)

Consensus

Policy Categorizations. Before assessing buy-in and action items for de-implementation, understanding how these common categorizations of policies resonate with practitioners provides context (Fig. 4; Fig. 5). We asked "Which of these state-level policies do you agree should be categorized as punitive or rehabilitative?" in the pre-survey for four main "punitive" policies and seven "rehabilitative" policies from the literature.(56–59) (Fig. 4). Participants could respond, "Agree", "Disagree", and "Not Sure" and prompted to explain disagreements. Four KY participants, disagreed to both policy 2 and policy 1, but qualitative responses seem to indicate disagreement with the policy itself rather than the categorization and were changed to "agree". For Fig. 5 on differences by role, we excluded "unsure" responses to create percentages.

Micro-interlocutor analysis. Measuring model consensus is a primary aim of PSD. To do so, we used micro-interlocutor analysis, a qualitative method for quantifying and assessing consensus in focus groups (Table 5). It quantifies participant involvement, response order, agreement or disagreement, and thus granular conversational patterns of a session, otherwise often reported in aggregate.(60, 61) For each theme that emerges, we documented participant agreements, disagreements, or no response—and compared the frequency of these responses to other themes.

Buy-In

Our study uses a buy-in measure that integrates across several domains (belief, motivation, self-efficacy, and commitment).(16, 62) A validated self-report, four-item Likert measure (Strongly Disagree, Disagree, Agree, or Strongly Agree) of participant buy-in was modified for "de-implementation" and "pregnant and parenting persons who use substances" language in the pre- and post-survey.(62) This included the following questions on 1) belief: "*I believe that punitive policy de-implementation is good for pregnant and parenting persons who use substances,*" 2) motivation: "*I am personally motivated to de-implement...*," 3) self-efficacy: "*I am capable of making the kind of changes needed to de-implement...*," and 4) commitment: "*I will take specific actions to support...*" There was also a "Not Sure" option. These occurred infrequently (four times across pre-/post- for NC and three times across both for KY) and are not factored numerically in the results scores. Wilcoxon signed-rank tests accounting for the paired, non-parametric distributions of the small sample size assessed changes in mean pre-/post- buy-in.

Results

Model

The full models for each state will be presented in an independent paper. We contextualize the evaluative results with a high-level causal loop diagram (CLD) with core variables from both states (Fig. 2). Fundamentally, participants felt punitive policy de-implementation requires countering evidence of how they limit medications for opioid use disorder (MOUD) and treatment access. They felt that translating these clinical contributions were most within their locus of

control (evidence-to-policy translation). Positive MOUD outcomes that promote policy change via elected official support are reinforced or threatened by larger political systems (policy / funding). This includes adjacent child welfare policies (such as treatment contingencies) that may reduce the likelihood of recovery and infant reunification (child welfare). These outcomes are strongly influenced by the presence of wraparound services. Services availability is dictated by funding and legislative priorities, (e.g. housing policy, residential treatment capacity limiting co-located care).

Model acceptability. The mean “acceptability” of the model after translation was 4.2 (out of 5) for NC and 4.75 for KY, with one request to the NC model and no requests for the KY model, indicating final models were accepted by participants.

Notes

RED arrow—variables change in opposite directions (as one variable increases, the other decreases); **GREEN** arrow—variables change in the same direction (as one variable increases or decreases, the other does the same)(63) For example, as more evidence is provided the MOUD works, public and provider stigma decreases (Micro-theme 6). *FFPSA – Family First Prevention Services Act; SOR – State Opioid Response grants; CARA – Comprehensive Addiction and Recovery Act (legislated SOR funding);* **BLACK BOX**—overlap with micro-Interlocutor analysis themes.

Qualitative analysis of action items

The list of top codes in the pre- and post-survey are listed in Fig. 3. Notably for both states, structural codes such as “funding”, making policies “more rehabilitative” (NC), and “evidence-to-policy pipeline” (KY) increased into the top three post-survey codes. This was seen alongside a decrease in action items targeted at individuals such as “education and training” and “advocacy” for North Carolina. For Kentucky, the top codes, making policies “less punitive”, “more rehabilitative”, and funding remained in the top three from pre- to post-survey.

Note

n values represent the total number of listed “action items” across included participants (9 Kentucky participants; 10 North Carolina participants), rather than the number of participants

In the analysis of retained and transferred codes, North Carolina participants had slightly fewer average “transfer” codes (e.g., they did not originally list the code in the pre-survey) in the post-survey (1.4) than retained codes (1.7) (Table 3). Kentucky had more transfer codes (2.2) than retained codes (1.2), indicating a high degree of knowledge sharing. Both states had more action items to make policies “less punitive” in the post-survey, indicating increased participant understanding of these mechanisms, a key motivation of the study aims.

Table 3
Modeling as Persuasion: Action Items Retained and Persuaded by Others (51)

North Carolina	Retained	Transferred	Transferred Codes
Participant 1	2	0	
Participant 2	2	1	Other policies (structural) ^a
Participant 3	0	2	Evidence-to-policy pipeline, stigma reduction
Participant 4	0	3	Less punitive, ^a more rehabilitative, other policies (structural) ^a
Participant 5	2	1	Funding ^a
Participant 6	1	0	
Participant 7	1	2	Funding, ^a less punitive ^a
Participant 8	4	1	Hospital-level policy
Participant 9	4	2	Education & training, funding ^a
Participant 10	1	2	Stigma reduction, more rehabilitative
Average	1.70	1.40	
Median	1.50	1.50	
Range	0–4	0–3	
Kentucky	Retained	Transferred	Transferred Codes
Participant 1	2	2	Evidence-to-policy pipeline, ^a advocacy, less punitive ^a
Participant 2	2	1	Evidence-to-policy pipeline ^a
Participant 3	0	4	Funding, less punitive, ^a remove barriers, more rehabilitative ^a
Participant 4	1	3	Stigma reduction, education & training, more rehabilitative ^a
Participant 5	2	1	More rehabilitative ^a
Participant 6	1	2	Advocacy & campaigning, funding
Participant 7	1	3	Education & training, funding, evidence-to-policy pipeline ^a
Participant 8	1	1	Less punitive ^a
Participant 9	1	3	Education & training, less punitive, ^a evidence-to-policy pipeline
Average	1.13	2.25	
Median	1	2	
Range	0–2	1–4	

^atop transfer code between pre- and post-survey.

Notes. Retained: Present in the participant's Pre- and Post- survey response themes, Transferred: theme that appeared in different participant pre-survey responses and appeared for the first time in the participant's post-response.

Broader political determinants of policy implementation

The qualitative findings showed a heightened post-participation understanding of broader political influences (e.g., evidence, funding, and political actors) of policy de-implementation. This was evidenced in Kentucky, where the "evidence-to-policy pipeline" (n = 4) code was the most common transfer code and responses had added specificity in the post-survey (Table 3). In the pre-survey these vaguely focused on a need for centers to "prioritize" certain treatments and local agencies to be "aware of" certain resources for referrals. In the post-survey, these codes specifically mentioned data sources and outcomes for evidence translation such as "cost savings" to inform funding decisions such as "understand(ing) the needs (needs assessment)... to decide where to distribute funds" and having lawmakers "evaluate alternatives to sentencing and what data supports" successful treatment connections.

"Funding" was the most common transfer code for North Carolina (n = 3) and had added specificity between pre- and post-surveys (Table 3). In the pre-survey (NC) codes were short and vague, such as "more funding" and "state [needs] more funding." However, in the post-survey, responses were more in-depth and more macro- in nature. This included funding on multiple levels, from "funding for system-level work", "state funding should prioritize recovery housing and residential treatment", and even federal "funders (such as CDC) need to create grants" that address dual mental health and substance use needs.

Deepened socio-political understanding extended to within-participant responses. For example, one participant listed two “education and training” tasks targeted at “adolescents” and “providers” in the pre-survey, replaced with stigma education for “public health” and “state officials” in the post-survey. This shift from provider-based to policymaker-, and public health official-based stigma reduction training indicates a more, macro-level prioritization and shift in and understanding of the major players involved in policy de-implementation.

De-implementation and policy change

Understanding change in participant conceptualizations of de-implementation was a primary study aim. Mechanisms to make policies less punitive were simplistic in the pre-survey for both states—often without specifically referencing a policy. This was seen in comments to “stop criminalizing substance use” (NC) or a need to “assess certain policies” (KY). In the post-survey, responses more explicitly targeted the process of who and how. For example, North Carolina individuals named specific reimbursement processes as an enforcement mechanism for de-implementation such as “mandat[ing] rehabilitative policies for recovery housing programs that receive federal funding” and changing specific state policies such as “Medicaid service definitions” to more explicitly reimburse only evidence-based practices and eliminate “state housing waiver penalties (e.g., barring housing applications) if a person is actively engaged in treatment.”

