SUPPLEMENTARY MATERIAL
Identifying General Mechanism Shifts in Linear Causal Representations

A Limitations and Broader Impacts

Limitations of this work include the need to relax the noise assumption and to consider similar settings under nonlinear mixing functions. These are promising directions to explore in the CRL field. The broader impact of this work is that CRL methods can be used to identify mechanism shifts and determine root causes, which can be utilized in the biological field to find disease genes or biomarkers. Currently, the negative impacts of this method are not clear.

B Illustration of our algorithm
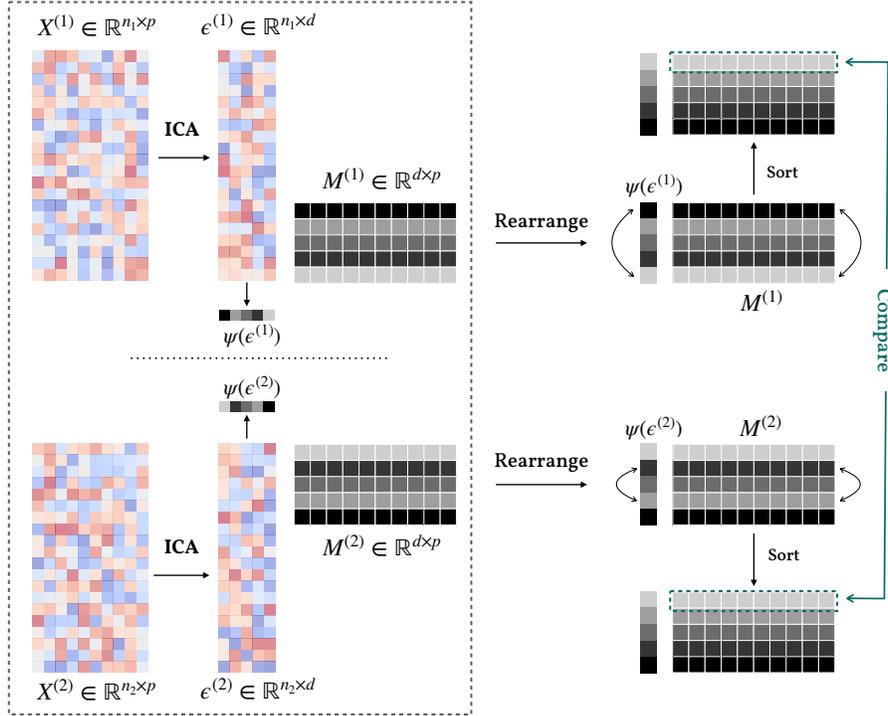
Figure 4: Overview of our method: For each context k, given the data $\boldsymbol{X}^{(k)}$, our method involves three main steps. First, we apply ICA to each dataset to estimate $\boldsymbol{\epsilon}^{(k)}$ and $M^{(k)}$. Second, we calculate $\psi(\epsilon^{(k)}) = \{\psi(\epsilon_1^{(k)}), \psi(\epsilon_2^{(k)}), \ldots, \psi(\epsilon_d^{(k)})\}$ for each noise component, sort these components in increasing order, and correspondingly arrange the rows of $M^{(k)}$. Third, we compare the sorted rows of $M^{(k)}$ to identify the shifted nodes.

C Discussion on Test Function

In Assumption B, we assume that there exists a test function ψ and that we can access it. Here we discuss ways to relax it. Recall that in Section 4.1, ψ is utilized to sort the noise component $\bar{\epsilon}^{(k)}$ to ensure that the post-sorting noise vector $\widetilde{\epsilon}^{(k)}$ has a consistent order across all environments.