Modifying policies to be “less punitive” as de-implementation. For both states, participants conceptualized de-implementation with more references to policy implementation and modifications (over removal) within a given statute in the post- compared to pre-survey. This included language on “enforcement”, “uniform policies”, “standardized approaches”, and changes to “specific language” within policies (e.g., reduce “variations by county”). This attention to policy implementation was particularly salient among Kentucky participants for treatment exceptions (e.g., exceptions from punitive statutes for substances taken as prescribed or for treatment); it was the most-referenced, specific, “less-punitive” policy (n = 3) across pre- and post-surveys. For example, there was a call for CPS to “look at hospital records and not penalize mothers with positive UDS [urine drug screen] who received treatments that could make UDS positive (e.g., MOUDs),” a direct reference to changing policy implementation as hospital policy institutionalization is a necessary precursor to substantiated reporting procedures. This was again seen in a call to change policies that “punish pregnant persons...actively seeking treatment” or “testing policies to reflect prescribed medications as non-punitive.” The only overt, active calls for policy change—that may be interpreted as de-implementation through removal—were to decriminalize punitive “definition” policies (n = 5). For example, “policies to decriminalize SUD with pregnant and parenting people and offer more rehabilitative solutions” (KY) was dually coded “more rehabilitative” and “stop criminalizing usage during pregnancy” was coded as “less punitive” but both make a call for decriminalization and every time mentioned it called for removal or explicit change. Interestingly, neither state has these policies in place, indicating strong beliefs on these policies in other states and nationally, or mis-implementation of these policies within the state.

Enacting policies to be “more rehabilitative” as de-implementation. Responses on needed actions to make policies more rehabilitative—were cited more frequently than changing punitive policies directly (e.g., the “less punitive” code). These emerged at about equivalent frequency pre- and post-session for North Carolina (n = 5; n = 6) and Kentucky (n = 6; n = 6), respectively. For both states, rehabilitative tasks were more likely to discuss active policy “changes” in the post-survey rather than “modifications” to existing policies—in a reversal of the above trend for “less punitive” codes. This was seen with language such as a need to “stop practice of ‘safe-keeping’ in state prisons” and state “mandat[ing of] evidence-based policies for recovery housing.” This was compared to language on a need to “clarify” notifications versus reports, “increase access to START teams... [within] reports”, and “loosen guidelines” on housing restrictions for this population in the pre-survey.

For Kentucky, making policies more rehabilitative was a major theme given its more punitive policy landscape. This included more specificity in funding and increased citation of co-locating mother and child needs (including childcare and housing). This practice increases treatment retention, often needed for mother-child reunification. Specifically, funding mechanisms for this code were cited more frequently in the post-survey. For example, one individual specifically mentioned “state-funded early intervention... during the first trimester” while another mentioned a need for more “state and federal funding for [co-located] treatment options.” Within the “more rehabilitative” codes, the importance of co-locating care for mother and child prevailed, making up 50% (6/12) of the rehabilitative codes across pre-and post-session. This is significant as it only emerged twice in the NC codes and once in the KY pre-codes. This included recommendations for more “rehab that parent and child can go to” and “more attempts [by CPS (child protective services)] at rehabilitation that include children versus excluding them.” While not specific to a policy, it does address context around policy implementation to be more rehabilitative by enabling chances of care retention and reducing relapse, which would otherwise activate punitive structures.

Consensus

Rehabilitative versus punitive policies. We ascertained pre-session agreement with common categorizations of punitive and rehabilitative policies. For North Carolina and Kentucky there was relatively consistent consensus on policy 1 (definition) & policy 2 (no treatment exceptions) as punitive (Fig. 4). There was mixed consensus, as indicated by a higher percentage of “disagree” and “unsure” selections on policy 3 (required testing) and policy 4 (mandated reporting). Interestingly, Kentucky participants had slightly more mixed views (e.g., 30% “disagree”) on required testing whereas North Carolina had more mixed (e.g., 30% “disagree”) views on mandated reporting. This was reflected in the qualitative responses where practitioners felt these policies were only truly “punitive” if not connected to treatment,

Infants with prenatal exposure to substances may need longer observation in the hospital or treatment for withdrawal symptoms after birth. It is imperative to know the substance of exposure to develop a treatment plan. Using that information simply to remove the infant from parental custody or as a threat of removal is punitive.

(Nurse Practitioner, Kentucky)

For both states, there was consistent agreement with almost all rehabilitative categorizations (Fig. 4). Only North Carolina had mixed views regarding policy 4 (required notifications for suspected substance use), the rehabilitative equivalent of mandated reporting, with 50% indicating “unsure”. However, the

qualitative responses seem to indicate that these participants may disagree with this categorization because they still see a notification-based system as punitive, where “there should be no need for reporting at all” and “requiring notification insinuates that there is confirmed substance use... that perpetuates stigma.”

These mixed views for punitive policy categorizations followed trends by role (Fig. 5). For both states, child welfare participants had a higher percentage (~50%) of their responses as “disagree” with categorization across all policies. This was also true for social work practitioners, for whom about one-third disagreed with these categorizations across states. Providers and clinic directors overwhelmingly agreed with the categorizations.

Note

Percentage was calculated as total number of indicated disagreements across all four possible punitive policy categorizations, divided by the total number of possible answers. For example, clinical provider participants in North Carolina (n = 4), five disagreements were marked across 15 total responses (n = 5/15, 33%)(excluding 1 “unsure”).

Micro-interlocutor analysis. The results of the micro-interlocutor analysis are listed in (Table 4) and contributions that garnered higher than median number of agreements corresponded to major themes in Fig. 2. Major themes for North Carolina (Table 4) corresponded directly with variables 1, 2, and 4 in the main feedback loop structure. For Kentucky, top themes corresponded with variables 3, 5, and 6 in the main feedback loop structures of the final models. Notably, wraparound services were predominantly featured for both states. This alignment helps validate several of the primary themes, intervening variables, and feedback loops for the states’ final models with no recorded disagreements. Full analysis is delineated in the model paper; here, we demonstrate this method as a novel validation tool for qualitative PSD model-building.

Table 4
Micro-interlocutor Analysis

NORTH CAROLINA	DSS perspective change over time	Rates of Maternal Overdose	Strong state leadership influence	County-level interpretation	Shifting substance landscape	Siloed MH /SUD / polysubstance	Public & provider stigma training	Time-delay in building hospital trust (S2)	Wraparound services (S2)
SUD Agency Staff Member	SE (S2)	–	A	SE	–	–	SE	–	A
Outpatient Social Worker	SE	0	A	–	–	0	A	–	–
Clinic Director 1	SE	–	0	–	0	SE	SE	A	SE
Clinic Director / MD Provider 1	A	–	A	SE (S2)	–	–	SE (S2)	–	–
Clinic Director / MD Provider 2	A	A	–	A	–	–	SE	SE	SE
Clinic Director / non-MD Provider 1	–	–	–	SE	–	–	SE (S2)	–	0
Inpatient Social Worker 1	0	–	–	A	–	–	SE (S2)	–	–
Inpatient Social Worker 2	–	–	–	–	–	–	–	–	SE
Licensed Practical Nurse (LPN)	–	–	A	–	–	–	SE	–	–
County-level child welfare staff	SE	A	–	0	A	–	A	–	–
Policy Advocate	SE (S2)	–	–	–	–	–	0	0	A
Count	7	2	4	5	1	1	9	2	5
KENTUCKY	Significant programing growth (last 5 years)	Medicaid expansion	Media attention	Family First Prevention Services Act (FFPSA)	Housing crisis putting pressure (S2)	Sustainability (S2)	Comorbid MH Tx integration (S2)	Residential treatment restrictions / Funding (S2)	Recovery Ready Communities designation
Title V Agency Staff member	SE	–	A (S2)	–	–	–	SE	SE	–
Nurse Practitioner	–	A (S2)	SE (S2)	–	–	–	–	–	–
Clinic Agency Director 1	A	0	–	SE	SE	0	SE	0	SE
Clinic Agency Director 2	–	–	A (S2)	A	SE	SE	0	SE	A
Clinical MD Provider 1	–	SE	–	–	–	–	–	–	–
Inpatient Social Worker 1	–	–	–	–	A	SE	A	A	–
Inpatient Social	0	–	–	0	–	–	–	A	–

NORTH CAROLINA	DSS perspective change over time	Rates of Maternal Overdose	Strong state leadership influence	County-level interpretation	Shifting substance landscape	Siloed MH /SUD / polysubstance	Public & provider stigma training	Time-delay in building hospital trust (S2)	Wraparound services (S2)
Worker 2									
Peer Recovery Coach 1	–	–	SE	–	O	–	–	SE	A
Peer Recovery Coach 2	–	–	–	–	–	–	–	A	–
Policy Advocate	–	A	O	–	–	–	–	A	O
Count	2	3	4	2	3	2	3	7	3
Notes. BOLD = higher than 4 agreements (NC) and higher than 2 agreements (KY) and qualified as major theme. S2 = theme emerged in session 2. O = Originator									
A = Agreement (verbal or non-verbal). D = Disagreement (verbal or non-verbal). SE = Provided significant statement or example suggesting agreement.									
SD = Provided significant statement or example suggesting dissent. – = Did not indicate agreement or dissent (non-response)									

Buy-in

The buy-in results were analyzed from completed pre-and post-surveys for NC (n = 10) and KY (n = 9). For North Carolina, mean scores increased non-significantly from pre- to post- for all buy-in domains; notably for belief and motivation domains and no “Strongly Disagree” or “Disagree” in the post-survey indicating possibly higher concordant buy-in (Table 5). For Kentucky, mean scores increased non-significantly; notably belief and self-efficacy, while decreasing for commitment domains (Table 5). Kentucky had lower post-session buy-in scores than North Carolina, with the most notable increases in the belief domain and lowest scores in self-efficacy (though notable increases) for both state contexts. This shifted from 60% of participants selecting “Strongly Agree” to 80% in the North Carolina post-survey, and an increase from 56% to 78% selecting “Strongly Agree” for Kentucky. Kentucky saw a strong post-session increase in the “self-efficacy” domain with an increase of 33% to 78% of individuals selecting “Agree”, driving the change of “agreement” (Strongly Agree + Agree) from 44% to 89% (Table 5). The self-efficacy scores were the lowest of the domains for both states.

Table 5
Buy-in Results by State

Domain	Indicator	Pre	Post	Change
North Carolina (n = 10)				
Belief	Mean	3.30	3.80	(+ 0.50)
	Range	1–4	3–4(b)	
	% Agreement(a)	80%	100%	20%
Motivation	Mean	3.50	3.80	(+ 0.30)
	Range	1–4	3–4(b)	
	% Agreement(a)	90%	100%	10%
Self-efficacy	Mean	3.13	3.22	(+ 0.09)
	Range	1–4	3–4(b)	
	% Agreement(a)	70%	100%	30%
Commitment	Mean	3.44	3.50	(+ 0.06)
	Range	1–4	3–4(b)	
	% Agreement(a)	90%	100%	10%
Composite	Mean	3.35	3.60	(+ 0.15)
Kentucky (n = 9)				
Belief	Mean	3.33	3.56	(+ 0.23)
	Range	1–4	1–4	
	% Agreement(a)	89%	89%	0%
Motivation	Mean	3.56	3.56	--
	Range	2–4	3–4(b)	
	% Agreement(a)	89%	100%	11%
Self-efficacy	Mean	2.71	3.00	(+ 0.29)
	Range	2–4	2–3	
	% Agreement(a)	44%	89%	45%
Commitment	Mean	3.50	3.33	(-0.17)
	Range	3–4	3–4	
	% Agreement(a)	89%	100%	11%
Composite	Mean	3.31	3.36	(+ 0.05)
Likert 1–4 (Strongly Disagree-Disagree-Agree-Strongly Agree)				
^a High Agreement = Agree + Strongly Agree				
^b Indicates no "Disagreements" appearing in the post-survey when present in the pre-survey				

Outcomes alignment with Theoretical Domains Framework (TDF)(44, 45)

We aligned study outcomes *experienced* by participants, but not action item themes, to the Theoretical Domains Framework (TDF) “sources of behavior” (capability, opportunity, and motivation). This revealed different groupings of subdomains and domains by outcome (Table 6). The micro-interlocutor analyses exclusively measured “social influences” within the opportunity source. The process of defining de-implementation through action items appeared primarily within the motivational sources, with these activities targeting five “goals” and one “intentions” subdomain. Notably, it also targeted at least one subdomain of all three capability sources. The majority of subdomains measuring policy consensus and buy-in fell within motivational aspects of behaviors in the “beliefs about capabilities”, “optimism” and “beliefs about consequences” domains.

Table 6
Study outcomes by Theoretical Domains Framework (TDF) domain

Domain	Construct	Micro-interlocutor Analysis	De-implementation action items	Policy consensus	Buy-In
CAPABILITY					
Knowledge	Knowledge		RAIC	APR	
	Knowledge of task environment		RAIC	APR	
Skills	Competence				BISE
	Ability				BISE
	Interpersonal skills		Transfer; MAP		
Behavioral Regulation	Action planning		DDIA		
MOTIVATION					
Professional Role and Identity	Professional role			DCR	
Beliefs about Capabilities	Self-confidence				
	Self-efficacy				BISE
	Perceived behavioral control				BISE
	Empowerment				BISE
	Professional confidence				BISE
Optimism	Optimism			DCS	BIM; DBS
	Pessimism			DCS	
Beliefs about Consequences	Beliefs			APR	BIB
	Outcome expectancies			APR; DCS; DCR	BIB
	Characteristics of outcome expectancies			APR	BIB
	Consequents			APR	BIB
Reinforcement	Rewards				
	Consequents			DCS	
	Contingencies			DCS	
Intentions	Stability of intentions				BIC
Goals	Goal priority		DDIA		
	Goal / target setting		DDIA		
	Action planning		DDIA		BIC
	Implementation intention		DDIA		BIC
OPPORTUNITY					
Environmental Context and Resources	Organizational culture / climate			DCS	
<p>APR = Agreement of policy and punitive or rehabilitative; BIB = Buy-In "belief" domain; Buy-in "commitment" domain ("I will <i>take specific actions</i>..."); BIM = Buy-In "motivation" domain; BISE = Buy-In "self-efficacy" domain ("I am <i>capable</i> of making changes..."); DBR = Differences in buy-in by role; DBS = Differences in buy-in by state; DCR = Differences in policy consensus by role; DCS = Differences in policy consensus by state; DDIA = Defining DI Actions; MAP = Modeling as persuasion; MI = Micro-interlocutor analysis; RAIC = Retained action item codes; Transfer = transfer codes</p> <p>Notes: The domains measures process-outcomes, not de-implementation themes. They did not make use of the following TDF domains and sub domains: Behavioral Regulation (2): Self-monitoring, breaking habit; Beliefs about capabilities (4): Self-confidence, perceived competence, beliefs, self-esteem; Beliefs about consequences (1): Anticipated regret; Emotion (all); Environmental contexts and resources (3): Environmental stressors, resources / material resources, salient events / critical incidents, person x environment interaction; Goals (2): Goals, Goals (autonomous / controlled); Intentions (2): Stages of change model, trans theoretical model and stages of change; Knowledge (1): Procedural knowledge; Memory, Attention and Decision Processes (all); Optimism (2): Unrealistic optimism, identity; Reinforcement (5): Rewards, Incentives, Punishment, Reinforcement, Sanctions; Skills (4): Skills, Skills development, Practice, Skill assessment; Social Influences (6): Social norms, social comparisons, social support, power, alienation, group identity; Professional role/identity (8): Prof. identity, social identity, identity, prof. boundaries, professional confidence, group identity, leadership, org. commitment</p>					

Domain	Construct	Micro-interlocutor Analysis	De-implementation action items	Policy consensus	Buy-In
Social influences	Barriers and facilitators		DDIA		
	Social pressure		MAP		
	Group conformity	MI			
	Group norms	MI			DBS; DBR
	Intergroup conflict	MI			
	Modelling		MAP		
Counts per domain		Micro-interlocutor Analysis	De-implementation action items	Policy consensus	Buy-In
CAPABILITY (9)					
Behavioral Regulation			1		
Knowledge			2	2	
Skills			2		2
Memory, Attention and Decision Processes					
MOTIVATION (27)					
Social/Professional Role and Identity				1	
Beliefs about Capabilities					4
Optimism				2	2
Beliefs about Consequences				5	4
Reinforcement				2	
Intentions					1
Goals			4		2
Emotion					
OPPORTUNITY (9)					
Environmental Context and Resources			1	1	
Social influences		3	2		2
APR = Agreement of policy and punitive or rehabilitative; BIB = Buy-In "belief" domain; Buy-in "commitment" domain ("I will <i>take specific actions</i> ..."); BIM = Buy-In "motivation" domain; BISE = Buy-In "self-efficacy" domain ("I am <i>capable</i> of making changes..."); DBR = Differences in buy-in by role; DBS = Differences in buy-in by state; DCR = Differences in policy consensus by role; DCS = Differences in policy consensus by state; DDIA = Defining DI Actions; MAP = Modeling as persuasion; MI = Micro-interlocutor analysis; RAIC = Retained action item codes; Transfer = transfer codes					
Notes: The domains measures process-outcomes, not de-implementation themes. They did not make use of the following TDF domains and sub domains: Behavioral Regulation (2): Self-monitoring, breaking habit; Beliefs about capabilities (4): Self-confidence, perceived competence, beliefs, self-esteem; Beliefs about consequences (1): Anticipated regret; Emotion (all); Environmental contexts and resources (3): Environmental stressors, resources / material resources, salient events / critical incidents, person x environment interaction; Goals (2): Goals, Goals (autonomous / controlled); Intentions (2): Stages of change model, trans theoretical model and stages of change; Knowledge (1): Procedural knowledge; Memory, Attention and Decision Processes (all); Optimism (2): Unrealistic optimism, identity; Reinforcement (5): Rewards, Incentives, Punishment, Reinforcement, Sanctions; Skills (4): Skills, Skills development, Practice, Skill assessment; Social Influences (6): Social norms, social comparisons, social support, power, alienation, group identity; Professional role/identity (8): Prof. identity, social identity, identity, prof. boundaries, professional confidence, group identity, leadership, org. commitment					

Discussion

In this study we evaluated participatory system dynamics' ability to measure changed participant definitions of, consensus on, and buy-in towards policy de-implementation. Practitioners viewed de-implementation as complex, politically-contingent, and influenced by proximity to implementation. This included prioritizing modifications to—rather than full dismantle of—certain policies. Partner buy-in from consistent and agreed-upon conceptualizations of de-implementation may be important determinants of punitive policy change.