An alternative approach to achieve this is to use distribution matching. We take the noise vector in the first environment as a reference and align all other noise vectors post-sorting with the reference vector. To do this, we can use a distribution distance metric D. First, define a signed permutation space as

$$S_d = \{S = PD \mid P \text{ is a permutation matrix}, D \text{ is a diagonal matrix with } D_{ii} \in \{-1, 1\}\}$$

Then, solve the optimization problem:

$$\min_{S \in S_d} D(\bar{\epsilon}^{(1)}, S\bar{\epsilon}^{(k)})$$

where D can be any distribution distance, such as Kullback-Leibler divergence. In Assumption A, we assume pair-wise different noise component, thus the optimization questions have minimums value 0 if and only if each noise component of $\bar{\epsilon}^{(1)}$ and $S\bar{\epsilon}^{(k)}$ have the same distribution, thus help us align the noise component order. We solve this optimization problem for each environment $k \geq 2$, thus obtaining $\overline{P}^{(k)}$. All following steps in our algorithm remain the same when using this alternative approach.

One small gap remains: even though all post-sorting noise vectors have a consistent order with $\bar{\epsilon}^{(1)}$, $\bar{\epsilon}^{(1)}$ is not the ground truth order of $\epsilon^{(1)}$. This ambiguity cannot be eliminated, consistent with the nature of the CRL method, and is the same with other CRL methods, such as [40, 28]. Fortunately, the ground truth order is not so important in practice. What people mainly care about is the semantic label for each latent node. Some CRL generative models, such as [54], may be helpful for performing fake interventions and manually assigning semantic labels. However, this is beyond the scope of this paper, and we will not discuss it further.

Even though the distribution matching optimization method offers greater flexibility, it is computationally expensive. First, note that the cardinality $|S_d| = d! \times 2^d$, which represents a vast search space when d is large. Furthermore, calculating $D(\cdot, \cdot)$ is generally computationally intensive. For example, the KL method requires density estimation, and the Maximum Mean Discrepancy (MMD) method necessitates the computation of pairwise distances among samples. These challenges render this alternative difficult to implement. Consequently, we opt to use the ψ function to facilitate efficient sorting, but it may need carefully design.

D Discussion on Sample Complexity

The sample complexity of our method must be considered from two perspectives: one involves using ICA to estimate $\bar{\epsilon}^{(k)}$ and $\overline{M}^{(k)}$, and the other pertains to utilizing $\bar{\epsilon}^{(k)}$ and a test function to sort the rows of $\overline{M}^{(k)}$. Since the sorting step depends on the choice of test function, we assume for simplicity that $\overline{M}^{(k)}$ is already sorted by the ground truth order. Thus, we only focus on the asymptotic behavior of $\widetilde{M}^{(k)}$, which closely relates to the properties of the ICA estimator.

There are various algorithms for solving ICA [18, 17, 41]; each algorithm exhibits different asymptotic statistical properties. If we apply the findings in Auddy and Yuan [2], we assume that the estimated ICA unmixing function has the following statistical accuracy:

Theorem 4. *If the sample size $n \geq g(d, \delta)$, then with probability at least $1 - h(n, d, \delta, \epsilon)$, we have:*

$$l(\widetilde{M}_i^{(k)} - M_i^{(k)}) \leq C \cdot p(d, n) f(\delta),$$

where $\widetilde{M}_i^{(k)}$ represents the i-th row of the estimated unmixing function $M^{(k)}$, C is a constant, and p, f, g, and h are known functions. For instance, in Auddy and Yuan [2], $p(d, n) = \sqrt{d/n}$ and $f(\delta) = \sqrt{\log(1/\delta)}$. Here, l denotes the loss function, and the L_2 norm can serve as an option.

Under this theorem, for two environments k and k', if node i does not shift, we have:

$$||\widetilde{M}_i^k - \widetilde{M}_i^{k'}||_2 \leq ||\widetilde{M}_i^k - M_i||_2 + ||\widetilde{M}_i^{k'} - M_i||_2 \leq 2 \cdot C \cdot p(d, n) f(\delta)$$

with a probability of at least $1 - 2h(n, d, \delta, \epsilon)$. Thus, by setting the threshold α as $2 \cdot C \cdot p(d, n) f(\delta)$, we can control the false discovery rate to be at most $2h(n, d, \delta, \epsilon)$. A similar sample complexity theorem can be extended to cases involving more than two environments, as long as the statistical properties of the ICA solution are known.