Conceptualizations of de-implementation

We assessed practitioner changes to definitions of de-implementation through proposed action items and to measure the method's ability for knowledge transfer. Practitioners increasingly framed de-implementation within structural determinants (e.g., funding incentives, lack of wraparound rehabilitative services, and lack of evidence-informed policymaking), aligning with emerging literature on the economic, political, and social contexts of perceived de-implementation.(20, 64)

Practitioners primarily conceptualized de-implementation as changing elements of rather than complete removal of extant policies by more frequently citing making policies "more rehabilitative" than "less punitive" and significant "modification" rather than overt "removal" language. This included enforcement and funding mechanisms to disincentivize and effectively remove "harmful" nonevidence-based practices, indicating a possible heightened understanding of mechanistic processes from session participation. These findings align with several de-implementation systematic reviews where the majority or all articles described participants centering "reductions" of a practice as the most amenable de-implementation target outcome.(20, 21) However, even with this prioritization, the literature demonstrates that the presence of rehabilitative policies is insufficient to advance protective MOUD outcomes alongside punitive policies.(65, 66) Conversely, there was overwhelming support for de-implementation of policies that categorize maternal substance use as child abuse and neglect as seen in both the strong consensus of these policies as punitive (90% and 100% agreements in NC, and KY, respectively) and the only policy action item with overt de-criminalization and removal language. Because viewing de-implementation of a practice as "acceptable" is more likely to have needed support for action, future work should account for diverging and aligned perspectives on de-implementation.(67)

Overall, application of implementation science concepts and frameworks remain limited in their application to de-implementation. The "transfer" code findings also show heightened participant understanding of how the broader policy environment (e.g., evidence, funding, and political actors) affects policy de-implementation, aligning with studies citing de-implementation as driven more by higher-level determinants than implementation.(29, 30, 36) A systems-level perspective, such as through systems science methods, may be best suited to uncover these outer-context, structural factors and future research should contextualize de-implementation within these frameworks and theories to advance understanding.(15, 68–70) Additionally, there is an imperative need to center health equity and de-implementation within punitive maternal substance use policies research as historically marginalized communities are most likely to be subject to harmful practices.(31–34)

Consensus and Buy-In

In summary, an abundance of literature utilizes categorizations of punitive policies (56, 58, 71–74), but we don't yet know how and if these categorizations (e.g. policy consensus) resonate with practitioners and how that affects buy-in. Regarding policy consensus, a novel finding was discordance in practitioner views on "punitive" policies, with some not seeing mandated reporting and testing as punitive when they facilitate treatment connection. These "disagreements" varied by state and role, as social workers and child welfare staff—more directly involved in reporting processes—and participants in mixed policy environments demonstrate more mixed views when the policy is present in their state context (e.g., higher "disagreements" in Kentucky). While our study wasn't sufficiently powered to detect significance, notable takeaways exist. The lower "self-efficacy" score across states aligns with other findings.(16) Increased "belief" and "motivation" buy-in domains, coupled with making policies "less punitive" as the most "transferred" code, suggest heightened post-participation motivation for de-implementation. Kentucky's more "mixed" views of punitive policies and fewer listed action items might be reflected in its lower overall buy-in score. Motivation is a known precursor to de-implementation; and several emerging de-implementation frameworks, including TDF, center the importance of belief and intentions as facilitators of de-implementation.(20–23) Therefore, Kentucky may be in earlier stages of de-implementation, constrained by political feasibility of their extant punitive policies. North Carolina's higher motivation scores, alongside their continued reference to a strong pipeline of evidence-to-policy translation (e.g., top code in pre- and post- survey)—may reflect a higher degree of within-state agency. These findings indicate a need to better understand how practitioners view their own agency within a political system, referenced by several TDF domains. For example, a person can "believe" de-implementation is effective and needed but not motivated if the change feels intractable in context. While "belief" is a known predictor of implementation success and many implementation frameworks reference self-efficacy, these domains have yet to be integrated into our understanding of buy-in, especially in directive policy change.(16, 75–79) In summary, buy-in is often cited in PSD literature and as an IS determinant, but rarely measured quantitatively or with dimension.(16, 18) Integrating our measure with multidimensional constructs, aligning with extant frame such as the TDF, may improve understanding of policy implementation, and contend implementation science frameworks for relevance to de-implementation.

PSD as a method

PSD enabled accurate depictions of the temporal, interdependent, and cyclical nature of the systems impacting policy de-implementation, expanding its use beyond a decision-making tool to reveal policy adoption and repeal processes.(6, 80, 81) This study answers a call to—not only increase practical applications of systems modeling in implementation science—but in understanding policy; the political realities and contexts within which policymaking occurs, a post-positivist approach that can amplify policy implementation efforts.(69, 82–84) Walsh et al., (2022) also explicitly state in their review of de-implementation frameworks, that projects lack methods that account for more "complex relationships", "policy feedback" and "incrementalism" that feedback and temporality from PSD can provide.(20) Further, the majority of our study process outcomes also fell within the "motivation" domain of TDF (Table 6), indicating PSD may be adept at specifically targeting belief and intentions, known facilitators of de-implementation.(20–23) Further, more PSD studies need to systematically align with known de-implementation and implementation frameworks, such as ours did with TDF, in order to expand its relevance as an implementation science method.

This study contributes to methodological validation techniques within participatory system dynamics as the micro-interlocutor analysis, modeling as persuasion, cross-state comparisons, and buy-in score scores were all novel measures to align with implementation science aims.(83, 85) The major themes from the micro-interlocutor analysis mirrored the major feedback loop themes, indicating it may be a useful method for checking, comparing, and validating model themes. It may also help researchers prioritize the inclusion of variables with the most consensus into model building and improve model validity. No disagreements in the micro-interlocutor analysis may alternatively be due to shifting subjects or non-endorsement and power dynamics between roles— inclusion of "clinicians" and "clinic directors" within workplace hierarchies—may have prevented explicit disagreement.(86) Future micro-interlocutor analyses

should account for more nuanced types of interactions and their impact on group participation. This is also the first, known PSD model to include two state political contexts, with the purpose of enhancing generalizability concerns that often accompany these methods; delineating between context-dependent versus shared factors affecting state-level de-implementation.

Limitations

Purposive sampling may limit generalizability due to network-based selection bias, though this was likely to occur regardless of sampling strategy given the small community of perinatal substance use professionals. In the policy consensus questions, some (n = 4) individuals selected “disagree”, but their responses indicated they disagreed with the policy itself, not the categorizations and were reassigned as “agree”, but it may indicate a need for improved prompt wording. Self-reported buy-in may also be subject to socially desirability bias.(87) While the inclusion of two states increases external generalizability, the small sample size limits interpretation of the discordant policy consensus and buy-in results. Replication with greater sample sizes and with intentional sampling by participant role in the child welfare and SUD treatment systems is needed. Despite these limitations, the study had high retention and response rates, reducing non-response bias based on motivation or availability.(88) PSD as a method also has notable limitations for implementation science, notably its context-specific and qualitative nature make it difficult to assign direct causality and possible loss of constructs in translation of the model process.(85)

Conclusions

The study demonstrates PSD’s ability to measure the complexity of processes impacting de-implementation. Buy-in and role-based perspectives toward these changes may be influenced by political context and proximity to implementation. PSD as a practitioner-engaged method seemed to enhance participant knowledge on de-implementation and non-significantly increase buy-in. By integrating measures of consensus, buy-in, knowledge transfer, model acceptability, and implementation determinants, the paper offers a methodological framework that future implementation studies can adapt across diverse policy and practice settings.

Abbreviations

PSD
Participatory System Dynamics
SD
System Dynamics
CLD
Causal Loop Diagram
COREQ
Consolidated Criteria for Reporting Qualitative Research
SUD
Substance Use Disorder
MOUD
Medications for Opioid Use Disorder
MH
Mental Health
UDS
Urine Drug Screen
CPS
Child Protective Services
FFPSA
Family First Prevention Services Act
SOR
State Opioid Response (grants)
CARA
Comprehensive Addiction and Recovery Act

Declarations

Ethics approval and consent to participate:

All activities were deemed exempt by the University of North Carolina (UNC) Institutional Review Board (IRB #24-1587) and all procedures involving human participants were conducted in accordance with the ethical standards of the institutional committee and with the principles of the Declaration of Helsinki and its subsequent amendments. Informed consent was obtained from all participants prior to their participation. Participant confidentiality was maintained throughout the study, and all data were de-identified prior to analysis.

Consent for publication:

Not applicable.

Funding

This research was partially supported by a National Research Service Award Pre-Doctoral/Post-Doctoral Traineeship from the Agency for Healthcare Research and Quality sponsored by The Cecil G. Sheps Center for Health Services Research, The University of North Carolina at Chapel Hill, Grant No. T32-HS000032.

Author Contribution

LC conceived of, designed, collected, received funding for, and analyzed data for the study. KHL assisted in data collection, methods, interpretation, and manuscript review. SG, LJM, CG, and MS assisted in methods preparation and significant manuscript revisions. CT made significant contributions to methods review, data collection, interpretation, and manuscript revisions. All authors approved the manuscript.

Acknowledgement

The authors would like to thank individuals across the KY and NC perinatal substance use community who generously helped with recruitment. They are also immeasurably grateful to the community partners involved in the study that contributed their time, lived experience, deep expertise, and radical candor enriched this project beyond measure. Jessica Simon at the Association of Maternal and Child Health Programs (AMCHP) and Hiba Fatima and Deborah Baron at the University of North Carolina contributed outstanding session facilitation and assistance. Finally, we deeply appreciate Katie Grimes and Talia Kieu for invaluable qualitative expertise, which added depth and clarity to findings.

Data Availability

The datasets used and/or analyzed during the current study, excluding qualitative responses, are available from the corresponding author in this published article.

References

1. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. *BMC Psychol*. 2015;3(1):32. 10.1186/s40359-015-0089-9.
2. Lobb R, Colditz GA. Implementation science and its application to population health. *Annu Rev Public Health*. 2013;34:235–51. 10.1146/annurev-publhealth-031912-114444.
3. Luke DA, Powell BJ, Paniagua-Avila A. Bridges and mechanisms: Integrating systems science thinking into implementation research. *Annu Rev Public Health*. 2024;45(1):7–25. 10.1146/annurev-publhealth-060922-040205.
4. Bauer MS, Kirchner J. Implementation science: What is it and why should I care? *Psychiatry Res*. 2020;283:112376. 10.1016/j.psychres.2019.04.025.
5. Cruden G, Powell BJ, Frerichs L, Lanier P, Brown CH, Saldana L, et al. Leveraging group model building to operationalize implementation strategies across implementation phases: an exemplar related to child maltreatment intervention selection. *Implement Sci Commun*. 2024;5(1):134. 10.1186/s43058-024-00660-2.
6. Estrada-Magbanua WM, Huang TTK, Lounsbury DW, Zito P, Iftikhar P, El-Bassel N, et al. Application of group model building in implementation research: A systematic review of the public health and healthcare literature. *PLoS ONE*. 2023;18(8):e0284765. 10.1371/journal.pone.0284765.
7. Zimmerman L, Lounsbury DW, Rosen CS, Kimerling R, Trafton JA, Lindley SE. Participatory system dynamics modeling: Increasing stakeholder engagement and precision to improve implementation planning in systems. *Adm Policy Ment Health*. 2016;43(6):834–49. 10.1007/s10488-016-0754-1.
8. Caton L, Zimmerman L, Kahkoska A, Goldstein BA, Sperber N. Participatory system dynamics in implementation science practice: a scoping review of methods, contexts and outcomes. *BMJ Open Qual*. 2026;15(2):e003919. 10.1136/bmjopen-2025-003919.
9. Sperber NR, Zimmerman LE, Caton L, Rust T, Bedoya A, Haas S, et al. Engaging hospital staff to identify levers for adoption of clinical decision support: Protocol for a single-site case study using system dynamics group model building. *JMIR Res Protocols*. 2026;15(1):e80848. 10.2196/80848.
10. Rouwette EAJA, Vennix JAM, Mullekom TV. Group model building effectiveness: a review of assessment studies. *Syst Dyn Rev*. 2002;18(1):5–45. 10.1002/sdr.229.
11. Richardson GP. Reflections on the foundations of system dynamics. *Syst Dyn Rev*. 2011;27(3):219–43. 10.1002/sdr.462.
12. Cilenti D, Issel M, Wells R, Link S, Lich KH. System dynamics approaches and collective action for community health: An integrative review. *Am J Comm Psychol*. 2019;63(3–4):527–45. 10.1002/ajcp.12305.
13. Gwynn I, Simon J, Anderson S, Klamen SL, Mullenix A, Cilenti D, Hassmiller Lich K. Tools for supporting the MCH workforce in addressing complex challenges: A scoping review of system dynamics modeling in maternal and child health. *Mat Child Health J*. 2022;26(Suppl 1):176–203. 10.1007/s10995-022-03376-8.
14. Hovmand PS. Introduction to community-based system dynamics. In: Hovmand PS, editor. *Community based system dynamics*. New York, NY: Springer; 2014. pp. 1–16. [cited 2025 Aug 18]. 10.1007/978-1-4614-8763-0_1.
15. Sterman J. *Business Dynamics*. Irwin/McGraw-Hill; 2010.