14

E Detailed Proofs

E.1 Proof of Proposition 2

Lemma 1. *Under problem setting, for any $x, y \in \mathbb{R}^{d \times 1}$, the equation $x^T H = y^T H$ holds if and only if $x = y$.*

Proof. Given that G possesses full column rank, it follows that $H = G^\dagger$ has full row rank. Consequently, the null space of H^T is $\{0\}$. Therefore, if $x^T H = y^T H$, it implies $H^T(x - y) = 0$. This leads to the conclusion that $x - y = 0$, which in turn implies $x = y$. $\square$

Proof of Proposition 2. Recall that $B^{(k)} = (\Omega^k)^{-\frac{1}{2}}(I_d - A^{(k)})$. Since $A^{(k)}$ is a weighted adjacency matrix, its diagonal entries are zero. Thus,

$$B_{ij}^{(k)} = -\left(\Omega_{ii}^{(k)}\right)^{-\frac{1}{2}} A_{ij}^{(k)} \quad \text{if} \quad i \neq j,$$
$$B_{ii}^{(k)} = \left(\Omega_{ii}^{(k)}\right)^{-\frac{1}{2}} \quad \text{if} \quad i = j.$$

Under Definition 1, if node Z_i is shifted, it implies either 1) $\Omega_{ii}^{(k)} \neq \Omega_{ii}^{(k')}$, 2) $A_i^{(k)} \neq A_i^{(k')}$, or 3) both conditions hold. In scenarios 1) and 3), $B_{ii}^{(k)} \neq B_{ii}^{(k')}$, resulting in $B_i^{(k)} \neq B_i^{(k')}$. In scenario 2), while $\Omega_{ii}^{(k)} = \Omega_{ii}^{(k')}$, there exists a $j \in [d]$ such that $A_{ij}^{(k)} \neq A_{ij}^{(k')}$, leading to $B_i^{(k)} \neq B_i^{(k')}$. If node Z_i is not shifted, then $A_i^{(k)} = A_i^{(k')}$ and $\Omega_{ii}^{(k)} = \Omega_{ii}^{(k')}$, implying $B_i^{(k)} = B_i^{(k')}$. Therefore, Z_i is shifted if and only if $B_i^{(k)} \neq B_i^{(k')}$. According to Lemma 1, $B_i^{(k)} \neq B_i^{(k')}$ if and only if $B_i^{(k)} H \neq B_i^{(k')} H$, which is equivalent to $M_i^{(k)} \neq M_i^{(k')}$.

In conclusion, Z_i is shifted if and only if $M_i^{(k)} \neq M_i^{(k')}$. $\square$

E.2 Proof of Theorem 3

Lemma 2. *Under problem setting, it is not possible for an intervention on the latent node Z_i to result in $M_i^{(k)} = -M_i^{(k')}$.*

Proof. We prove this by contradiction. Suppose that $M_i^{(k)} = -M_i^{(k')}$. According to Lemma 1, this would imply $B_i^{(k)} = -B_i^{(k')}$. However, we know $B^{(k)} = (\Omega^{(k)})^{-1}(I_d - A^{(k)})$ where $A^{(k)}$ is the weight matrix for a DAG. Since $A_{ii}^{(k)} = 0$, it follows that $B_{ii}^{(k)} = (\Omega^{(k)})_{ii}^{-1}$. Therefore, both $B_{ii}^{(k)}$ and $B_{ii}^{(k')}$ are positive. It is impossible for $B_i^{(k)}$ to be equal to $-B_i^{(k')}$. Consequently, the scenario where $M_i^{(k)} = -M_i^{(k')}$ cannot occur. $\square$

Proof of Theorem 3. Recall from the data generation process that

$$M^{(k)} X^{(k)} = \epsilon^{(k)}.$$

When input $X^{(k)}$ to ICA, we have $\overline{M}^{(k)} = P^{(k)} D^{(k)} M^{(k)}$ and $\bar{\epsilon}^{(k)} = P^{(k)} D^{(k)} \epsilon^{(k)}$. Without loss of generality, we assume that $\epsilon^{(k)}$ is ordered increasingly with respect to ψ. Thus, post sorting with respect to ψ, we eliminate the ambiguity of $P^{(k)}$, and we get $\widetilde{M}^{(k)} = D^{(k)} M^{(k)}$ and $\widetilde{\epsilon}^{(k)} = D^{(k)} \epsilon^{(k)}$.

We are now ready to prove that Z_i is not shifted if and only if $\widetilde{M}^{(k)} = \pm\widetilde{M}^{(k')}$. This immediately implies that if Z_i is not shifted, then $M_i^{(k)} = M_i^{(k')}$, thus satisfying $\widetilde{M}^{(k)} = \pm\widetilde{M}^{(k')}$.

If $\widetilde{M}^{(k)} = \pm\widetilde{M}^{(k')}$, there are two cases: $M_i^{(k)} = M_i^{(k')}$ or $M_i^{(k)} = -M_i^{(k')}$. We prove in Lemma 2 that the scenario $M_i^{(k)} = -M_i^{(k')}$ cannot exist. The only surviving situation is $M_i^{(k)} = M_i^{(k')}$, which indicates that Z_i is not shifted.

$\square$

15

F Experiments on Synthetic Data Compared with DCI

As described in Section 5.1, instead of generating the mixing function G from $\text{Unif}[-0.25, 0.25]$, we set $G = I$, such that $X = Z$ and Z can be directly observed. In this setup, finding general interventions in linear causal representations reduces to identifying general interventions in linear SEM, a setting for which the existing method DCI [52] is designed. Table 2 presents the performance comparison between our method and DCI under these conditions, demonstrating that our method outperforms DCI in most cases.

Graph Type	d	Method	Precision	Recall	F1
ER 2	5	DCI	0.60	0.60	0.60
		Ours	0.80	0.80	0.80
	10	DCI	0.87	1.00	0.92
		Ours	1.00	1.00	1.00
	15	DCI	0.74	1.00	0.84
		Ours	0.66	1.00	0.78
ER 4	10	DCI	0.83	1.00	0.89
		Ours	1.00	1.00	1.00
	15	DCI	0.71	1.00	0.81
		Ours	0.62	0.93	0.73
SF 2	5	DCI	0.70	0.80	0.73
		Ours	1.00	1.00	1.00
	10	DCI	0.67	1.00	0.79
		Ours	1.00	1.00	1.00
	15	DCI	0.65	1.00	0.78
		Ours	0.70	0.93	0.78
SF 4	5	DCI	0.60	0.60	0.60
		Ours	0.80	0.80	0.80
	10	DCI	0.77	1.00	0.85
		Ours	1.00	1.00	1.00
	15	DCI	0.56	0.93	0.68
		Ours	0.67	1.00	0.79

Table 2: Comparison of Precision, Recall, and F1 scores for different graph types, d values, and methods between our method and DCI.

G Additional Information on Real Data

This section provides detailed information on the procedures employed in analyzing the real dataset.

Preprocessing The initial dataset comprised 19,719 observations, which can be downloaded from https://www.kaggle.com/datasets/lucasgreenwell/ocean-five-factor-personality-test-responses/data. In the preprocessing phase, any observation with a missing value in any column was excluded, leaving a total of 19,710 observations for further analysis. Subsequently, we applied max-min value normalization to the scores of each question, ensuring that all scores were normalized to fall within the range $[0, 1]$. This normalization step is crucial for achieving uniformity in the data scale, thereby facilitating accurate analysis and comparison across the dataset.

Labeling the Noise To derive meaningful psychological insights, it is crucial to assign semantic labels to all latent nodes. Given that the noise components are pairwise distinct and unique to the latent node Z_i, we consider intervening on each noise component, then remixing and observing the changes in the observational space. This approach enables us to assign semantic labels to both the noise components and their corresponding latent nodes. We utilize observations from the male dataset as the reference context, which comprises 7,603 observations. Following the initial step of our method, we obtain the sorted $\widetilde{M}^{male}$ and $\widetilde{\epsilon}^{male}$. The mixing function G is derived from $(\widetilde{M}^{male})^{\dagger}$.