16. Grebing EM, Edmunds JA, Arshavsky NP. The relationship between buy-in and implementation: Measuring teacher buy-in to a high school reform effort. *Eval Program Plann.* 2023;97:102224. 10.1016/j.evalprogplan.2023.102224.
17. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q.* 2004;82(4):581–629. 10.1111/j.0887-378X.2004.00325.x.
18. Hohl SD, Bird JE, Nguyen CVT, D'Angelo H, Minion M, Pauk D, et al. Operationalizing leadership and clinician buy-in to implement evidence-based tobacco treatment programs in routine oncology care: A mixed-method study of the U.S. Cancer Center Cessation Initiative. *Curr Onc.* 2022;29(4):2406–21. 10.3390/curroncol29040195.
19. O'Malley D, Hudson SV, Nekhlyudov L, Howard J, Rubinstein E, Lee HS et al. Learning the landscape: implementation challenges of primary care innovators around cancer survivorship care. *J Cancer Surv.* 2017;11:13–23. doi:0.1007/s11764-016-0555-2.
20. Walsh-Bailey C, Tsai E, Tabak RG, Morshed AB, Norton WE, McKay VR, et al. A scoping review of de-implementation frameworks and models. *Implement Sci.* 2021;16(1):100. 10.1186/s13012-021-01173-5.
21. Dunsmore J, Duncan E, MacLennan S, N'Dow J, MacLennan S. Effectiveness of de-implementation strategies for low-value prescribing in secondary care: a systematic review. *Implement Sci Commun.* 2023;4(1):115. 10.1186/s43058-023-00498-0.
22. Whelan J, Fraser P, Bolton KA, Love P, Strugnelli C, Boelsen-Robinson T, et al. Combining systems thinking approaches and implementation science constructs within community-based prevention: a systematic review. *Health Res Policy Sys.* 2023;21(1):85. 10.1186/s12961-023-01023-4.
23. Nilsen P, Ingvarsson S, Hasson H, von Thiele Schwarz U, Augustsson H. Theories, models, and frameworks for de-implementation of low-value care: A scoping review of the literature. *Implement Res Pract.* 2020;1:2633489520953762. 10.1177/2633489520953762.
24. Cary MA, Plamondon K, Banner-Lukaris D, Oelke N, Sibley KM, Baxter K, et al. Building consensus in research partnerships: a scoping review of consensus methods. *Evid Policy.* 2023;19(3):485–511. 10.1332/174426421X16645354235140.
25. Djulbegovic B, Guyatt G. Evidence vs consensus in clinical practice guidelines. *JAMA.* 2019;322(8):725–6. 10.1001/jama.2019.9751.
26. Smith NR, Levy DE, Falbe J, Purtle J, Chiqui JF. Design considerations for developing measures of policy implementation in quantitative evaluations of public health policy. *Front Health Serv.* 2024;4. 10.3389/frhs.2024.1322702.
27. McKay VR, Morshed AB, Brownson RC, Proctor EK, Prusaczyk B. Letting Go: Conceptualizing intervention de-implementation in public health and social service settings. *Am J Comm Psychol.* 2018;62(1–2):189–202. 10.1002/ajcp.12258.
28. Raudasoja AJ, Falkenbach P, Vernooij RWM, Mustonen JMJ, Agarwal A, Aoki Y, et al. Randomized controlled trials in de-implementation research: a systematic scoping review. *Implement Sci.* 2022;17(1):65. 10.1186/s13012-022-01238-z.
29. Augustsson H, Ingvarsson S, Nilsen P, von Thiele Schwarz U, Muli I, Dervish J, et al. Determinants for the use and de-implementation of low-value care in health care: a scoping review. *Implement Sci Commun.* 2021;2(1):13. 10.1186/s43058-021-00110-3.
30. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci.* 2015;10(1):53. 10.1186/s13012-015-0242-0.
31. Helfrich CD, Rose AJ, Hartmann CW, van Bodegom-Vos L, Graham ID, Wood SJ, et al. How the dual process model of human cognition can inform efforts to de-implement ineffective and harmful clinical practices: A preliminary model of unlearning and substitution. *J Eval Clin Pract.* 2018;24(1):198–205. 10.1111/jep.12855.
32. Bridges KM. Race, Pregnancy, and the Opioid Epidemic. *Harv Law Rev.* 2020;133(3):770–851.
33. Bridges KM. White privilege and white disadvantage. *Va L Rev.* 2019;105:449.
34. Hansen H, Netherland J, Herzberg D, Whiteout. How racial capitalism changed the color of opioids in America. Oakland; Univ of California; 2023.
35. deLeon P. Public policy termination: An end and a beginning. *Policy Anal.* 1978;4(3):369–92. 10.1093/oxfordjournals.jpart.a003544.
36. deLeon P, deLeon L. What ever happened to policy implementation? An alternative approach. *J Pub Admin Res Theor.* 2002;12(4):467–92. 10.1093/oxfordjournals.jpart.a003544.
37. Ingvarsson S, Hasson H, von Thiele Schwarz U, Nilsen P, Powell BJ, Lindberg C, et al. Strategies for de-implementation of low-value care—a scoping review. *Implement Sci.* 2022;17(1):73. 10.1186/s13012-022-01247-y.
38. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care.* 2007;19(6):349–57.
39. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. *Qual Health Res.* 2019;29(10):1483–96. 10.1177/1049732318821692.
40. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Soc Sci Med.* 2022;292:114523. 10.1016/j.socscimed.2021.114523.
41. Zimmermann NS, Black LJ, Shrubsole C, Davies M. Meaning-making in the process of participatory system dynamics research. In: Proceedings of the 33rd International Conference of the System Dynamics Society* Cambridge, Massachusetts, USA–July 19–23, 2015.
42. Kim D. Behavior over time diagrams: Seeing dynamic interrelationships. *The Systems Thinker* [Internet]. 2016 Jan 13 [cited 2026 Feb 11]. Available from: <https://thesystemsthinker.com/behavior-over-time-diagrams-seeing-dynamic-interrelationships/>
43. Meadows DH. Thinking in systems: A primer. Junction. VT:Chelsea Green Publishing; 2008.
44. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. *Implement Sci.* 2017;12(1):77. 10.1186/s13012-017-0605-9.
45. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci.* 2012;7(1):37. 10.1186/1748-5908-7-37.

46. Neuman W. Social research methods: Qualitative and quantitative approaches. 3rd Edition. Boston, MA: Allyn & Bacon; 1997.
47. Corbin J, Strauss A. Basics of qualitative research: Techniques and procedures for developing grounded theory. SAGE; 2014.
48. Brown KK, Lemke MK, Fallah-Fini S, Hall A, Obasanya M. Planning, implementing, and evaluating an online group-model-building workshop during the COVID-19 pandemic: celebrating successes and learning from shortcomings. *Syst Dyn Rev.* 2022;38(1):93–112. 10.1002/sdr.1704.
49. Calancie L, Fullerton K, Appel JM, Korn AR, Hennessy E, Hovmand P, et al. Implementing group model building with the Shape Up Under 5 community committee working to prevent early childhood obesity in Somerville. *Mass JPHMP.* 2022;28(1):E43–55. 10.5888/pcd19.210181.
50. Egbuonye NC, Sesterhenn L, Goldman B, Pikora J, Parmater J. Bridging the gap between transformative practices and traditional public health approaches. *JPHMP.* 2022;28(Supplement 4):S151–8. 10.1097/PHH.0000000000001502.
51. Felmingham T, Bolton KA, Fraser P, Allender S, Brown AD. Measuring shifts in mental models in the prevention of childhood obesity in rural Australia. *Health Educ Behav.* 2023;50(5):662–70. 10.1177/10901981231165339.
52. Fokkinga B, Bleijenbergh I, Vennix JAM. Group Model building evaluation in single cases: a method to assess changes in mental models. In: *Proceedings of the 27th International Conference of the System Dynamics Society.* 2009.
53. Scott RJ, Cavana RY, Cameron D. Evaluating immediate and long-term impacts of qualitative group model building workshops on participants' mental models. *Syst Dyn Rev.* 2013;29(4):216–36. 10.1002/sdr.1505.
54. Wensing M, Grol R. Knowledge translation in health: how implementation science could contribute more. *BMC Med.* 2019;17(1):88. 10.1186/s12916-019-1322-9.
55. Frerichs L, Young TL, Dave G, Stith D, Corbie-Smith G, Hassmiller Lich K. Mind maps and network analysis to evaluate conceptualization of complex issues: A case example evaluating systems science workshops for childhood obesity prevention. *Eval Program Plann.* 2018;68:135–47. 10.1016/j.evalprogplan.2018.03.003.
56. Faherty LJ, Kranz AM, Russell-Fritch J, Patrick SW, Cantor J, Stein BD. Association of punitive and reporting state policies related to substance use in pregnancy with rates of Neonatal Abstinence Syndrome. *JAMA Netw Open.* 2019;2(11):e1914078. 10.1001/jamanetworkopen.2019.14078.
57. Faherty LJ, Stein BD, Terplan M. Consensus guidelines and state policies: The gap between principle and practice at the intersection of substance use and pregnancy. *Am J Obstet Gynecol MFM.* 2020;2(3):100137. 10.1016/j.ajogmf.2020.100137.
58. Faherty LJ, Heins S, Kranz AM, Patrick SW, Stein BD. Association between punitive policies and neonatal abstinence syndrome among Medicaid-insured infants in complex policy environments. *Addiction.* 2022;117(1):162–71. 10.1111/add.15602.
59. Thomas KA, Struble CA, Stenersen MR, Moore K. The impact of state-level prenatal substance use policies on rates of maternal and infant mortality in the United States: A legal epidemiology Study. 2022 Nov 18. 10.1101/2022.11.16.22282429
60. Onwuegbuzie AJ, Dickinson WB, Leech NL, Zoran AG. Toward more rigor in focus group research: A new framework for collecting and analyzing focus group data. *Int J Qual Methods.* 2009;8(3):1–21.
61. Onwuegbuzie AJ, Dickinson WB, Leech NL, Zoran AG. A qualitative framework for collecting and analyzing data in focus group research. *Int J Qual Methods.* 2009;8(3):1–21. 10.1177/160940690900800301.
62. Turnbull B. Teacher participation and buy-in: Implications for school reform initiatives. *Learn Environ Res.* 2002;5(3):235–52. 10.1023/A:1021981622041.
63. Brennan LK, Sabounchi NS, Kemner AL, Hovmand P. Systems thinking in 49 communities related to healthy eating, active living, and childhood obesity. *J Pub Health Manag Pract.* 2015;21:S55. 10.1097/PHH.0000000000000248.
64. Upvall MJ, Bourgault AM. De-implementation: A concept analysis. *Nurs Forum.* 2018;53(3):376–82. 10.1111/nuf.12256.
65. McCourt AD, White SA, Bandara S, Schall T, Goodman DJ, Patel E, et al. Development and Implementation of state and federal child welfare laws related to drug use in pregnancy. *Milbank Q.* 2022;100(4):1076–120. 10.1111/1468-0009.12591.
66. White SA, McCourt A, Bandara S, Goodman DJ, Patel E, McGinty EE. Implementation of state laws giving pregnant people priority access to drug treatment programs in the context of coexisting punitive laws. *Women's Health Issues.* 2022;33(2). 10.1016/j.whi.2022.09.001.
67. Desai A, DeMatteo D, Heilbrun K, Holliday R, Lankford C, Rotrosen J. Public perception on policies to address prenatal substance use: Recommendations regarding maternal criminal prosecution and child welfare. *PPPL.* 2023;29(3):402–15. 10.1037/law0000387.
68. Hovmand P, Rouwette E, Andersen D, Richardson G, Calhoun A. Scriptapedia: a handbook of scripts for developing structured group model building sessions [Internet]. 2011. Available from: <https://repository.ubn.ru.nl/bitstream/handle/2066/95406/95406.pdf>
69. Crable EL, Lengnick-Hall R, Stadnick NA, Moullin JC, Aarons GA. Where is policy in dissemination and implementation science? Recommendations to advance theories, models, and frameworks: EPIS as a case example. *Implement Sci.* 2022;17(1):80. 10.1186/s13012-022-01256-x.
70. Moullin JC, Dickson KS, Stadnick NA, Rabin B, Aarons GA. Systematic review of the Exploration, Preparation, Implementation, Sustainment (EPIS) framework. *Implement Sci.* 2019;14(1):1. 10.1186/s13012-018-0842-6.
71. Bandara S, Schall TE, White SA, McCourt AD, Goodman D, McGinty EE. Association between U.S. state prenatal drug use laws and child welfare reporting in Alabama, Maryland and Utah. *Int J Drug Policy.* 2022;108:103806. 10.1016/j.drugpo.2022.103806.
72. Maclean JC, Witman A, Durrance CP, Atkins DN, Meinhofer A. Prenatal substance use policies and infant maltreatment reports. *Health Aff (Millwood).* 2022;41(5):703–12. 10.1377/hlthaff.2021.01755.
73. Austin AE, Naumann RB, Simmons E. Association of state child abuse policies and mandated reporting policies with prenatal and postpartum care among women who engaged in substance use during pregnancy. *JAMA Pediatr.* 2022;176(11):1123–30. 10.1001/jamapediatrics.2022.3396.
74. Patrick SW, Schiff DM, COMMITTEE ON SUBSTANCE USE AND PREVENTION, Ryan SA, Quigley J, Gonzalez PK, et al. A Public Health Response to Opioid Use in Pregnancy. *Peds.* 2017;139(3):e20164070. 10.1542/peds.2016-4070.

75. Cavanagh AJ, Aragón OR, Chen X, Couch BA, Durham MF, Bobrownicki A, et al. Student buy-in to active learning in a college science course. *CBE Life Sci Educ.* 2016;15(4):ar76. 10.1187/cbe.16-07-0212.
76. Durlak JA, DuPre EP. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *AJCP.* 2008;41:327–50.
77. Coleman T, Antoniak M, Britton J, Thornton J, Lewis S, Watts K. Recruiting pregnant smokers for a placebo-randomised controlled trial of nicotine replacement therapy. *BMC Health Serv Res.* 2004;4(1):29. 10.1186/1472-6963-4-29.
78. Malshe A, Sohi RS. Sales buy-in of marketing strategies: Exploration of its nuances, antecedents, and contextual conditions. *J Pers Sell Sales Manag.* 2009;29(3):207–25. 10.2753/PSS0885-3134290301.
79. Meyers SA, Earnshaw VA, D'Ambrosio B, Courchesne N, Werb D, Smith LR. The intersection of gender and drug use-related stigma: A mixed methods systematic review and synthesis of the literature. *Drug Alcohol Depend.* 2021;223:108706. 10.1016/j.drugalcdep.2021.108706.
80. Ghaffarzadegan N, Lyneis J, Richardson GP. How small system dynamics models can help the public policy process. *Syst Dyn Rev.* 2011;27(1):22–44. 10.1002/sdr.442.
81. Grant A. Beyond evidence-based policy in public health: The interplay of ideas of ideas. *Soc Policy Adm.* 2015;49(5):672–4. 10.1111/spol.12159.
82. Astbury CC, McGill E, Egan M, Penney TL. Systems thinking and complexity science methods and the policy process in non-communicable disease prevention. *Int J Health Policy Manag.* 2021;12:6772. 10.1136/bmjopen-2021-049878.
83. Carey G, Malbon E, Carey N, Joyce A, Crammond B, Carey A. Systems science and systems thinking for public health: a systematic review of the field. *BMJ Open.* 2015;5(12):e009002. 10.1136/bmjopen-2015-009002.
84. Lich KH, Ginexi EM, Osgood ND, Mabry PL. A call to address complexity in prevention science research. *Prev Sci.* 2013;14(3):279–89. 10.1007/s1121-012-0285-2.
85. Luna-Reyes LF, Andersen DL. Collecting and analyzing qualitative data for system dynamics: methods and models. *Syst Dyn Rev.* 2003;19(4):271–96. 10.1002/sdr.280.
86. Bransby DP, Mayo AT, Cronin MA, Park K, Yuan CT. A systematic review of respect between acute care nurses and physicians. *Health Care Manag Rev.* 2023;48(3):237–48.
87. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual Quant.* 2013;47(4):2025–47. 10.1007/s11135-011-9640-9.
88. Abshire M, Dinglas VD, Cajita MIA, Eakin MN, Needham DM, Himmelfarb CD. Participant retention practices in longitudinal clinical research studies with high retention rates. *BMC Med Res Methodol.* 2017;17(1):30. 10.1186/s12874-017-0310-z.

Figures

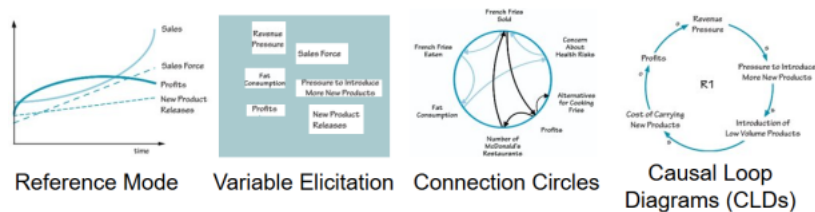


Figure 1

Sample session artifacts by PSD stage (42)

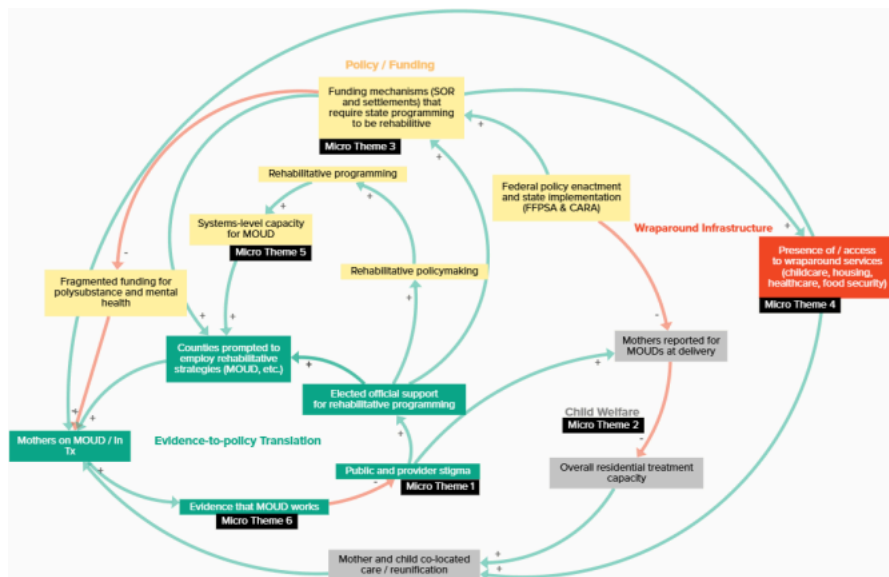
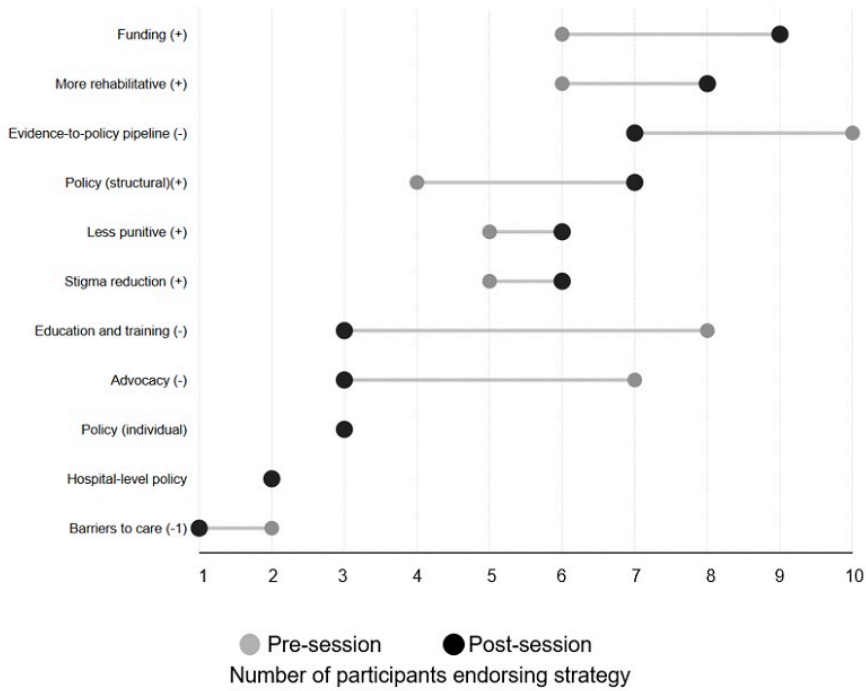