16

To identify the semantic label for the first component of $\widetilde{\epsilon}$, we set its corresponding noise vector component to 0, effectively nullifying the first component of $\widetilde{\epsilon}^{male}$. This intervention yields an estimated noise matrix samples from $\widetilde{\epsilon}_{inv}^{male}$, denoted as $\widehat{\epsilon}_{inv}^{male}$. The intervened reconstruction, $\boldsymbol{X}_{inv}^{male} = G(\widehat{\epsilon}_{intv}^{male})^T$, and the original score distribution, $\boldsymbol{X}^{male} = G(\widehat{\epsilon}^{male})^T$, allow us to compare question scores pre- and post-intervention. Figure 7 plots these distributions, revealing significant shifts for questions pertaining to the *Agreeableness* dimension, with minimal impact on other scores, thereby identifying the first noise component as *Agreeableness*. This process is replicated for the second through fifth columns of ϵ^{male}, with results illustrated in Figures 9, 8, 5, and 6. Each plot demonstrates that interventions result in significant distribution changes for questions related to a single personality dimension, with negligible effects on others. Consequently, we label these noise components as *Openness*, *Conscientiousness*, *Extraversion*, and *Neuroticism*, respectively. These labels will be used for all the following analysis.

Shifted Nodes Detection We then applied our method to data from the male and female contexts. The calculated $L_i^{male,female}$ values are $\{0.074, 0.0497, 0.078, 0.638, 0.633\}$. Based on these results, we identify shifts in the last two personality dimensions, specifically labeled as *Extraversion* and *Neuroticism*. Additionally, we conducted a comparative analysis of personality dimensions between the US and UK, which have 8,753 and 1,531 observations, respectively. The computed $L_i^{US,UK}$ values are $\{0.302, 0.258, 0.109, 0.189, 0.088\}$, indicating that no latent node is considered as having undergone shifts between these two countries.
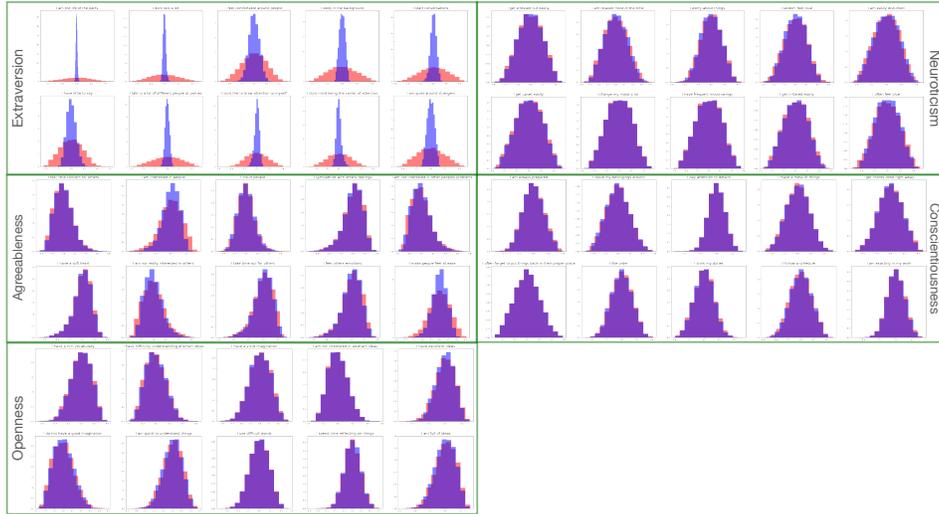
Figure 5: Intervention on the fourth component of the noise vector and subsequent re-mixing generate a new observed space — a new score distribution. Notably, only *Extraversion* exhibits significant changes after intervention, leading us to label the fourth component of the noise vector (after sorting) as *Extraversion*.
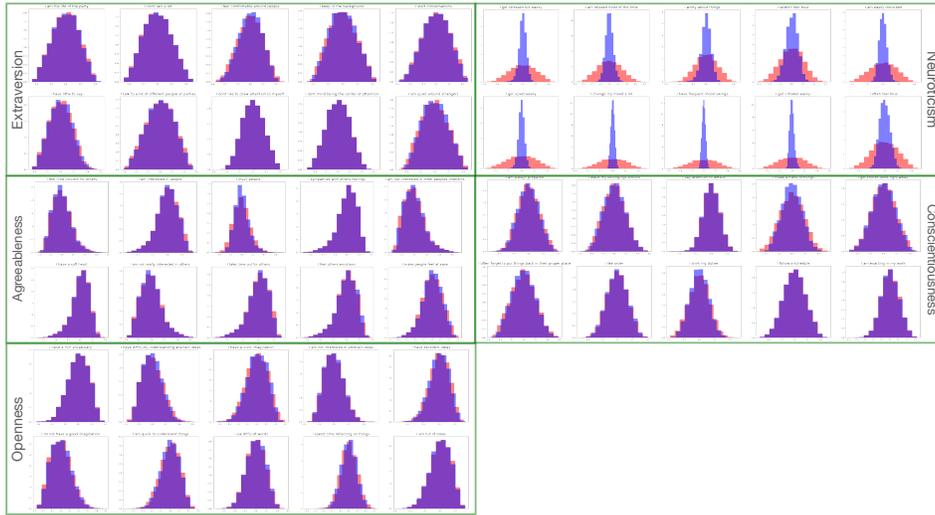
Figure 6: Intervention on the fifth component of the noise vector and subsequent re-mixing generate a new observed space — a new score distribution. Notably, only *Neuroticism* exhibits significant changes after intervention, leading us to label the fifth component of the noise vector (after sorting) as *Neuroticism*.
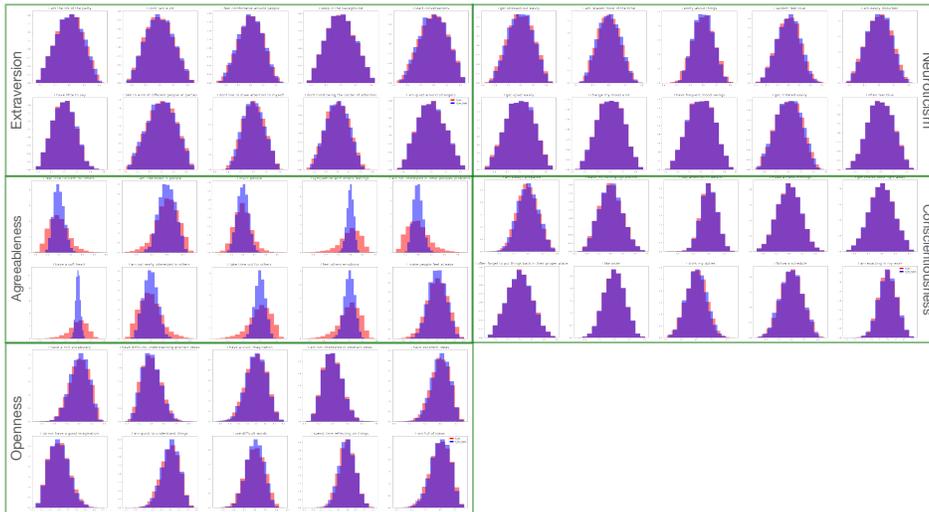
Figure 7: Intervention on the first component of the noise vector and subsequent re-mixing generate a new observed space — a new score distribution. Notably, only *Agreeableness* exhibits significant changes after intervention, leading us to label the first component of the noise vector (after sorting) as *Agreeableness*.