Figure 2

Causal Loop Diagram (CLD) with Core Variables from Both States

North Carolina
Pre (n=47) vs. Post (n=44)



Kentucky
Pre (n=42) vs. Post (n=41)

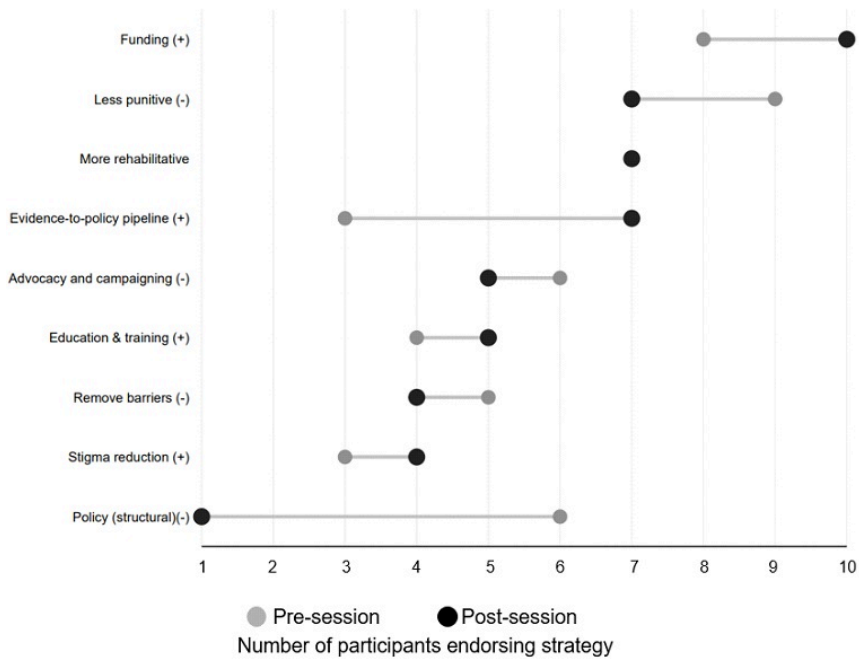
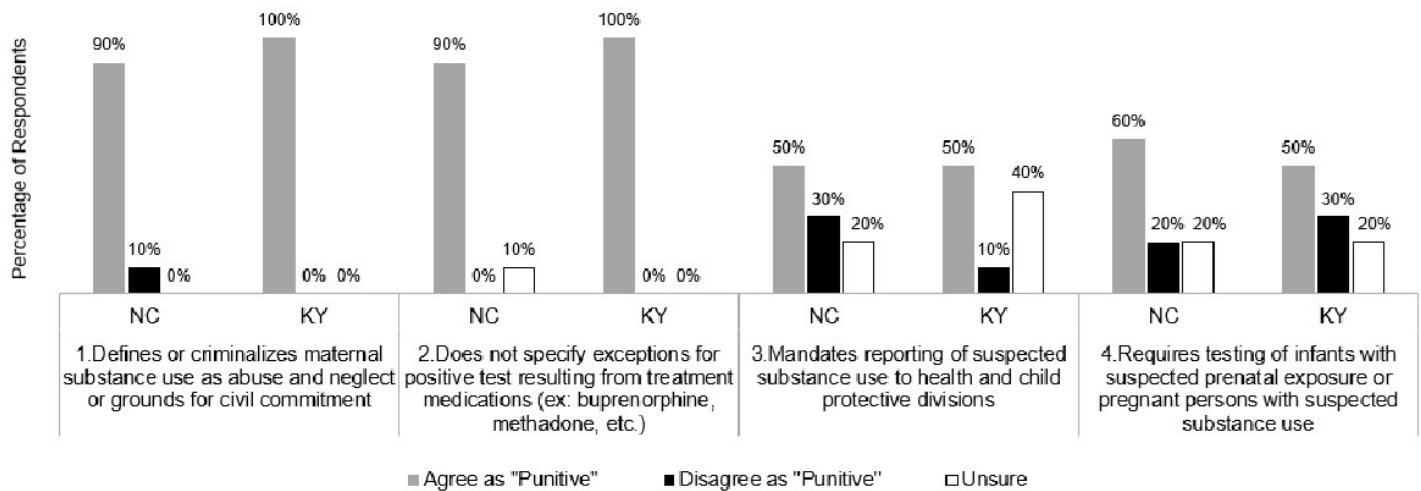


Figure 3

Pre-Post Shifts in Action Item Prioritization (55)

Punitive Policy Consensus by State and Policy NC (n=10), KY (n=10)



Rehabilitative Policy Consensus by State and Policy NC (n=10), KY (n=10)

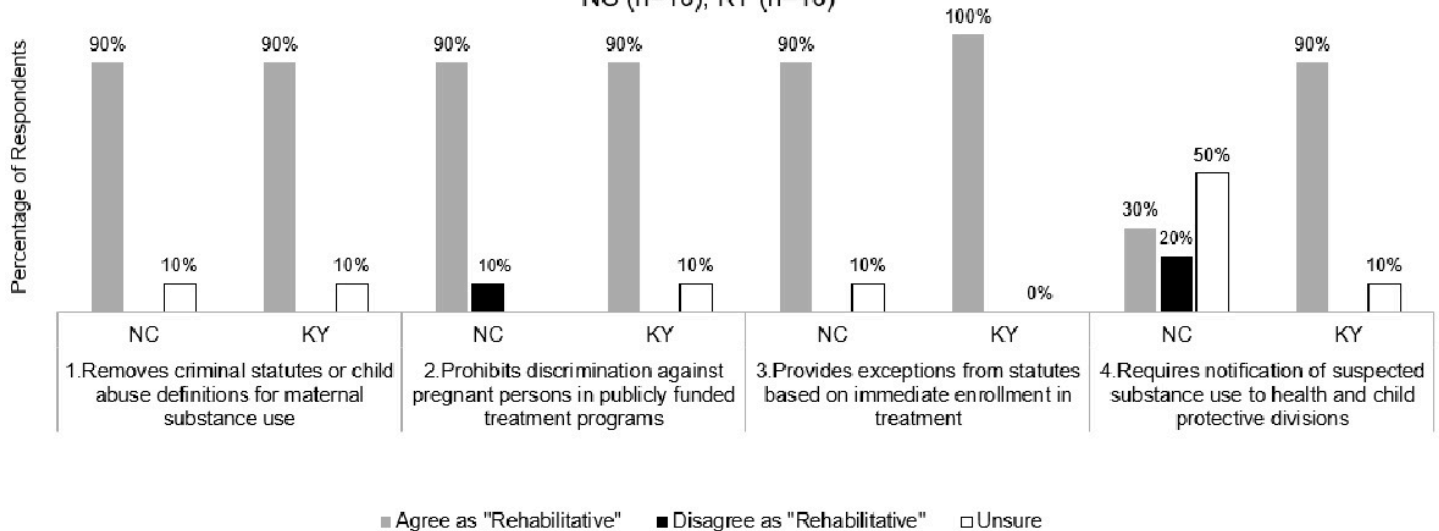


Figure 4

Punitive and Rehabilitative Consensus by State and Policy

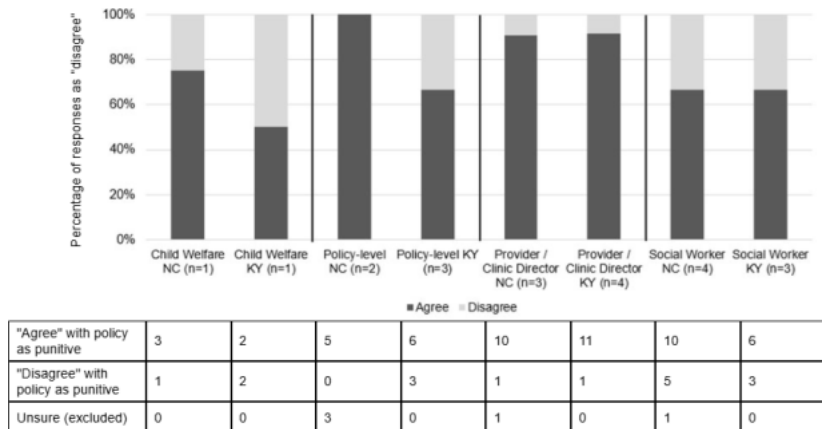


Figure 5

Percentage of Disagree Responses (for Punitive Policies) by Role

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementalFiles3.docx](#)