18
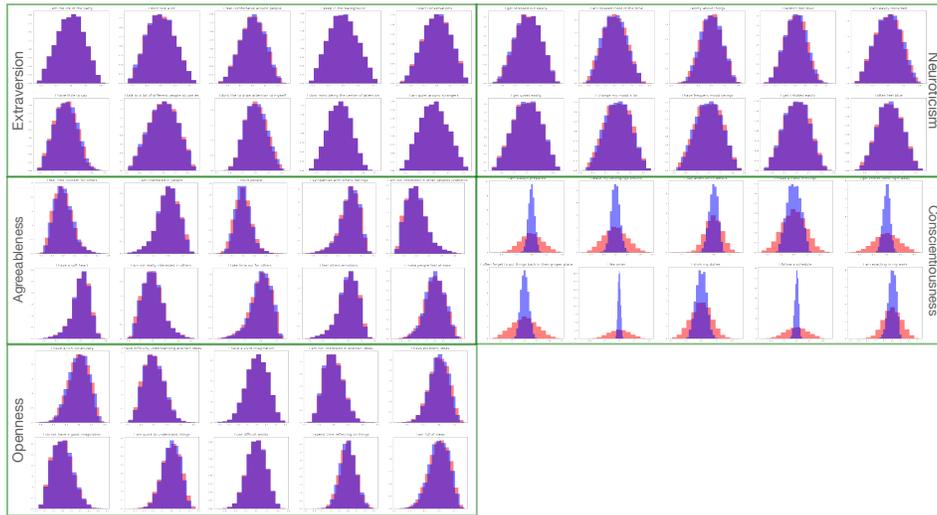
Figure 8: Intervention on the third component of the noise vector and subsequent re-mixing generate a new observed space — a new score distribution. Notably, only *Conscientiousness* exhibits significant changes after intervention, leading us to label the third component of the noise vector (after sorting) as *Conscientiousness*.
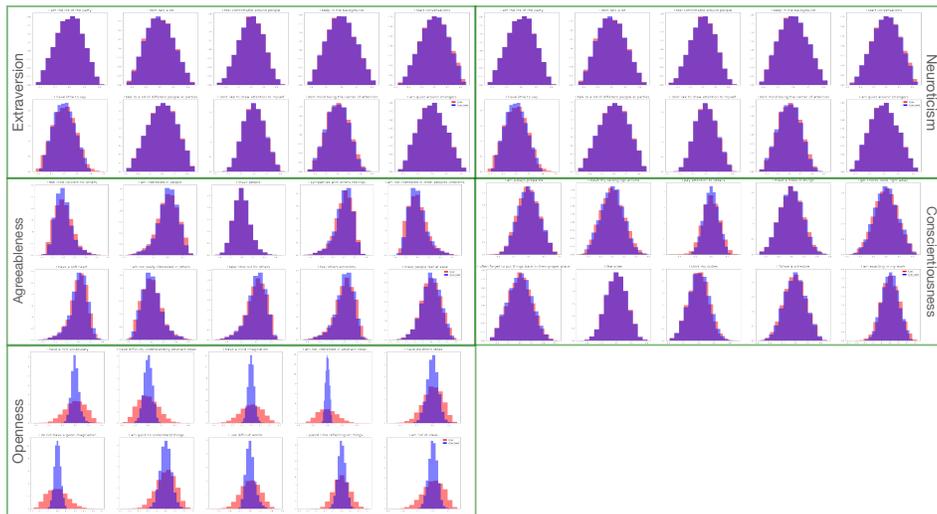
Figure 9: Intervention on the second component of the noise vector and subsequent re-mixing generate a new observed space — a new score distribution. Notably, only *Openness* exhibits significant changes after intervention, leading us to label the second component of the noise vector (after sorting) as *Openness*.

19

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The main claims made in the abstract and introduction accurately reflect the paper's contributions and scope.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: It has been discussed in Appendix A.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

20

Justification: Theorems and proposition has been completely proven in Appendix.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: The whole algorithm has been discussed in Algorithm 1.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

21

Answer: [Yes]

Justification: The code is published in `https://github.com/TianyuCodings/iLCS`.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: The algorithm is detailed in Algorithm 1, and the implementation details are discussed in Section 5.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [No]

Justification: We do not report statistical significance of the experiments as it follows the traditions of CRL literature, such as [40] and [4].

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)

22

- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [No]

   Justification: We do not report the computer resources used as it is not the primary concern of this paper.

   Guidelines:
   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics https://neurips.cc/public/EthicsGuidelines?

   Answer: [Yes]

   Justification: The research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics.

   Guidelines:
   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: It is discussed in Section A.

    Guidelines:
    - The answer NA means that there is no societal impact of the work performed.
    - If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
    - Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.

23

- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: The paper has no such risks.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: One dataset is generated by simulation, and the source of the real data has been explicitly mentioned.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, paperswithcode.com/datasets has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.

24

- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes]

    Justification: The primary new asset of the paper is the code, and the code is published in https://github.com/TianyuCodings/iLCS.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
    - We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
    - For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

25